NUMBER 37

STATISTICAL METHODS FOR THE ESTIMATION OF COSTS IN THE MEDICARE WAITING PERIOD FOR SOCIAL SECURITY DISABLED-WORKER BENEFICIARIES

Barry V. Bye*
Gerald F. Riley**

*Office of Research and Statistics
Social Security Administration

**Office of Research
Health Care Financing Administration

March 1989

Social Security Administration
Office of Policy
Office of Research and Statistics

This paper presents the statistical methods used to estimate Medicare costs in the waiting period that were presented in text tables 2-3 of Bye and Riley (1989).[1] The first part describes the development of Medicare utilization equations for each Social Security Disability Insurance (DI) program status group. The second part describes how these equations were used to predict expected costs per month and how the monthly estimates were aggregated to yield estimates of costs in the full two year waiting period and in the second year only. Finally, there is a brief discussion of the accuracy of the predictions.

MEDICARE UTILIZATION EQUATIONS

Multipart Models

Following the approach recommended by Duan et al. (1983), the model for Medicare utilization was assumed to have four parts.[2] Two of the equations dealt with the use of any Medicare services and the use of inpatient hospital services in a given year. The remaining two equations predict average Medicare reimbursement per month during the year for those beneficiaries who had 1 or more hospital stays and for those who used Medicare services but had no inpatient hospital stay. The four part model was expected to more accurately reflect the distribution of Medicare reimbursements than simpler models, but no specific test of the predictive power of various alternatives was performed as was done by Duan et al..

Logistic regression models were chosen for the categorical outcome variables. Consider the outcome as a 3-category variable, L, with values 1(no Medicare use), 2(Medicare use, no hospital use), and 3(hospital use). Two equations were defined by

$$\ln\left(\sum_{i > m} P(L=i) \Big/ P(L=m) \right) = x\alpha^{m} \quad , \quad m = 1,2 \qquad (1)$$

where x is a row vector of independent variables with a "1" in its first component corresponding to the constant, P(L=j) is the probability of a randomly selected observation falling into the jth category of L, and α^{m} is a column vector of unknown parameters to be estimated for the mth equation.

This specification has several nice features. First, the vectors of coefficients are easy to interpret. The components of α^{1} indicate which of the variables are predictive of Medicare utilization. The components of α^{2} indicate the factors associated with inpatient hospital use given Medicare use. A second advantage is that the likelihood function for the data can be written as the product of two binomial likelihoods; thus, the parameters can be estimated separately. A third feature is that the standard errors of the estimated coefficients can easily be corrected for the time dependence among the observations (see the discussion below).

To estimate the level of Medicare reimbursements, linear regression models were used with the logarithm of the reimbursement amount as the dependent variable. Using the logarithm of reimbursement as the dependent variable reduces the skewness (long right tail) of the distribution, thus improving prediction. It also insures the positivity of predictions from the estimated equations. The two equations were defined by

$$\ln(y_m) = x\beta^m , \quad m = 1,2 \qquad (2)$$

where x is a row vector of regressors including a "1" in the first component corresponding to the constant, β^m is a column vector of parameters to be estimated, y_m is the reimbursement amount. The case m=1 is taken to be the equation for log reimbursement given 1 or more inpatient hospital stays. The second equation, m=2, is the equation corresponding to no inpatient stays.

Outcome Variables

The principal Medicare outcomes examined in this article are derived from the Medicare reimbursement amounts for the years 1974 - 1981. In order to remove the effects of inflation, adjustments were made to the reimbursement data for each of the years 1974-1980 to convert reimbursement dollar amounts in those years to 1981 levels. The adjustments were applied separately to Part A and Part B reimbursements. For Part A services, an extension of the National Hospital Input Price Index was used.[3] This is an index measuring inflation in hospital input prices. The index is appropriate for the years in question because hospitals were reimbursed on a reasonable cost basis for that period of time. Therefore an index measuring inflation effect on hospital input prices should also measure reasonably well the effects of inflation on hospital reimbursements. Although the index does not take into account skilled nursing facility and home health agency services, separate indices were not developed for these factors because they represent a very small part of Part A reimbursements.

Part B reimbursements were indexed using data prepared by Health Care Financing Administration for the 1983 Annual report of the Board of Trustees of the Supplemental Medical Insurance Trust Fund.[4] The data show net increases in reasonable charges for physicians services, due to inflation, for disabled Medicare enrollees (excluding end stage renal disease cases). Technically, the index applies only to physician services, but these comprise the bulk of Part B services.

For each sample person with a period of Medicare eligibility, the average reimbursement per month of eligibility was computed for each of the years 1974-1981. No Medicare utilization in a particular year was defined to be an average reimbursement of $0. Inpatient hospital utilization is indicated by 1 or more inpatient discharges in the specified year.

It should be noted that some beneficiaries with no reimbursements may nonetheless have used some services. Costs of care that do not exceed the Medicare Part A or Part B deductible are not paid by Medicare and are therefore not recorded in the Medicare Statistical System. In addition, the Part B deductible remained at $50 during the period 1974-1981; high inflation during that period caused more beneficiaries to exceed the Part B deductible threshold over time, resulting in an increase in identified users of services, independent of any real changes in utilization.

Regressors

Newhouse (1981) reports a large literature on the demand for medical care that establishes relationships between medical care use and a number of demographic and economic factors.[5] Many of these factors were included in the analysis, but their interpretation may not be the same as that found elsewhere. The literature does not treat Medicare use by DI program beneficiaries specifically, and it may be erroneous to assume that interpretations of these determinants in other contexts apply to disabled beneficiaries as well. For disabled beneficiaries, such factors as age, education, and occupation can have a significant impact on the determination of eligibility for cash benefits. Applicants with intermediate impairment severity who are older, less educated, and lacking in transferable job skills are more likely to be awarded DI program benefits than other applicants with similar severity levels. Thus the younger, the more highly educated, and the more highly skilled beneficiaries are likely to have impairments of greater severity than beneficiaries with other characteristics. In the absense of a strong control for severity, associations between these three factors and Medicare use are likely to reflect variation in impairment severity as well as the variation in taste for medical care services among persons of differing demographic and economic backgrounds.

Much of the literature reported by Newhouse focusses on the cost of medical care to the individual as a primary determinant of medical care use. Variation in costs for the general public relates partly to possession of health insurance and the premium cost and provisions -- coverage, deductibles, coinsurance -- of the variety of health insurance policies. For Medicare enrollees, some of the variation in cost has been removed due to entitlement for Part A services that covers inpatient hospital costs and Part B that covers physician services. Still, some significant sources of cost variation remain unaccounted for -- for example, the possession of private supplementary insurance to cover Part A and Part B deductibles and coinsurance. In addition, DI program beneficiaries who also receive cash benefits from the Supplemental Security Insurance (SSI) program are also entitled to Medicaid benefits that cover Medicare premium, deductible and coinsurance, as well as additional costs such as nursing home care and the cost of perscription drugs. Unfortunately, none of the information on other forms of insurance was available in the data used to estimate the equations, leaving an important gap in the set of regressors.

The full set of regressors used in the analysis is described in Figure 1 (page 29). Several of these factors require further explanation.

Primary Diagnosis--The primary diagnosis identifies the principal medical cause of the disabling condition as determined at the time cash benefits were awarded. The diagnosis has been coded according to the ICD-8-CM system which was in effect in 1972.[6] The diagnoses have been categorized into body system groups. One might expect, and in fact observes, substantial variation in DI programmatic and Medicare outcomes among these groups. One also would expect there to remain large within group variations because the groups are so broadly defined.

Occupation--For most persons in the study, the occupation reported represents the individual's major occupation in the 15 year period prior to application for benefits. Occupations are coded using the 3-digit codes found in the Dictionary of Occupational Titles.[7] The categories presented in the following analysis are the major occupational groups based on the first digit of the code.

Primary Insurance Amount--The primary insurance amount (PIA) is the dollar figure upon which cash benefits are based.[8] Computationally, the PIA is a function of the number of years of SSA covered earnings prior to onset of disability and the level of earnings in those years. Analytically, the PIA serves roughly as a proxy for level of lifetime earnings and also as a rough indication of current economic status since it is directly related to the amount of cash benefits received.

The PIA amount used in this analysis is PIA as of December 31, 1985. The data available for this study did not include a uniform reporting of PIA at the time of entitlement, 1972. The 1985 PIA reflects adjustments that were made over the years due to legislative changes and inflation. In the case of recovered beneficiaries, the PIA might also reflect recent work prior to a second DI or old age entitlement.

Time to the end of the Observation Period -- For this analysis, the observation period for an individual beneficiary can end in one of 3 ways: death, recovery, and attainment of age 65 or still in the program. The time to the end of the observation period is a set of dummy variables (variables 38-44 in Figure 1.) indicating the number of years prior to the end of the period of the particular observation at hand. If the observation year is in the last year -- for example, the year of death -- then variable 38 = 1 and all others (39-44) are set to 0 (See Figure 1.). If the observation year is the year prior to the last year, then variable 38 is set to 0, variable 39 is set to 1, and all others in the set are given the value 0. The earliest observation year is 7 years prior to the last year of observation -- that is, 1974 for periods that ended in 1981. In this case all variables (38 - 44) are set to 0.

