Projected Outcomes and Length of Time in the Disability Insurance Program

by John C. Hennessey
and Janice M. Dykacz*

This article reports on the first phase of a longitudinal analysis of a random sample of 18,816 Social Security beneficiaries who were first entitled to disabled-worker benefits in 1972. These individuals were observed from the time of benefit entitlement to January 1981. This phase examines the first event of interest after entitlement—that is, recovery, death, or retirement. Mathematical models are used to project these events beyond the observation period and to calculate the proportion of beneficiaries who ultimately leave the program rolls for these reasons. Average length of time in the program is also estimated. The analysis relates these outcomes to a set of covariates that includes primary diagnosis, educational level, past occupation, primary insurance amount (PIA), sex, race, and age at entitlement. The second phase of the project analyzes the postrecovery period of the same cohort. The results appear in the second article of this issue.

This study projects that 11 percent of the beneficiaries will ultimately leave the Disability Insurance program because they recover, 36 percent will have their disabled-worker benefit terminated at death, and 53 percent will have their benefits converted to retired-worker benefits at age 65—that is, they retire. Mean length of time in the program is estimated to be 9.3 years. Considerable variation was found in the outcomes and mean number of years in the program by primary diagnosis, educational level, PIA, age, and sex.

*Division of Disability Program Information and Studies, Office of Disability, Social Security Administration. The authors would like to thank Barry V. Bye for his expert guidance and Joseph G. Lindamood, Jr., and Audrey B. Coe for explaining the Continuing Disability Review process.

This article presents a detailed longitudinal analysis of a random sample of Social Security beneficiaries who were first entitled to Disability Insurance (DI) benefits in 1972. Mathematical models are used to project events such as recovery, death, or retirement beyond the observation period and to calculate the proportions of disabled workers whose benefits were ultimately terminated for these reasons. Average length of time beneficiaries remain in the program is also estimated. The analysis relates these outcomes to a set of covariates that includes the primary diagnosis of the disabling condition; educational level; type of past occupation; primary insurance amount (PIA); and the demographic variables of sex, race, and age at entitlement.

Information on outcomes and length of time in the DI program is of particular interest because it describes the general composition of the disabled-worker population entitled in 1972, and has a direct bearing on program costs. This study pioneers the collection and in-depth analysis of such information. Studies have been done on Medicare utilization using this same cohort [1], [2]. A further study is planned that will combine the results derived in this article and in [3], [1], and [2] to estimate lifetime Medicare costs.

Analysis Plan

Disabled-worker benefits may be terminated when the beneficiary has a medical recovery or work recovery, dies, or attains age 65 (when an automatic conversion from the DI program to the retirement program occurs). This study estimates the proportion of disabled workers whose benefits were terminated during their first period in the DI program. The average length of time in the first benefit entitlement period is also estimated. For those whose benefits were not terminated, survival analysis techniques were used to project beyond the observation period. Mathematical models of the tendency to recover or to die at a given point in time before retirement are estimated separately as functions of the covariates. For each combination of covariate values, the models are used to estimate the probabilities of benefit termination because of recovery or death before attaining age 65. The probabilities are then used to compute the average length of time in the program for each covariate combination. The proportions of projected terminations based on recovery, death, or retirement for the 1972 cohort are computed overall. The Technical Appendix details the techniques used in this study.

Description of Data File

The data file used for this study is based on a 5-percent random sample—23,062 disabled-worker beneficiaries first entitled under the Disability Insurance program in 1972—of whom 18,816 were used in the analysis. It is assumed that the observation period ends on January 1, 1981, because substantial administrative changes in the disability review process and court intervention regarding benefit terminations began about this time. Although data were available for the period 1981-86, they were excluded from the analysis and the 1972 model was developed as though no atypical intervention had taken place. This censored observation period provides a more realistic representation of the "pre-1980's" program on which to base long-run estimates of program duration and outcomes.

Distribution of Covariates

The distribution of covariates for the 18,816 persons in the data file used in the analysis is shown in table 1. The following sections