The primary purpose of this set of variable is to account for changes in the last several years before death and recovery where utilization and costs are expected to increase and decrease respectively. It might also pick up an upward drift in the probability of Medicare use (because the level of Part B Medicare deductibles was not adjusted for inflation during this period of time) and a corresponding downward drift in probability of hospital use given Medicare use (because reimbursement for hospital use would not be affected by the Part B deductible.)

Census Division -- It is well known that medical care utilization and costs vary geographically.[9] States were grouped into the nine Census Bureau geographic divisions as the geographic construct for this analysis. The classification is shown in Figure 2 (page 31).

Months in the observation year -- While utilization of services was known for each observation year, no information was provided as to when the services were used within the year. The first year (1974) and last year of observation can contain less than 12 months of eligibility. Since the utilization equations predict utilization at yearly levels, probabilities of Medicare or inpatient hospital use will be lower, on average, in those years with fewer months of exposure. At the same time, the average cost per month in a year, given Medicare use in that year, is likely to be higher, on average, in those years with less months, especially when the part year happened to contain an inpatient hospital stay.

In order to account for variation in utilization due to the number of months, the number of months was added to the equations as a single continuous variable under the assumption that its effect would be roughly linear and additive. Several equations were examined with the number of months entered as a set of discrete dummy variables. The effects appeared to be more or less linear; a formal test of linearity was not performed.

Conditioning on DI Program Outcome

Bye et al. (1987) report that Medicare utilization and costs for DI program beneficiaries increase greatly in the year of death and the year prior to death.[10] Medicare use also appears to decrease in the year of recovery and is relatively low in all years for beneficiaries who ultimately recover. These findings suggest that separate sets of utilization equations should be estimated depending on DI program status as of 1981. Because program status in the first two years also is known for beneficiaries who did not survive the waiting period, such equations can be applied directly to the sample cases to predict Medicare costs.

In addition, persons whose disability period ends in death are generally less healthy while enrolled than persons who stayed on the rolls through 1981 or recover and leave the rolls before that time. And in turn, beneficiaries who stay on the rolls are likely to represent persons who are less heatlhy than those who recover. Thus conditioning on outcome might increase the accuracy of the prediction of costs in the first two years for those beneficiaries who remained in the DI program for more than two years.

To condition on program outcome, the sample cases were divided into 3 groups: (1) death prior to 1982, (2) recovered prior to 1982, and (3) attained age 65 prior to 1982 or still in the program in 1982. A separate four part model was estimated for each of the three groups.

Multiple Observations

Each beneficiary in the sample could supply from 1 to 8 observations to each of the equations estimated, depending on the particular equation in question and length of the observation period. For the logistic regression of Medicare use, each sample person supplied one data record for each observation year. The dependent variable indicated whether $1 or more of Medicare reimbursement (1=yes, 0=no) occurred in the year. The number of observations supplied by a single beneficiary ranged from 1 to 8 depending on whether cash benefits were terminated in the period 1974 - 1981. Given Medicare use in at least one year, the sample beneficiary supplied one observation for each year of Medicare use to the estimation of the logistic regression for inpatient hospital use. (For a given year, the dependent variable indicating hospital use was assigned a value of "1" if 1 or more inpatient hospital discharges occurred in the year, a value of "0" was assigned if no discharges occurred.) If Medicare services were used in all 8 years, then eight observations would be supplied to the hospital use regression.

Also, one observation was supplied to each of the two cost equations for each year of Medicare use. An observation was supplied to the first cost equation, conditional on an inpatient stay, for each year with 1 or more inpatient hospital stays. The remaining observations were supplied to the equation for average monthly costs in the years when there was no inpatient stay.

Intertemporal Dependence

The description of the analysis approach given thus far treats each year's Medicare use as though it arises as a self-contained event independent of utilization that may have occurred in previous years. This, of course, is not the correct representation of the actual processes that produced the data. Those processes consist of episodes of ill health and medical care use resulting in sequences of doctor visits, possible hospitalizations and rehospitalizations and the costs associated with the provision of these services. To arbitrarily divide the outcomes of these processes, which unfold continuously in time, by calendar year most certainly does not result in a set of observations for an individual that are stochastically independent. Because episodes of Medicare use will often cross calendar years, one can expect strong positive associations between observations of Medicare use and hospital use and their associated costs from one year to the next.[11]

An analysis based on likelihood theory which presupposes independence in the observations will clearly result in a misspecification of the likelihood functions. Fortunately the misspecification does not affect the consistency of the estimates or their asymptotic normality. The estimated standard errors associated with the classical maximum likelihood estimator are incorrect but have been recomputed by methods developed by Liang and Zeger.[12]

Regression Results

The regression results for the 12 equations (4 model parts by 3 program status conditions) are given in tables 1-12 of the appendix. The tables show the estimated coefficients and the standard errors and the t-values and a set of prediction coefficient to be discussed below. The sample sizes for the recovered group (Tables 5-8) were too small to permit estimation of the full set of parameters (Figure 1). In all cases except for age and occupation, a reduction in the number of parameters was achieved by increasing the number of categories in the reference group. For the age variable the highest two categories -- 50-59, and 60 and over-- also were combined. For the occupation variable, the reference group now contains the first two categories, -- professional and clerical occupations. All other occupations were combined into a single variable labeled "blue collar."

A substantive interpretation of the coefficients contained in the Appendix tables is complicated by the use of the 4-part model, as well as the stratification on DI Program outcomes. Nonetheless, where the results present a consistent picture, some of the relationships between reggressors and outcome variables are briefly described.

Female beneficiaries tended to use services significantly more often and incur higher costs than male beneficiaries, regardless of program outcome as indicated by the significant positive coefficients in each of the utilization equations. One exception was that there was a significantly lower probability of hospital use, given Medicare use, among women who turned 65 or remained on the rolls. Blacks were less likely to use services than nonblacks (negative coefficients in the utilization equations); given use, however, blacks incurred higher average expenses (positive coefficients in the reimbursement equations). This held true for both hospital and nonhospital services. This result may reflect high costs of inner city hospitals and other inner city health care providers.

The characteristic most consistently associated with high use and expenses was End Stage Renal Disease (ESRD) status. Beneficiaries entitled to Medicare through their end stage renal disease tend to have very high health care costs due to their need for dialysis and for transplant services.

There were consistent differences among Census divisions in the use and cost of services, particularly with respect to the use and cost of hospital services. Divisions 1 (New England) and 9 (Pacific) were least similar to other sections of the country. Among decedents, the probability of hospital use, given Medicare use, was significantly lower in New England than all other divisions except the Pacific division. Among those turning 65 or remaining in the program, hospital use was also relatively lower in both the New England and Pacific Divisions. Hospital expenses, given hospital use, were significantly higher in the New England and Pacific divisions among both decedents and those turning 65 or staying in the program.

Among decedents, the proximity to death was an important determinant of expense levels. The dichotomous variable indicating year of death had large positive regression coefficients in the use and expense models. The first and second years before death were also characterized by relatively high levels of expense; in general, use and costs were lowest for years farthest from death, as expected.

Lastly, the number of months in the year of event was included in the models to control for length of exposure during the year. As expected, the probability of use was significantly positively associated with the number of months of exposure in each of the models. Average monthly costs, given use, were negatively associated with months of exposure.

Prediction equations

As indicated by the t-values in the appendix tables, many of the estimated coefficients across the 12 equations had standard errors that were quite large relative to the size of the coefficients. This suggests that the exclusion of many of these variables from a final set of prediction equations might produce equations with smaller mean square errors of prediction than that obtained from the full equations. Although omitting variables might result in some bias in the predictions if in fact the variable is associated with outcome measured, the increase in bias will probably be more than made up for by the reduction in variance obtained from the equations with fewer variables, thus leading to smaller mean square errors of prediction.

In general, a t-value of 2.00 was used to determine whether a variable should be included or not. It is unlikely that variables that have coefficients with standard errors as large as one half of the estimated coefficient will make a significant contribution to the prediction. However, variables with coefficients having smaller standard errors often do not make a significant contribution either. Thus we suspect that optimal prediction equations would have numbers of included variables somewhat less than those presented here.

Notice that in several instances variables were included with t-values less than 2, and in several cases, variables were excluded with t-values greater than 2. The former usually pertained to the inclusion of regional dummy variables. The latter occurred mainly in cases where the only significant coefficient in a set was the unknown group. It was not clear whether keeping such variables in the equation would improve predictions since we had no understanding as to why the results turned out the way they did.

After omitting the variables, all 12 equations were reestimated. The resulting coefficients are shown in the last column of each of the appendix tables. The adjusted R^2s for the reimbursement equations indicate little or no slippage in the prediction power of these equations. We suspect that the same is true for the logistic regression equations of Medicare and hospital use but have no summary statistics to present as evidence.

Waiting Period Cost Projections

Since all of the equations provide estimates that are calendar year specific, the cost in the two year waiting period was estimated in three parts corresponding to the amount that would be incurred in each of the calendar years 1972, 1973, and 1974. The projected costs in each of these calendar years was the expected reimbursement per month in the year times the number of months of eligibility in the year that were contained in the waiting period. The expected reimbursement per month in the ith year is given by the following:[13]

$$R_i = R_i^H P_i^H + R_i^{NH} P_i^{NH} \tag{3}$$

where

R_i^H is the average reimbursement per month given 1 or more inpatient hospital stays in the ith year

P_i^H is the probability of one or more inpatient stays in the ith year

R_i^{NH} is the average reimbursement per month assuming no inpatient hospital stays in the ith year

P_i^{NH} is the probability of no inpatient stays in the ith year but some Medicare use.