Table 1.—Distribution of covariates, by sex and age at entitlement

Covariate	Total	Sex							
		Men				Women			
		Age at entitlement				Age at entitlement			
		Total	18-34	35-49	50-61	Total	18-34	35-49	50-61
Diagnostic group									
Total	18,816	13,139	1,596	3,146	8,397	5,677	502	1,300	3,875
Infective	326	258	30	102	126	68	9	25	34
Neoplasms	1,525	955	66	229	660	570	39	144	387
Endocrine	620	390	30	88	272	230	23	44	163
Mental disorders	1,752	1,217	386	378	453	535	117	170	248
Nervous system	1,085	704	124	177	403	381	73	95	213
Circulatory	5,389	3,978	67	752	3159	1,411	28	232	1,151
Respiratory	1,152	952	10	126	816	200	5	34	161
Digestive	545	398	24	127	247	147	10	43	94
Genitourinary	170	109	25	34	50	61	3	22	36
Musculoskeletal	2,917	1,832	193	468	1,171	1,085	67	223	795
Congenital anomalies	198	131	39	43	49	67	10	27	30
Accidents	1,276	1,023	370	300	353	253	41	62	150
Other	1,861	1,192	232	322	638	669	77	179	413
Years of education									
Total	18,816	13,139	1,596	3,146	8,397	5,677	502	1,300	3,875
0-8	6,798	5,204	252	1,111	3,841	1,594	39	262	1,293
9-11	3,949	2,671	337	721	1,613	1,278	72	321	885
12	4,330	2,810	530	695	1,585	1,520	220	392	908
13 or more	1,495	996	186	253	557	499	75	121	303
Unknown	2,244	1458	291	366	801	786	96	204	486
Primary insurance amount									
Total	18,816	13,139	1,596	3,146	8,397	5,677	502	1,300	3,875
$1-$299	3,041	1,262	292	344	626	1,779	118	469	1,192
$300-$499	8,117	5,010	785	1,301	2,924	3,107	281	705	2,121
$500-$699	7,416	6,646	383	1,479	4,784	770	88	125	557
$700 or more	242	221	136	22	63	21	15	1	5
Occupation									
Total	18,816	13,139	1,596	3,146	8,397	5,677	502	1,300	3,875
White collar	4,092	2,211	205	494	1,512	1,881	187	445	1,249
Service	2,690	1,188	192	275	721	1,502	92	280	1,130
Farming	763	719	57	129	533	44	5	9	30
Manufacturing	5,613	4,487	440	1,110	2,937	1,126	98	266	762
Unknown and miscellaneous	5,658	4,534	702	1,138	2,694	1,124	120	300	704
Race									
Total	18,816	13,139	1,596	3,146	8,397	5,677	502	1,300	3,875
Nonblack	16,159	11,299	1,339	2,628	7,332	4,860	429	1,083	3,348
Black	2,657	1,840	257	518	1,065	817	73	217	527

describe these covariates and their distribution in the sample.

Age, Sex, and Race

About 70 percent of individuals in the sample are male and 86 percent are nonblack. At the time of entitlement (1972), 65 percent of the sample were aged 50-61; 24 percent were aged 35-49; and 11 percent were aged 34 or younger (the youngest group). Beneficiaries aged 62 or older at the time of entitlement were excluded from the analysis because the available data do not distinguish disabled-worker beneficiaries from retired-worker beneficiaries.

Primary Diagnosis

One of the most important covariates in this study is the primary diagnosis of the disabling condition—that is, the primary underlying medical impairment on which the original entitlement decision was based. The diagnostic groups in this study were taken from the International Classification of Diseases [4]. They are

- Infective and parasitic diseases;
- Neoplasms;
- Endocrine, nutritional, and metabolic diseases;
- Mental disorders;
- Diseases of the—
  - Nervous system and sense organs;
  - Circulatory system;
  - Respiratory system;
  - Digestive system;
  - Genitourinary system;
  - Musculoskeletal system and connective tissue;
- Congenital anomalies;
- Accidents, poisonings, and violence; and
- Other (includes diseases of blood and blood forming organs and diseases of the skin and subcutaneous tissue).

Primary Insurance Amount

The primary insurance amount is the dollar figure on which cash benefits are based. It is a function of the number of years of covered earnings under the Social Security program before the onset of disability and the level of earnings for those years. It serves as a rough proxy for the level of lifetime earnings. The PIA also gives a rough indication of economic status because it is directly related to the cash benefit received.

The PIA at the time of entitlement in 1972 was not available. Therefore, this study uses the PIA in effect on December 31, 1985. This PIA reflects changes over time in response to adjustments for inflation, legislation, and additional work performed after a recovery termination. The PIA was categorized into four levels: $1-$299; $300-$499; $500-$699; and $700 or more.

Education

Beneficiaries are classified according to educational level attained at time of entitlement. Thirty-six percent had not gone beyond eighth grade; 21 percent completed 9-11 years; 23 percent graduated from high school; and 8 percent have some years of college. For 12 percent, educational attainment is unknown.

Occupation

For most persons in the sample, the occupation reported is the major occupation in the 15-year period preceding their application for disabled-worker benefits. The Dictionary of Occupational Titles (DOT) 1965 codes [5] are used to coarsely group occupations into white collar (codes 1-29), service (codes 30-38), farming (codes 40-46), manufacturing (codes 50-89), and unknown and miscellaneous. Twenty-two percent are classified white collar, 14 percent service, 4 percent farming, 30 percent manufacturing, and 30 percent unknown or miscellaneous.

Understanding the Termination Process

To understand how a covariate affects the proportions of beneficiaries who ultimately leave the DI program rolls because they recovered, died, or retired, it is necessary to develop an understanding of the complex termination process. The first focus of the analysis is on the process at an instant in time after entitlement and before the beneficiary attains age 65, at which time the beneficiary is at risk of termination from the program based on either recovery or death. Then, the retirement event at age 65 is taken into account.

At a point in time after entitlement, a beneficiary who is younger than age 65 could possibly leave the program because of recovery or death. Thus, a "recovery tendency" and a "tendency for death" at this instant in time compete with each other. If both tendencies are very low, the probability of a program termination of either type about this time will tend to be low. If both tendencies are high, the probability of a termination of either type will tend to be high. The strength of the two tendencies relative to each other will dictate which of the two probabilities is larger.

The value of a covariate, such as age, at this instant in time (after entitlement and before age 65) will affect the tendency toward a program termination because of

recovery or death. The probability of a recovery termination about this time depends on the net effect of the covariate on both tendencies. The same is true for the probability of a death.

For example, it is expected that the age of a beneficiary affects both tendencies. The recovery tendency is expected to be lower for an older beneficiary than for a younger beneficiary. Similarly, the death tendency would be higher for the older beneficiary. Therefore, it is not clear whether the probability of a program termination by recovery or death about this time will be high or low. The net effect of the covariate on the two tendencies determines the effect on the probability.

The strength of the recovery and death tendencies over the time period from entitlement to age 65 together with the length of this time period will determine the percentage of beneficiaries who ultimately leave the program rolls because of recovery, death, or retirement. If both tendencies are strong over the period, the probability of a retirement termination is small because most beneficiaries will recover or die. Conversely, if both tendencies are weak, the probability of a retirement termination will be high. Also, given the same tendencies to die or to recover, the longer the period from entitlement to age 65 the greater the chance of a recovery or death because more time exists for such program terminations to occur. Thus, the three factors—age at entitlement and the two tendencies—must be examined simultaneously to compute the probabilities of ultimately leaving the DI program rolls because of recovery, death, or retirement.

Older beneficiaries, for example, have less time between entitlement and age 65. This factor together with the increase in the death tendency and the decrease in the recovery tendency relative to the younger beneficiaries determines the effect of age on the proportions of beneficiaries who ultimately recover, die, or retire. In fact, it will be shown that the probability of a termination due to death is lower for those in the oldest age group than for those in the middle age group, although the older persons have a stronger death tendency than the middle-aged persons.

The following sections focus on the effect of the covariates on the tendencies toward recovery or death as the reason for program terminations and the results of those tendencies on the probabilities of ultimate terminations due to recovery, death, or retirement.

Recovery and Death Tendencies

Because the mathematical models are not derived from a formal social or economic theory, causal interpretations of the covariate effects should be made with great caution. Still, the effects are of interest in that they represent basic patterns in the DI process regarding program terminations. The statistical analysis of the tendency for a beneficiary to recover or to die yields the following findings. (See the Technical Appendix for details of the models and coefficient estimates.)

Tendencies Over Time

The tendency for a recovery termination for all beneficiaries appears to increase for about the first 27 months after benefit entitlement. Then, a substantial drop in the recovery tendency occurs and it remains fairly constant and low until age 65. In conjunction with this phenomenon, there is an interesting shift in the overall pattern of association of the covariates with the tendency toward a recovery termination. Covariates have a significant effect on the recovery tendency more often after the 27-month drop than before the 27-month drop. The reason for these findings is not known with certainty; but, the data indicate two termination processes. The first process, during the first 27 months, is affected in part by SSA administrative rules whereby certain cases are reviewed after a specified period of time. The second process, after 27 months, is not affected by such reviews. In contrast, the tendency for a death termination did not indicate two separate processes. The covariates of age and sex had the expected association with the tendency to terminate from the program because of death.

Covariate Effects

Primary diagnosis.—As expected, the disabling condition has a profound and complex effect on the tendencies to recover or die over time. For example, the recovery tendency for males who had a white collar job, with 0-8 years of education, a PIA less than $299, aged 18-34 at entitlement, and were in the infective diagnostic group is higher during the first 27 months than for the same subgroup in the accidents group. After 27 months, the reverse is true. Also, during the first 27 months the covariates of sex, occupation, and education are not associated with the recovery tendency for the infective diagnostic group whereas they are for the accidents group. After 27 months, there is a different pattern of association of the

covariates with the recovery tendency for the infective group and the accidents group.