These quantities are obtained from the four part models described above in the following way.

$$P_i^H = \text{Prob}(\text{Medicare}_i)\text{Prob}(\text{hospital}_i|\text{Medicare}_i)$$

$$P_i^{NH} = \text{Prob}(\text{Medicare}_i)(1 - \text{Prob}(\text{hospital}_i|\text{Medicare}_i)) \tag{4}$$

where Prob(.) are obtained directly from equation (1).

Computation of R_i^H and R_i^{NH} is somewhat more complex because the reimbursement amounts were transformed to the logarithmic scale before the equations were estimated. Letting y represent the reimbursement amount, the transformed equations take the form

$$\ln y = x\beta + \epsilon$$

where ϵ is a random disturbance.

One can show that the expected value of y for a given set of regressors, x, is given by:

$$E(y|x) = \exp(x\beta)E(\exp(\epsilon|x)). \tag{5}$$

Thus an estimate of $E(y|x)$ not only involves x and β but requires an estimate of $E(\exp(\epsilon)|x)$ as well. If the error distribution does not depend on x (the usual regression assumption) then the second term of (5) is a constant, $\phi = E(\exp(\epsilon))$. Further, if ϵ

is normally distributed $N(0,\sigma^2)$, then $\phi = \exp(\sigma^2/2)$.

When there is uncertainty about the distribution of ϵ, a nonparametric estimate of ϕ may be more appropriate. Duan et al. have proposed such an estimate, the smearing estimate, which is the average of the exponentiated least squares residuals.[14] This estimate is given by:

$$\hat{\phi} = 1/n \sum_{k=1}^{n} \exp(\hat{\epsilon}_k) , \tag{6}$$

where n is the number of sample cases and $\epsilon_k = y_k - x_k\beta$ is the kth residual. Equation (6) provides a consistent estimate of ϕ for all distributions of ϵ in those cases where the distribution of ϵ does not depend on x. Using equation (6), the estimates of the average reimbursements per month are obtained from equation (2) by:

$$\begin{aligned} R_i^H &= \exp(x_i\beta_H)\hat{\phi}_H \\ R_i^{NH} &= \exp(x_i\beta_{NH})\hat{\phi}_{NH} \end{aligned} \tag{7}$$

Equations (4) and (7) are substituted into equation (3) to obtain expected reimbursement per month in a specific year.

Eliminating both years or the second year of the waiting period

For each sample case, the cost in the full two year waiting period was computed by summing the estimated costs for each of the parts of the waiting period that fell in the years 1972, 1973, and 1974. For each year, the expected cost per month, given by equation (3), was multiplied by the number of months of eligibility in the specified calendar year that was also part of the waiting period. As an example, consider a beneficiary entitled in August, 1972, and still in the program two years later. This beneficiary would have 5 months (August - December) of eligibility in 1972, 12 months in 1973, and 7 months in 1974 if the waiting period were completely eliminated. The estimated Medicare costs in this case would be:

$$5R_{72} + 12R_{73} + 7R_{74}$$

where R_i is the expected monthly cost in the ith year as given by equation (3).

If the beneficiary terminated within the first two years, then only that time in the program was accounted for. For example, a beneficiary who was entitled in April, 1972, and terminated in October 1973 would have estimated Medicare costs:

$$9R_{72} + 10R_{73}$$

Keep in mind that the equations used to compute the R_i are conditioned on the reason for termination, as discussed above.

Estimation of costs of only the second year if the waiting period were eliminated follows a similar computational scheme. In this case, estimates of costs were made just for the relevant portions of 1973 and 1974. Again, an estimate was made for only part of the second year if the beneficiary terminated within two years. If the beneficiary terminated within one year then no second year costs were incurred.

Text tables 2 and 3 in Bye and Riley (1989)

The average costs in the waiting period per beneficiary shown in text tables 2 and 3 were obtained by estimating the costs for each beneficiary and then taking the arithmetic mean of the estimated costs for the beneficiaries in each table cell. Cases with End Stage Renal Disease (ESRD) were excluded from this analysis. Beneficiaries with this condition have no Medicare waiting period and therefore there would be no additional costs incurred by this group. ESRD cases represent about 0.5 percent of the entitlement cohort under age 62.

Standard errors for the estimated averages were not computed because the computation would have been very complex, depending on the covariance structure of the estimated coefficients, the formulas used to obtain the average waiting period estimates and the effects of sampling from the particular population of interest. In addition, there are also nonsampling errors in the estimates to the extent that the prediction equations are not appropriate for the waiting period; and the nonsampling errors may be large relative to the sampling errors. There is no way of quantifying the nonsampling errors and thus no way to directly estimate the precision of the predictions even if sampling errors had been estimated.

Accuracy of the Estimating Equations

In this section some information is provided by which the overall fit of the prediction equations to the data from which they were estimated can be gauged. A detailed analysis of the fit of each of the 12 equations was not made due to resource constraints. A general comparison of estimated and actual reimbursements based on all equations using the estimation approach described above was performed. Over the entire observation period the models performed reasonably well. For the full sample, the average observed Medicare costs were \$5722. The average estimated costs were \$6018, higher than the observed costs by about 5 percent. Approximately the same overall differences were obtained for each of the DI program groups as shown in the table on the following page.

Actual and Estimated Medicare costs 1974-1981
by DI Program Termination Status

Termination status	Number of cases	Average Direct Estimate 1974-81	Average Model Estimate 1974-81	Percent difference
Total	15375	$5722	$6018	5.2%
Death	3128	5967	6202	3.9
Recovery	805	596	604	1.3
Age 65	5320	4748	5024	5.8
None	6122	7855	8307	5.8

The differences between actual and estimated costs, however, were somewhat higher in the first three years as shown in the next table which compares direct and model estimates by year.[15]

Actual and Estimated Medicare Costs by Year

Year	Number of cases	Average Direct Estimate 1974-81	Average Model Estimate 1974-81	Percent difference
1974	15375	$567	$672	18.5%
1975	14826	1158	1323	14.2
1976	13525	1255	1347	7.3
1977	11821	1346	1330	-1.2
1978	10330	1379	1351	-2.0
1979	9026	1452	1375	-5.3
1980	7873	1498	1497	0.0
1981	6928	1381	1586	14.8

Of concern in these tables is the relatively large overestimates from the models in 1974 and 1975. The models were not structured to account for such a pattern. These patterns might be an artifact of the inception of the program of Medicare coverage for the disabled in 1974, that is, underutilization due to newness of the program. If this is true, then the average experience over time of the various groups probably better reflects a typical cohort experience in the early years than that which would be obtained from using just the early years experience itself. However, if Medicare utilization is really substantially smaller in the first several years of eligibility as these data nominally show, then using overall averages to project the costs of eliminating the waiting period might have resulted in an overestimate of the waiting period costs. Further investigation of this matter is beyond the scope of this analysis because it would require the longitudinal analysis of later cohorts to see whether this type of pattern persists for them as well.

The model also overestimates costs for 1981. The reason for this is probably that the CMHS contains data from all Medicare claims received through June 1982. Some claims for Medicare services rendered in 1981 would have been received after that date, resulting in an undercount of actual reimbursements in 1981.

NOTES

1. Bye, Barry V. and Riley, Gerald F. (1989) "Eliminating the Medicare Waiting Period for Social Security Disabled-Worker Beneficiaries," Social Security Bulletin, Forthcoming.

2. Duan, Naihua; Manning, William G. Jr.; Morris, Carl N.; and Newhouse, Joseph P. (1982) "A Comparison of Alternative Models for the Demand for Medical Care," report R-2754-HHS, The Rand Corporation.

3. Freeland, Mark S.; Anderson, Gerard; and Schendler, Carol Ellen (1979) "National Hospital Input Price Index," Health Care Financing Review, Summer, 37-61.

4. Health Care Financing Administration Annual Report of the Board of Trustees of the Federal Supplemental Insurance Trust Fund, June, 1983.

5. Newhouse, Joseph P. (1981) "The Demand for Medical Care Services: A Retrospect and Prospect," report R-2691-HHS, The Rand Corporation.

6. World Health Organization (1967) Manual of the International Classification of Diseases, Injuries, and Causes of Death, Geneva.

7. Department of Labor (1965) Dictionary of Occupation Titles. Washington D. C. .

8. Social Security Administration (1982) Social Security Handbook, Washington D. C.: U. S. Government Printing Office.

9. Wennberg, J.; Freeman, J.; and Culp, W., "Are hospital services rationed in New Haven over-utilized in Boston?" The Lancet, May 23, 1987, pages 1185-1189.

Wennberg, J. and Gittelsohn, A., "Variations in medical care among small areas," Scientific American, Vol. 246 No. 4, pages 120-134, 1982.

Wennberg, J. and Gittelsohn, A., "Small area variations in health care delivery," Science, Vol. 182 No. 4117, pages 1102-1108, December 14, 1973.

Chassin, M.; Brook, R.; and Park, R. E., "Variations in the use of medical and surgical services by the Medicare population," New England Journal of Medicine, 314, pages 285-290, January 30, 1986.