The effect of primary diagnosis on the recovery tendency is so complex that estimates for the recovery and death tendencies are made separately for each diagnostic group. The results of the separate estimates are used to compute recovery, death, and retirement probabilities.

Sex.—Men have a stronger tendency toward a recovery termination than women. They also have a stronger tendency than women for death to be the reason for program termination. These stronger tendencies for men tend to decrease their probability of retirement, compared with the retirement probability for women.

Age.—Younger workers have a stronger recovery tendency and a weaker death tendency. However, when the analysis factors in the longer time period from entitlement to age 65 for younger workers, the effect on the proportion of disabled workers who ultimately retire is not obvious because age also affects the length of time to retirement.

Education.—Educational level, in general, is important. Workers with only a grade school education have a weaker recovery tendency. No clear trend on the effect of years of education on the death tendency is noted.

The education variable serves two purposes. On the one hand, it stands as a proxy for human capital. As such, one would expect higher educational attainment to increase the recovery tendency. On the other hand, the SSA disability determination process for awarding benefits allows vocational factors to be used in determining whether or not to award benefits to persons who cannot be denied or awarded benefits strictly on medical factors. Because educational attainment is considered a vocational factor, beneficiaries with more education may have been more severely disabled than those with less education. In this sense, the education variable is a proxy for the level of severity of the disabling condition. As such, one would expect beneficiaries with more education to have a weaker recovery tendency because they would be more severely disabled. Because there was no control for the level of severity of the disabling condition in the present data, the data exhibit the net effect of the two conflicting associations of education with the recovery tendency.

Occupation.—In some diagnostic groups, occupation appears to be important but the data show no clear pattern common to all diagnoses. The general trend is that white-collar workers have a stronger recovery tendency than other occupational groups. These results indicate a number of possibilities—that the grouping is not appropriate, that the DOT codes do not constitute a meaningful covariate, or that occupation is not relevant to the process being examined.

Primary insurance amount.—The PIA appears to have a complex association with the tendency toward a recovery termination. For the most part, as the PIA level increases the recovery tendency decreases until the very high PIA level is reached. At that point, the recovery tendency begins to increase.

Because benefits are related to lifecycle earnings, this variable may represent potential earnings capability. Therefore, one would expect that persons with a higher PIA have a greater tendency to recover. But, this expectation is not substantiated by the data, except for the highest PIA group. It is necessary to remember that the data set does not contain a variable that measures the level of severity of the disabling condition. It may be that the level of severity increases across the first three PIA groups and thus the recovery tendency decreases across these groups.

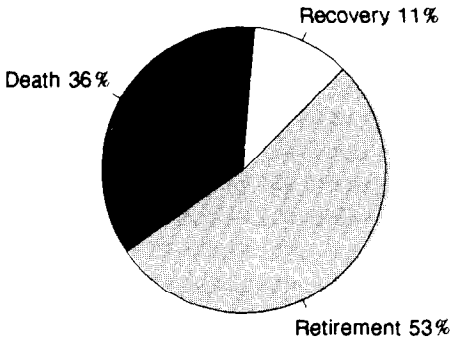
Projected Outcomes and Length of Time in the Program

The estimates of the recovery and death tendencies for disabled workers by various covariates are combined with retirement at age 65 to compute the probabilities of ultimately leaving the DI program rolls because of recovery, death, or retirement for various subgroups of beneficiaries. Table 1 presents the numbers of beneficiaries in different covariate combinations. These numbers are shown to indicate the prevalence of a given type of beneficiary in the program.

Chart 1 presents the projected outcomes of the first entitlement period for the entire 1972 cohort: About 53 percent remain in the program until retirement, about 11 percent recover, and 36 percent die. The fact that retirement is the major reason for program termination is not too surprising because most of the beneficiaries—65 percent—are in the oldest group aged 50-61 at entitlement.