10. Bye, Barry; Riley, Gerald; and Lubitz, James (1987) "Medicare Utilization by Disabled-Worker Beneficiaries: A Longitudinal Analysis," Social Security Bulletin, Vol. 50, No. 12, 13-28.

11. Beebe J., "Medicare Reimbursement and Regression to the Mean," Health Care Financing Review, 9,3, pages 9-22, 1988.

McCall G. and Wai N., "An Analysis of the Use of Medicare Services by the Continuously Enrolled Aged," Medical Care, 21,6 pages 567-585, 1983.

12. Bye, Barry V., and Riley, Gerald F., "Model Estimation When Observations Are Not Independent: Application of Liang and Zeger's Methodology to Linear and Logistic Regression Models," Sociological Methods and Research, May, 1989, Forthcoming.

13. Strictly speaking, the values for P_i^H and P_i^{NH} are annual estimates. For ease of exposition, it is convenient to take them as a probability distribution over any month in the year as well. As will be seen in the next section, the annual character of their values is essentially reestablished when the expected monthly reimbursement in a year is multiplied by the number of eligible months in the year, yielding an annual reimbursement estimate to which the annual probabilities apply.

14. See 2. op cit.

15. The average actual reimbursement for 1974 is smaller than that of other years because 1974 was the first year of Medicare eligibility for this cohort. Thus, virtually all enrollees had only part year eligibility.

Figure 1. Regressor Variables

Variable	Omitted Category
Sex	
1. sex = female	sex = male
Race	
2. Race = black	Race = white
3. Race = other	
Occupation	
4. Occupation = clerical	
5. Occupation = service	
6. Occupation = farm	
6. Occupation = farm	
7. Occupation = processing	Occupation = professional
8. Occupation = Machine	
9. Occupation = Bench work	
10. Occupation = Structural	
11. Occupation = Miscellaneous	
12. Occupation = Unknown	
Diagnosis	
13. Diagnosis = neoplasm	
14. Diagnosis = endocrine	
15. Diagnosis = mental	
16. Diagnosis = nervous	
17. Diagnosis = eye and ear	
18. Diagnosis = circulatory	Diagnosis = Infectious and Parasitic
19. Diagnosis = respiratory	
20. Diagnosis = digestive	
21. Diagnosis = genito-urinary	
22. Diagnosis = musculoskeletal	
23. Diagnosis = traumatic	
24. Diagnosis = other, unknown	
Primary Insurance Amount	
25. PIA = 300 - 399	
26. PIA = 400 - 499	PIA = less than 300
27. PIA = 500 - 599	
28. PIA = 600 and over	

Figure 1. (continued)

Education

29. Education = 0 - 8 years	
30. Education = 9 - 12 years	Education = none
31. Education = 13 years or more	
32. Education = other, unknown	

End Stage Renal Disease

33. ESRD = yes	ESRD = no

Age in Year of Observation

34. Age = 30 - 39	
35. Age = 40 - 49	Age = less than 30
36. Age = 50 - 59	
37. Age = 60 - 61	

Time Before End of Observation

38. Time = last year	
39. Time = 1 year before last	
40. Time = 2 years before last	
41. Time = 3 years before last	Time = 7 years before last
42. Time = 4 years before last	
43. Time = 5 years before last	
44. Time = 6 years before last	

Census Division

45. Division2 = Mid Atlantic	
46. Division3 = East North Central	
47. Division4 = West North Central	
48. Division5 = South Atlantic	Division1 = New England
49. Division6 = East South Central	
50. Division7 = West South Central	
51. Division8 = Mountain	
52. Division9 = Pacific	

Months of Eligibility in Observation Year

53. Months = 01 - 12 = 1 to 12 months	(not applicable)

Figure 2. Census Divisions

Division	State
1. New England	Maine New Hampshire Vermont Massachusetts Rhode Island Connecticut
2. Middle Alantic	New York New Jersey Pennsylvania
3. East North Central	Ohio Indiana Illinois Michigan Wisconsin
4. West North Central	Minnesota Iowa Missouri North Dakota South Dakota Nebraska Kansas
5. South Atlantic	Delaware Maryland District of Columbia Virginia West Virginia North Carolina South Carolina Georgia Florida
6. East South Central	Kentucky Tennessee Alabama Mississippi
7. West South Central	Arkansas Louisiana Okalahoma Texas

Figure 2. (continued)

8. Mountain	Montana Idaho Wyoming Colorado New Mexico Arizona Utah Nevada
9. Pacific	Washington Oregon California Alaska Hawaii

Table 1. Logistic regression coefficients, Medicare use, death cases

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	-1.4522	.5201	-2.7922	-1.3356
	Sex				
1	FEMALE	.6955	.0929	7.4869	.7121
	Race				
2	BLACK	-.2366	.0880	-2.6904	-.2865
3	OTHER	-.0314	.2899	-.1085	
	Occupation				
4	CLERICAL	.1036	.1467	.7063	
5	SERVICE	-.1460	.1420	-1.0284	
6	FARM	-.1707	.1905	-.8961	
7	PROCESSING	.0059	.1986	.0295	
8	MACHINE	-.1820	.1476	-1.2331	
9	BENCH WORK	.1083	.1885	.5745	
10	STRUCTURAL	-.0892	.1386	-.6436	
11	MISCELLANEOUS	-.1418	.1370	-1.0351	
12	UNKNOWN	-.5992	.1732	-3.4588	
	Diagnosis				
13	NEOPLASMS	.0466	.2909	.1602	
14	ENDOCRINE	.2153	.3088	.6971	
15	MENTAL	-.5678	.2871	-1.9776	-.5493
16	NERVOUS	-.0538	.3125	-.1721	
17	EYE AND EAR	-.2249	.3350	-.6714	
18	CIRCULATORY	-.0241	.2673	-.0902	
19	RESPIRATORY	.0759	.2807	.2704	
20	DIGESTIVE	-.1244	.3086	-.4030	
21	GENITO-URINARY	-.2845	.4263	-.6673	
22	MUSCULOSKELETAL	-.0593	.2802	-.2115	
23	TRAUMATIC	-.4088	.3089	-1.3236	
24	OTHER,UNKNOWN	-.2631	.3115	-.8448	
	PIA				
25	300-399	.1368	.1086	1.2604	
26	400-499	.0640	.1143	.5602	
27	500-599	.1628	.1168	1.3937	
28	600 AND OVER	.3744	.1408	2.6590	.2789

Table 1. continued

Education				
29 1-8 YEARS	.0685	.2721	.2518	
30 9-12 YEARS	.1647	.2805	.5872	
31 13 AND OVER	.0054	.2816	.0191	
32 OTHER AND UNKNOWN	.6836	.3205	2.1326	
ESRD				
33 ESRD	1.5254	.3481	4.3818	1.4335
Age				
34 30-39	-.5339	.2438	-2.1901	-.3862
35 40-49	-.4225	.2303	-1.8344	-.2872
36 50-59	-.6266	.2182	-2.8712	-.4465
37 60 AND OVER	-.3746	.2231	-1.6790	-.1812
Time to Death				
38 YR OF DEATH	1.7280	.1655	10.4380	1.5487
39 1 YR BEFORE DEATH	.6132	.1688	3.6330	.4392
40 2 YRS BEFORE DEATH	.3803	.1681	2.2618	.2086
41 3 YRS BEFORE DEATH	.2431	.1685	1.4429	
42 4 YRS BEFORE DEATH	.0917	.1649	.5562	
43 5 YRS BEFORE DEATH	.2294	.1684	1.3620	
44 6 YRS BEFORE DEATH	.0419	.1689	.2481	
Census Division				
45 DIVISION 2	.1694	.1586	1.0680	
46 DIVISION 3	.0548	.1588	.3454	
47 DIVISION 4	-.0315	.1860	-.1691	
48 DIVISION 5	.0454	.1564	.2904	
49 DIVISION 6	-.0100	.1735	-.0577	
50 DIVISION 7	.0703	.1715	.4101	
51 DIVISION 8	-.3408	.2162	-1.5761	
52 DIVISION 9	.1865	.1648	1.1312	
Months				
53 MTH IN YR OF EVENT	.1487	.0068	21.9254	.1482
NUMBER OF OBSERVATIONS	12052			
NUMBER OF CASES	3126			

Table 2. Logistic regression coefficients, hospital use, death cases

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	-.9476	.6233	-1.5202	-.7174
	Sex				
1	FEMALE	.0451	.0959	.4709	.0778
	Race				
2	BLACK	-.2680	.0964	-2.7791	-.2209
3	OTHER	-.5839	.3096	-1.8860	
	Occupation				
4	CLERICAL	.1169	.1375	.8499	
5	SERVICE	.1943	.1470	1.3219	
6	FARM	.0522	.2231	.2342	
7	PROCESSING	.4541	.2214	2.0512	.2598
8	MACHINE	.2143	.1515	1.4148	
9	BENCH WORK	.2732	.1748	1.5629	
10	STRUCTURAL	.3935	.1463	2.6904	.2283
11	MISCELLANEOUS	.2152	.1410	1.5264	
12	UNKNOWN	.4154	.1941	2.1399	
	Diagnosis				
13	NEOPLASMS	-.2672	.3052	-.8754	
14	ENDOCRINE	-.0765	.3097	-.2470	
15	MENTAL	-.1126	.3098	-.3634	
16	NERVOUS	-.5866	.3188	-1.8401	
17	EYE AND EAR	.4686	.4661	1.0054	
18	CIRCULATORY	-.3748	.2809	-1.3343	
19	RESPIRATORY	-.0479	.2977	-.1609	
20	DIGESTIVE	.2796	.3208	.8715	
21	GENITO-URINARY	-.8762	.3877	-2.2603	-.6114
22	MUSCULOSKELETAL	-.3214	.2936	-1.0946	
23	TRAUMATIC	-.3313	.3496	-.9475	
24	OTHER,UNKNOWN	-.1605	.3393	-.4731	
	PIA				
25	300-399	.0685	.1204	.5688	
26	400-499	-.0637	.1219	-.5227	
27	500-599	.0036	.1237	.0288	
28	600 AND OVER	-.1875	.1428	-1.3125	

Table 2. continued

Education				
29 1-8 YEARS	.2605	.3839	.6784	
30 9-12 YEARS	.2709	.3887	.6969	
31 13 AND OVER	.3670	.3913	.9378	
32 OTHER AND UNKNOWN	.0784	.4154	.1888	
ESRD				
33 ESRD	.7905	.2219	3.5620	.8494
Age				
34 30-39	.0129	.2935	.0438	
35 40-49	-.0372	.2848	-.1307	
36 50-59	-.2591	.2747	-.9432	
37 60 AND OVER	-.2859	.2795	-1.0227	
Time to Death				
38 YR OF DEATH	2.2030	.2642	8.3376	1.8453
39 1 YR BEFORE DEATH	.8876	.2629	3.3766	.5403
40 2 YRS BEFORE DEATH	.5919	.2642	2.2407	.2583
41 3 YRS BEFORE DEATH	.5093	.2637	1.9313	.1749
42 4 YRS BEFORE DEATH	.3219	.2627	1.2252	
43 5 YRS BEFORE DEATH	.3960	.2694	1.4699	
44 6 YRS BEFORE DEATH	.2553	.2615	.9763	
Census Division				
45 DIVISION 2	.3431	.1558	2.2027	.3559
46 DIVISION 3	.4948	.1587	3.1182	.5303
47 DIVISION 4	.4361	.1925	2.2661	.4456
48 DIVISION 5	.4286	.1583	2.7076	.4164
49 DIVISION 6	.5601	.1775	3.1560	.6204
50 DIVISION 7	.4089	.1813	2.2556	.4399
51 DIVISION 8	.6685	.2412	2.7720	.6759
52 DIVISION 9	.2002	.1618	1.2374	.2248
Months				
53 MTH IN YR OF EVENT	.0289	.0104	2.7865	.0309
NUMBER OF OBSERVATIONS	7028			
NUMBER OF CASES	2518			

Table 3. Regression coefficients for log reimbursement, inpatient hospital stay, death cases

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	6.7915	.4108	16.5308	6.8438
	Sex				
1	FEMALE	.2711	.0602	4.5058	.2800
	Race				
2	BLACK	.3258	.0677	4.8148	.2839
3	OTHER	.2511	.1960	1.2814	
	Occupation				
4	CLERICAL	-.0140	.0906	-.1543	
5	SERVICE	-.1665	.0923	-1.8043	
6	FARM	-.4038	.1357	-2.9763	-.3539
7	PROCESSING	-.0618	.1512	-.4086	
8	MACHINE	-.0785	.0941	-.8342	
9	BENCH WORK	-.1094	.1152	-.9500	
10	STRUCTURAL	-.1771	.0918	-1.9288	-.1070
11	MISCELLANEOUS	-.1048	.0890	-1.1766	
12	UNKNOWN	-.1806	.1316	-1.3720	
	Diagnosis				
13	NEOPLASMS	-.2385	.1509	-1.5809	
14	ENDOCRINE	.0410	.1661	.2471	
15	MENTAL	-.3611	.1650	-2.1892	-.2453
16	NERVOUS	-.3808	.1727	-2.2045	-.1931
17	EYE AND EAR	-.3235	.2134	-1.5159	
18	CIRCULATORY	-.3445	.1307	-2.6353	-.2302
19	RESPIRATORY	-.1192	.1424	-.8371	
20	DIGESTIVE	-.1479	.1589	-.9304	
21	GENITO-URINARY	-.2139	.2373	-.9014	
22	MUSCULOSKELETAL	-.1905	.1412	-1.3493	
23	TRAUMATIC	-.2614	.1924	-1.3583	
24	OTHER,UNKNOWN	.0617	.1815	.3399	
	PIA				
25	300-399	.0380	.0737	.5159	
26	400-499	.0711	.0798	.8904	
27	500-599	.1933	.0761	2.5395	.1589
28	600 AND OVER	.1351	.0886	1.5245	.1665

Table 3. continued

Education				
29 1-8 YEARS	.2108	.2030	1.0383	
30 9-12 YEARS	.3241	.2049	1.5819	
31 13 AND OVER	.3392	.2066	1.6419	
32 OTHER AND UNKNOWN	.2443	.2345	1.0416	
ESRD				
33 ESRD	1.3918	.1418	9.8159	1.4270
Age				
34 30-39	.2109	.2205	.9564	
35 40-49	-.1741	.2049	-.8495	
36 50-59	-.1620	.2010	-.8060	
37 60 AND OVER	-.1449	.2026	-.7150	
Time to Death				
38 YR OF DEATH	.6272	.2274	2.7583	.5175
39 1 YR BEFORE DEATH	.4533	.2297	1.9734	.3500
40 2 YRS BEFORE DEATH	.1585	.2310	.6862	
41 3 YRS BEFORE DEATH	.2190	.2320	.9441	
42 4 YRS BEFORE DEATH	-.0065	.2338	-.0278	
43 5 YRS BEFORE DEATH	-.0537	.2321	-.2315	
44 6 YRS BEFORE DEATH	-.3182	.2300	-1.3834	
Census Division				
45 DIVISION 2	-.1018	.1130	-.9015	-.0798
46 DIVISION 3	-.1890	.1114	-1.6968	-.1678
47 DIVISION 4	-.4036	.1363	-2.9605	-.3960
48 DIVISION 5	-.3646	.1111	-3.2816	-.3588
49 DIVISION 6	-.6345	.1265	-5.0149	-.6401
50 DIVISION 7	-.4529	.1270	-3.5673	-.4552
51 DIVISION 8	-.3927	.1599	-2.4561	-.3552
52 DIVISION 9	.0396	.1178	.3361	.0622
Months				
53 MTH IN YR OF EVENT	-.0886	.0070	-12.6505	-.0883
NUMBER OF OBSERVATIONS	4689			
NUMBER OF CASES	2291			
STANDARD ERROR	1.2010			1.2076
ADJUSTED R2	.2070			.1982
PHI - SMEARING				1.7730
PHI - NORMAL				2.0732

Table 4. Regression coefficients for log reimbursement, no inpatient hospital stay, death cases

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	2.0680	.6455	3.2034	3.3495
	Sex				
1	FEMALE	.3414	.0965	3.5356	.3586
	Race				
2	BLACK	.4422	.1187	3.7265	.3101
3	OTHER	.0568	.2710	.2095	
	Occupation				
4	CLERICAL	.1025	.1506	.6810	
5	SERVICE	-.2432	.1461	-1.6638	
6	FARM	-.1465	.2066	-.7090	
7	PROCESSING	-.0952	.2138	-.4451	
8	MACHINE	-.1999	.1508	-1.3259	
9	BENCH WORK	.0370	.1692	.2185	
10	STRUCTURAL	-.0913	.1559	-.5856	
11	MISCELLANEOUS	-.2486	.1411	-1.7623	
12	UNKNOWN	-.2643	.2357	-1.1213	
	Diagnosis				
13	NEOPLASMS	.6049	.4188	1.4442	
14	ENDOCRINE	.1924	.4396	.4375	
15	MENTAL	.0479	.4254	.1127	
16	NERVOUS	.2401	.4226	.5681	
17	EYE AND EAR	-.2565	.4806	-.5337	
18	CIRCULATORY	.0093	.4046	.0231	
19	RESPIRATORY	.4646	.4210	1.1035	
20	DIGESTIVE	.0531	.4731	.1123	
21	GENITO-URINARY	.3642	.5372	.6780	
22	MUSCULOSKELETAL	-.0398	.4155	-.0959	
23	TRAUMATIC	-.0731	.4392	-.1664	
24	OTHER, UNKNOWN	.1896	.4507	.4207	
	PIA				
25	300-399	.0055	.1147	.0476	
26	400-499	.0186	.1199	.1552	
27	500-599	.1061	.1298	.8177	
28	600 AND OVER	.2188	.1449	1.5098	

Table 4. continued

	Education				
29	1-8 YEARS	1.0234	.2562	3.9940	
30	9-12 YEARS	1.1609	.2618	4.4338	.2010
31	13 AND OVER	1.1310	.2652	4.2652	.2927
32	OTHER AND UNKNOWN	1.3658	.3456	3.9516	.3730
	ESRD				
33	ESRD	3.4009	.3932	8.6501	3.4760
	Age				
34	30-39	.2989	.3323	.8995	
35	40-49	-.1445	.3412	-.4235	
36	50-59	-.0952	.3365	-.2828	
37	60 AND OVER	-.0298	.3396	-.0877	
	Time to Death				
38	YR OF DEATH	.7220	.2166	3.3337	.4369
39	1 YR BEFORE DEATH	.5198	.2200	2.3633	.2601
40	2 YRS BEFORE DEATH	.4439	.2229	1.9919	.1789
41	3 YRS BEFORE DEATH	.3695	.2225	1.6601	
42	4 YRS BEFORE DEATH	.2144	.2244	.9553	
43	5 YRS BEFORE DEATH	.2431	.2196	1.1069	
44	6 YRS BEFORE DEATH	.4696	.2150	2.1844	
	Census Division				
45	DIVISION 2	-.0865	.1500	-.5764	-.0610
46	DIVISION 3	-.2555	.1567	-1.6301	-.2676
47	DIVISION 4	-.1948	.2471	-.7885	-.1652
48	DIVISION 5	-.1634	.1590	-1.0282	-.1611
49	DIVISION 6	-.4976	.1753	-2.8381	-.5421
50	DIVISION 7	-.3933	.1599	-2.4601	-.3412
51	DIVISION 8	-.3823	.2185	-1.7500	-.3029
52	DIVISION 9	.2999	.1563	1.9182	.3206
	Months				
53	MTH IN YR OF EVENT	-.0871	.0108	-8.0306	-.0923
	NUMBER OF OBSERVATIONS	2339			
	NUMBER OF CASES	1811			
	STANDARD ERROR	1.3420			1.3610
	ADJUSTED R2	.2242			.2014
	PHI - SMEARING				2.7706
	PHI - NORMAL				2.5248

Table 5. Logistic regression coefficients, Medicare use, recovery cases

VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
CONSTANT	-2.5232	.4731	-5.3332	-2.6194
Sex				
1 FEMALE	.4448	.1821	2.4428	.5764
Race				
2 BLACK	-.0554	.2190	-.2531	-.1044
Occupation				
4 BLUE COLLAR	-.0851	.1925	-.4422	
Diagnosis				
15 MENTAL	-.4439	.2422	-1.8324	
18 CIRCULATORY	.3649	.2509	1.4544	
22 MUSCULOSKELETAL	.2084	.2290	.9102	
23 TRAUMATIC	-.2278	.2472	-.9217	
PIA				
25 300-399	-.1920	.1980	-.9694	
26 400-499	-.2362	.2405	-.9823	
27 500-599	.0004	.2686	.0015	
28 600 AND OVER	-.2824	.3074	-.9186	
Education				
30 9-12 YEARS	.1550	.2485	.6238	
31 13 AND OVER	.2676	.2199	1.2168	
32 OTHER AND UNKNOWN	-.0302	.2928	-.1031	
ESRD				
33 ESRD	1.4790	.8402	1.7602	
Age				
34 30-39	-.0473	.1926	-.2454	
35 40-49	.0124	.2107	.0591	
36 50 AND OVER	-.4616	.2320	-1.9898	-.3720

Table 5. continued

Time to recovery				
38 YR OF RECOV	-.8854	.1933	-4.5808	-.6972
39 1 YR BEFORE RECOV	-.7667	.1897	-4.0408	-.5854
40 2 YRS BEFORE RECOV	-.4778	.1900	-2.5148	-.2958
41 3 YRS BEFORE RECOV	-.5747	.1805	-3.1830	-.4037
42 4 YRS BEFORE RECOV	-.2098	.1829	-1.1470	
Census division				
45 DIV 2	.4330	.3611	1.1990	
46 DIV 3	.0407	.3491	.1167	
47 DIV 4	.0674	.4170	.1616	
48 DIV 5	-.3016	.3544	-.8508	
49 DIV 6	.3844	.3923	.9799	
50 DIV 7	-.2664	.3638	-.7324	
51 DIV 8	.3061	.4166	.7349	
52 DIV 9	.3760	.3365	1.1173	
Months				
53 MTH IN YR OF EVENT	.1894	.0145	13.0364	.1840
NUMBER OF CASES	2777			
NUMBER OF OBSERVATIONS	803			

Table 6. Logistic regression coefficients, hospital use recovery cases

VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
CONSTANT	.1711	.7282	.2350	-.4358
Sex				
1 FEMALE	-.2805	.2279	-1.2308	-.3319
Race				
2 BLACK	-.6187	.3156	-1.9603	-.3037
Occupation				
4 BLUE COLLAR	-.2931	.2603	-1.1257	
Diagnosis				
15 MENTAL	-.1422	.3508	-.4055	
18 CIRCULATORY	-.6453	.3400	-1.8979	
22 MUSCULOSKELETAL	-.4578	.3051	-1.5004	
23 TRAUMATIC	-.0082	.3051	-.0267	
PIA				
25 300-399	.0340	.2648	.1283	
26 400-499	.0350	.3095	.1132	
27 500-599	.1942	.3468	.5598	
28 600 AND OVER	-.4043	.3708	-1.0904	
Education				
30 9-12 YEARS	.7101	.3402	2.0870	.4616
31 13 AND OVER	.5209	.2977	1.7496	.4096
32 OTHER AND UNKNOWN	.7778	.3890	1.9995	.6163
ESRD				
33 ESRD	1.2571	.4629	2.7157	1.2700
Age				
34 30-39	.0493	.2799	.1763	
35 40-49	.3825	.2900	1.3190	
36 50 AND OVER	-.3719	.2927	-1.2705	

Table 6. continued

Time to recovery			
38 YR OF RECOV	-.3905	.3146	-1.2413
39 1 YR BEFORE RECOV	-.3951	.3040	-1.2996
40 2 YRS BEFORE RECOV	-.1981	.3131	-.6325
41 3 YRS BEFORE RECOV	-.2164	.2717	-.7966
42 4 YRS BEFORE RECOV	-.1273	.3001	-.4243
Census division			
45 DIV 2	-.1173	.4705	-.2493
46 DIV 3	-.6263	.4673	-1.3403
47 DIV 4	.0920	.5991	.1535
48 DIV 5	-.3079	.4881	-.6308
49 DIV 6	.6715	.5348	1.2556
50 DIV 7	.4381	.5080	.8624
51 DIV 8	-.9141	.5549	-1.6474
52 DIV 9	-.3837	.4594	-.8353
Months			
53 MTH IN YR OF EVENT	.0091	.0384	.2378
NUMBER OF OBSERVATIONS	664		
NUMBER OF CASES	268		

Table 7. Regression coefficients for log reimbursement, inpatient hospital stay, recovery cases

VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
CONSTANT	6.2449	.5173	12.0720	6.7028
Sex				
1 FEMALE	.5123	.2027	2.5275	.3934
Race				
2 BLACK	.5943	.2507	2.3707	.3075
Occupation				
4 BLUE COLLAR	-.4391	.1989	-2.2077	-.5846
Diagnosis				
15 MENTAL	-.0535	.3310	-.1615	
18 CIRCULATORY	-.2597	.2923	-.8885	
22 MUSCULOSKELETAL	.1858	.2879	.6453	
23 TRAUMATIC	.3177	.2414	1.3159	
PIA				
25 300-399	.2423	.2161	1.1215	
26 400-499	.1362	.2277	.5983	
27 500-599	.5918	.2524	2.3447	.3821
28 600 AND OVER	.8929	.2906	3.0730	.6501
Education				
30 9-12 YEARS	.1698	.2954	.5749	
31 13 AND OVER	-.0765	.2283	-.3351	
32 OTHER AND UNKNOWN	.0726	.2740	.2651	
ESRD				
33 ESRD	1.3858	.6343	2.1849	1.3023
Age				
34 30-39	.0159	.2268	.0699	
35 40-49	-.0605	.2354	-.2571	
36 50 AND OVER	-.0783	.2597	-.3016	

Table 7. continued

Time to recovery				
38 YR OF RECOV	.1897	.2165	.8762	
39 1 YR BEFORE RECOV	.0661	.2294	.2881	
40 2 YRS BEFORE RECOV	.2034	.1780	1.1424	
41 3 YRS BEFORE RECOV	.2728	.2058	1.3257	
42 4 YRS BEFORE RECOV	-.0552	.2729	-.2021	
Census division				
45 DIV 2	.1440	.3502	.4112	
46 DIV 3	.0301	.3749	.0804	
47 DIV 4	-.0832	.3166	-.2629	
48 DIV 5	.0460	.3926	.1171	
49 DIV 6	.0600	.2980	.2014	
50 DIV 7	-.3925	.3595	-1.0918	
51 DIV 8	.0899	.3815	.2356	
52 DIV 9	.3803	.3006	1.2653	
Months				
53 MTH IN YR OF EVENT	-.1348	.0276	-4.8861	-.1303
NUMBER OF OBSERVATIONS	303			
NUMBER OF CASES	171			
STANDARD ERROR	1.1778			1.1774
ADJUSTED R2	.1667			.1674
PHI - SMEARING				1.7755
PHI - NORMAL				2.0000

Table 8. Regression coefficients for log reimbursement, no inpatient hospital stay, recovery cases

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	3.4469	.5009	6.8809	2.5009
	Sex				
1	FEMALE	-.1162	.1827	-.6359	
	Race				
2	BLACK	.1766	.2024	.8727	
	Occupation				
4	BLUE COLLAR	-.0581	.1903	-.3050	
	Diagnosis				
15	MENTAL	-.7787	.2895	-2.6901	-.6340
18	CIRCULATORY	-.0615	.2705	-.2275	
22	MUSCULOSKELETAL	-.0885	.2146	-.4122	
23	TRAUMATIC	-.0937	.2492	-.3760	
	PIA				
25	300-399	-.0491	.1880	-.2611	
26	400-499	-.0238	.2286	-.1043	
27	500-599	.1747	.2648	.6599	
28	600 AND OVER	.2064	.2482	.8316	
	Education				
30	9-12 YEARS	.0732	.3066	.2387	
31	13 AND OVER	-.0979	.1832	-.5345	
32	OTHER AND UNKNOWN	.2832	.2682	1.0558	
	ESRD				
33	ESRD	2.5493	.4844	5.2630	2.0759
	Age				
34	30-39	-.1474	.2314	-.6371	
35	40-49	.0969	.2299	.4214	
36	50 AND OVER	.1817	.2630	.6908	

Table 8. continued

Time to recovery				
38 YR OF RECOV	-.1677	.2073	-.8092	
39 1 YR BEFORE RECOV	-.4269	.2038	-2.0950	
40 2 YRS BEFORE RECOV	-.1515	.1959	-.7736	
41 3 YRS BEFORE RECOV	.0096	.2056	.0469	
42 4 YRS BEFORE RECOV	.0279	.1863	.1495	
Census division				
45 DIV 2	.0333	.3799	.0876	
46 DIV 3	-.5067	.3581	-1.4152	
47 DIV 4	-.7391	.4966	-1.4884	
48 DIV 5	-.3427	.3410	-1.0051	
49 DIV 6	-.6718	.3984	-1.6862	
50 DIV 7	-.5133	.3538	-1.4508	
51 DIV 8	-.5904	.3704	-1.5940	
52 DIV 9	.0261	.3511	.0743	
Months				
53 MTH IN YR OF EVENT	-.0391	.0298	-1.3137	
NUMBER OF OBSERVATIONS	361			
NUMBER OF CASES	199			
STANDARD ERROR	1.1890			1.2000
ADJUSTED R2	.0725			.0552
PHI - SMEARING				2.8770
PHI - NORMAL				2.0544

Table 9. Logistic regression coefficients, Medicare use age 65/still in program*

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	-2.1072	.2319	-9.0882	-1.8420
	Sex				
1	FEMALE	.6211	.0420	14.7991	.6146
	Race				
2	BLACK	-.0747	.0469	-1.5925	-.0771
3	OTHER	-.1801	.1616	-1.1145	
	Occupation				
4	CLERICAL	-.0863	.0706	-1.2212	-.0936
5	SERVICE	-.1758	.0713	-2.4661	-.1793
6	FARM	-.2075	.0981	-2.1152	-.2225
7	PROCESSING	-.2102	.1061	-1.9806	-.2151
8	MACHINE	-.1568	.0758	-2.0697	-.1579
9	BENCH WORK	-.1929	.0829	-2.3269	-.2000
10	STRUCTURAL	-.2064	.0738	-2.7973	-.2117
11	MISCELLANEOUS	-.2153	.0695	-3.0983	-.2200
12	UNKNOWN	-.1334	.0903	-1.4776	-.1366
	Diagnosis				
13	NEOPLASS	.4045	.1618	2.5002	.3005
14	ENDOCRINE	.4931	.1448	3.4056	.3820
15	MENTAL	.0964	.1286	.7498	
16	NERVOUS	.1262	.1413	.8929	
17	EYE AND EAR	-.1195	.1561	-.7657	
18	CIRCULATORY	.4522	.1233	3.6685	.3434
19	RESPIRATORY	.3403	.1357	2.5071	.2250
20	DIGESTIVE	.3061	.1611	1.9004	.1937
21	GENIT-URINARY	.4998	.2331	2.1439	.4025
22	MUSCUOSKELETAL	.1598	.1245	1.2842	
23	TRAUMTIC	.0241	.1347	.1792	
24	OTHER,UNKNOWN	.3421	.1430	2.3930	.2692
	PIA				
25	300-399	-.0180	.0518	-.3479	
26	400-499	-.0655	.0556	-1.1794	
27	500-599	-.0271	.0574	-.4715	
28	600 AND OVER	.0570	.0684	.8334	

Table 9. continued

Education				
29 1-8 YEARS	.2964	.1428	2.0758	.1732
30 9-12 YEARS	.3645	.1455	2.5057	.2511
31 13 AND OVER	.2565	.1460	1.7566	.1490
32 OTHER AND UNKNOWN	.1481	.1643	.9014	
ESRD				
33 ESRD	1.8889	.4457	4.2379	1.9098
Age				
34 30-39	.1430	.0922	1.5508	
35 40-49	.2154	.0957	2.2505	.1714
36 50-59	.1791	.0921	1.9453	.1410
37 60 AND OVER	.0145	.0951	.1527	
Time to year 1981 or age 65				
38 1981 or AGE 65	1.6309	.0391	41.6659	1.6116
39 1 YR BEFORE	1.0497	.0408	25.7473	1.0350
40 2 YRS BEFORE	.9032	.0396	22.8276	.8899
41 3 YRS BEFORE	.8360	.0390	21.4307	.8257
42 4 YRS BEFORE	.7162	.0382	18.7581	.7088
43 5 YRS BEFORE	.6229	.0367	16.9771	.6204
44 6 YRS BEFORE	.4612	.0342	13.5013	.4626
Census Division				
45 DIVISION 2	-.0873	.0816	-1.0702	-.0932
46 DIVISION 3	-.1842	.0806	-2.2860	-.1866
47 DIVISION 4	-.2622	.0949	-2.7625	-.2664
48 DIVISION 5	-.1656	.0804	-2.0591	-.1681
49 DIVISION 6	-.2529	.0879	-2.8774	-.2537
50 DIVISION 7	-.2652	.0870	-3.0498	-.2715
51 DIVISION 8	-.1815	.0998	-1.8181	-.1912
52 DIVISION 9	.2133	.0834	2.5581	.2048
Months				
53 MTH IN YR OF EVENT	.0656	.0025	26.1799	.0649
NUMBER OF OBSERVATIONS	62002			
NUMBER OF CASES	9561			

* One seventh of the cases were excluded at random from the analysis due to software limitations.

Table 10. Logistic regression coefficients, hospital use age 65/still in the program

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	-.2155	.2599	-.8293	-.1702
	Sex				
1	FEMALE	-.1937	.0403	-4.8116	-.1892
	Race				
2	BLACK	-.3321	.0473	-7.0173	-.3017
3	OTHER	-.0766	.1583	-.4839	
	Occupation				
4	CLERICAL	.0016	.0645	.0248	
5	SERVICE	.0371	.0664	.5584	
6	FARM	-.1110	.0978	-1.1344	
7	PROCESSING	.1798	.1009	1.7826	
8	MACHINE	-.0605	.0735	-.8237	
9	BENCH WORK	-.0054	.0781	-.0688	
10	STRUCTURAL	.0355	.0693	.5118	
11	MISCELLANEOUS	.0345	.0661	.5219	
12	UNKNOWN	.0324	.0880	.3686	
	Diagnosis				
13	NEOPLASS	.0454	.1868	.2430	
14	ENDOCRINE	.3971	.1687	2.3539	.2408
15	MENTAL	.0253	.1584	.1596	
16	NERVOUS	.0899	.1667	.5392	
17	EYE AND EAR	-.0413	.1792	-.2304	
18	CIRCULATORY	.2379	.1535	1.5504	
19	RESPIRATORY	.3053	.1634	1.8688	.1460
20	DIGESTIVE	.4797	.1813	2.6460	.3225
21	GENIT-URINARY	.0709	.2427	.2923	
22	MUSCUOSKELETAL	.1616	.1552	1.0413	
23	TRAUMTIC	.1295	.1633	.7928	
24	OTHER,UNKNOWN	.2424	.1667	1.4541	
	PIA				
25	300-399	-.0434	.0486	-.8923	
26	400-499	-.0795	.0525	-1.5124	
27	500-599	-.1790	.0556	-3.2201	-.1279
28	600 AND OVER	-.3845	.0654	-5.8789	-.3238

Table 10. continued

Education				
29 1-8 YEARS	.1006	.1552	.6477	
30 9-12 YEARS	.1507	.1575	.9571	
31 13 AND OVER	.0425	.1582	.2687	
32 OTHER AND UNKNOWN	.0186	.1760	.1056	
ESRD				
33 ESRD	.9269	.2158	4.2949	.9238
Age				
34 30-39	-.0754	.0981	-.7689	
35 40-49	-.1394	.0985	-1.4143	
36 50-59	-.3200	.0945	-3.3845	-.1757
37 60 AND OVER	-.3625	.0973	-3.7257	-.2074
Time to year 1981 or age 65				
38 1981 or AGE 65	-.1378	.0581	-2.3702	-.1200
39 1 YR BEFORE	-.1557	.0586	-2.6571	-.1339
40 2 YRS BEFORE	-.1859	.0588	-3.1618	-.1618
41 3 YRS BEFORE	-.2000	.0589	-3.3967	-.1750
42 4 YRS BEFORE	-.1365	.0593	-2.3018	-.1103
43 5 YRS BEFORE	-.1718	.0591	-2.9063	-.1413
44 6 YRS BEFORE	-.0515	.0583	-.8847	
Census Division				
45 DIVISION 2	-.0077	.0789	-.0977	.0007
46 DIVISION 3	.3358	.0788	4.2598	.3439
47 DIVISION 4	.3949	.0938	4.2119	.3991
48 DIVISION 5	.2843	.0790	3.6011	.2942
49 DIVISION 6	.5562	.0873	6.3729	.5649
50 DIVISION 7	.4768	.0857	5.5610	.4846
51 DIVISION 8	.1650	.1046	1.5768	.1631
52 DIVISION 9	-.1980	.0803	-2.4648	-.2015
Months				
53 MTH IN YR OF EVENT	.0112	.0037	3.0094	.0106
NUMBER OF OBSERVATIONS	36052			
NUMBER OF CASES	9059			

Table 11. Regression coefficients for log reimbursement, inpatient hospital stay, age 65/still in the program

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	6.0873	.2180	27.9203	5.9580
	Sex				
1	FEMALE	.1263	.0314	4.0255	.1105
	Race				
2	BLACK	.1926	.0363	5.3125	.1584
3	OTHER	.1262	.0907	1.3911	
	Occupation				
4	CLERICAL	-.0086	.0500	-.1724	
5	SERVICE	-.0527	.0527	-1.0003	
6	FARM	-.1814	.0786	-2.3066	-.1725
7	PROCESSING	-.0737	.0726	-1.0144	
8	MACHINE	-.0775	.0549	-1.4121	
9	BENCH WORK	-.1111	.0572	-1.9427	-.0793
10	STRUCTURAL	-.1378	.0547	-2.5191	-.1006
11	MISCELLANEOUS	-.1386	.0505	-2.7455	-.1113
12	UNKNOWN	.0273	.0675	.4052	
	Diagnosis				
13	NEOPLASS	.2022	.1372	1.4731	
14	ENDOCRINE	.2795	.1237	2.2601	.2426
15	MENTAL	.0254	.1163	.2187	
16	NERVOUS	.0863	.1263	.6830	
17	EYE AND EAR	.0327	.1355	.2411	
18	CIRCULATORY	.0650	.1125	.5777	
19	RESPIRATORY	.1120	.1205	.9298	
20	DIGESTIVE	.0990	.1314	.7535	
21	GENIT-URINARY	.4409	.2003	2.2007	.4187
22	MUSCUOSKELETAL	.0127	.1139	.1119	
23	TRAUMTIC	.0437	.1204	.3631	
24	OTHER,UNKNOWN	.0091	.1236	.0735	
	PIA				
25	300-399	.0223	.0369	.6031	
26	400-499	.0887	.0406	2.1829	
27	500-599	.0655	.0417	1.5685	
28	600 AND OVER	.0413	.0531	.7768	

Table 11. continued

Education				
29 1-8 YEARS	-.1558	.1475	-1.0564	
30 9-12 YEARS	-.0697	.1487	-.4683	
31 13 AND OVER	-.0363	.1492	-.2434	
32 OTHER AND UNKNOWN	-.1753	.1599	-1.0963	
ESRD				
33 ESRD	1.7654	.1500	11.7654	1.7713
Age				
34 30-39	-.0720	.0855	-.8419	
35 40-49	-.1065	.0861	-1.2357	
36 50-59	-.0883	.0836	-1.0563	
37 60 AND OVER	-.0211	.0856	-.2469	
Time to year 1981 or age 65				
38 1981 or AGE 65	-.1058	.0489	-2.1645	-.0609
39 1 YR BEFORE	-.0062	.0480	-.1284	.0184
40 2 YRS BEFORE	-.0808	.0481	-1.6777	-.0605
41 3 YRS BEFORE	-.1187	.0480	-2.4742	-.1042
42 4 YRS BEFORE	-.1733	.0488	-3.5507	-.1653
43 5 YRS BEFORE	-.1667	.0483	-3.4524	-.1744
44 6 YRS BEFORE	-.2028	.0478	-4.2451	-.2148
Census Division				
45 DIVISION 2	-.1296	.0592	-2.1900	-.1208
46 DIVISION 3	-.1424	.0581	-2.4515	-.1391
47 DIVISION 4	-.3635	.0703	-5.1742	-.3641
48 DIVISION 5	-.4100	.0582	-7.0481	-.4221
49 DIVISION 6	-.5224	.0665	-7.8572	-.5436
50 DIVISION 7	-.5205	.0642	-8.1068	-.5284
51 DIVISION 8	-.3646	.0787	-4.6345	-.3403
52 DIVISION 9	-.0282	.0597	-.4724	-.0117
Months				
53 MTH IN YR OF EVENT	-.0255	.0032	-8.0439	-.0228
NUMBER OF OBSERVATIONS	15439			
NUMBER OF CASES	6624			
STANDARD ERROR	1.0825			
ADJUSTED R2	.0848			.0805
PHI - SMEARING				1.7140
PHI - NORMAL				1.8017

Table 12. Regression coefficients for log reimbursement, no inpatient hospital stay, age 65/still in the program

	VARIABLE	COEFF	STAND ERROR	t-VALUE	PREDICT COEFF
	CONSTANT	2.6002	.1995	13.0348	2.7967
	Sex				
1	FEMALE	.1130	.0306	3.6889	.1229
	Race				
2	BLACK	.1495	.0357	4.1897	.1246
3	OTHER	.0592	.1013	.5841	
	Occupation				
4	CLERICAL	.0757	.0497	1.5224	
5	SERVICE	-.0745	.0502	-1.4831	
6	FARM	-.2050	.0729	-2.8126	-.2014
7	PROCESSING	-.1807	.0765	-2.3615	-.1691
8	MACHINE	-.1240	.0551	-2.2518	-.1161
9	BENCH WORK	-.1741	.0574	-3.0339	-.1663
10	STRUCTURAL	-.1122	.0526	-2.1319	-.1004
11	MISCELLANEOUS	-.0672	.0504	-1.3329	
12	UNKNOWN	.0669	.0763	.8765	
	Diagnosis				
13	NEOPLASS	.2940	.1477	1.9904	.2910
14	ENDOCRINE	.0309	.1297	.2381	
15	MENTAL	.0227	.1163	.1954	
16	NERVOUS	.0094	.1264	.0745	
17	EYE AND EAR	-.0378	.1311	-.2880	
18	CIRCULATORY	-.0069	.1129	-.0612	
19	RESPIRATORY	.1799	.1240	1.4513	
20	DIGESTIVE	.1026	.1414	.7255	
21	GENIT-URINARY	.2276	.1899	1.1986	
22	MUSCUOSKELETAL	-.0332	.1137	-.2920	
23	TRAUMTIC	-.0179	.1238	-.1443	
24	OTHER,UNKNOWN	.0429	.1327	.3232	
	PIA				
25	300-399	.0869	.0364	2.3859	
26	400-499	.0105	.0397	.2636	
27	500-599	.0606	.0418	1.4521	
28	600 AND OVER	-.0252	.0496	-.5093	

Table 12. continued

Education				
29 1-8 YEARS	.2045	.1179	1.7349	-.0233
30 9-12 YEARS	.2796	.1194	2.3408	.0549
31 13 AND OVER	.3103	.1199	2.5875	.1004
32 OTHER AND UNKNOWN	.1373	.1389	.9885	
ESRD				
33 ESRD	3.9502	.3082	12.8155	4.0508
Age				
34 30-39	-.0617	.0942	-.6547	
35 40-49	-.0918	.0968	-.9477	
36 50-59	-.0660	.0947	-.6969	
37 60 AND OVER	-.0752	.0970	-.7756	
Time to year 1981 or age 65				
38 1981 or AGE 65	-.1786	.0492	-3.6279	-.1870
39 1 YR BEFORE	-.1087	.0493	-2.2034	-.1148
40 2 YRS BEFORE	-.0992	.0492	-2.0182	-.1020
41 3 YRS BEFORE	-.1006	.0488	-2.0606	-.1033
42 4 YRS BEFORE	-.0821	.0496	-1.6554	-.0833
43 5 YRS BEFORE	-.0972	.0493	-1.9729	-.0971
44 6 YRS BEFORE	-.1072	.0503	-2.1298	-.1049
Census Division				
45 DIVISION 2	.0803	.0572	1.4041	.0776
46 DIVISION 3	-.1066	.0575	-1.8545	-.1133
47 DIVISION 4	-.3625	.0684	-5.2996	-.3774
48 DIVISION 5	-.2264	.0568	-3.9885	-.2381
49 DIVISION 6	-.4596	.0630	-7.3003	-.4663
50 DIVISION 7	-.2827	.0639	-4.4207	-.3010
51 DIVISION 8	-.1836	.0703	-2.6118	-.1908
52 DIVISION 9	.2757	.0564	4.8893	.2625
Months				
53 MTH IN YR OF EVENT	-.0204	.0029	-6.9571	-.0207
NUMBER OF OBSERVATIONS	20613			
NUMBER OF CASES	6459			
STANDARD ERROR	1.1952			1.1970
ADJUSTED R2	.0962			.0928
PHI - SMEARING				2.0547
PHI - NORMAL				2.0471