Chart 2 displays the distribution of length of time from benefit entitlement to program termination for all types of program terminations and for the total population. The mean length of time in the DI program is 9.3 years. However, the mean number of years in the

Chart 1.—Projected outcomes of first entitlement period



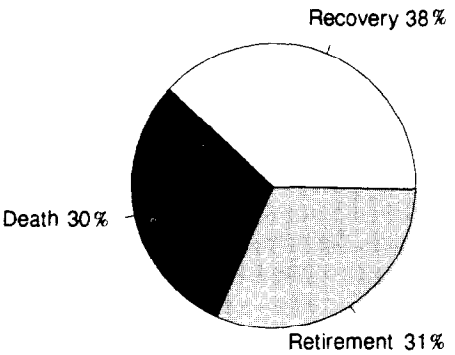
Reason for program termination

program cannot be viewed as the time when most beneficiaries leave the program rolls. The highest percentage of terminations—about 9.6 percent—occurs in the fourth year. The percentage of terminations in each of the first 2 years is about 8.4 percent. The proportion of terminations declines to 5.7 percent in the third year and then increases to 9.6 in the fourth year. This increase occurs because the fourth year is the first year in which an individual can retire (recalling that persons who were entitled to benefits after age 61 are excluded). Then, the yearly percentage of terminations declines over the remaining years. Half of the beneficiaries will have left the program rolls before the end of their seventh year after entitlement to benefits, although their mean length of time in the program is 9.3 years. Three-fourths of them will have experienced a termination by the end of their twelfth year after benefit entitlement.

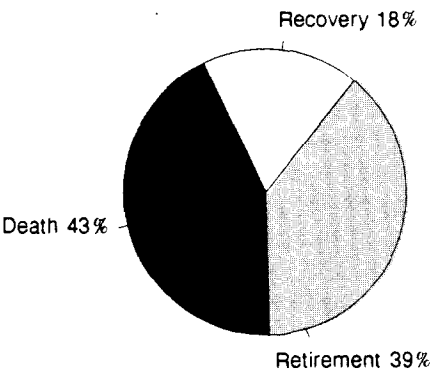
Age at Entitlement

Chart 3 shows how the projected reasons for program termination are associated with age at entitlement. In the youngest age group, which constitutes 11 percent of the beneficiary population, the estimated percentages of terminations of each type are about the same, with the largest percentage—38 percent—due to recovery. The percentage of expected recoveries is substantially lower for those in the middle age group—18 percent—with death and retirement percentages both larger than those in the younger group. In the oldest age group, containing 65 percent of the beneficiary population, most individuals leave the program rolls because of retirement (62 percent), 34 percent

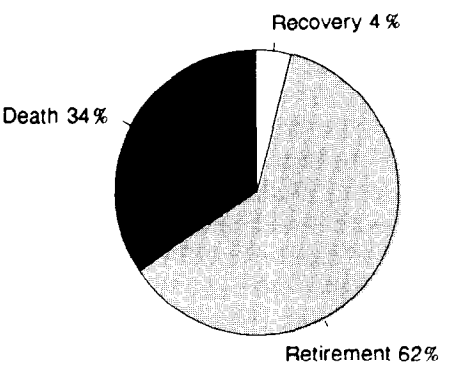
Chart 3.—Projected outcomes of first entitlement period, by age at entitlement



Age 18-34

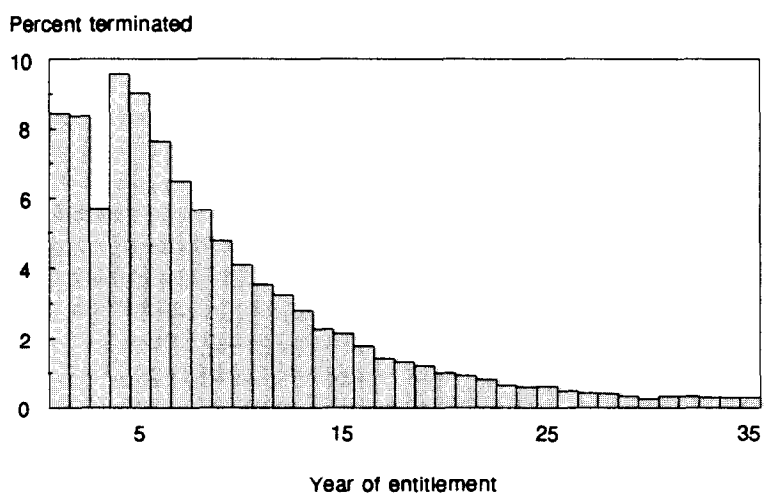


Age 35-49



Age 50-61

Chart 2.—Distribution of length of time from entitlement to termination



leave because of death, and 4 percent leave because of recovery. It is interesting to note that although the death tendency is stronger for older beneficiaries they have a lower probability of leaving the program rolls because of death. This difference occurs because the length of time from entitlement to age 65 is considerably shorter and the increased death tendency at any time during this period does not compensate for the shorter length of time in the program.

Age is also associated with the mean length of time in the program (chart 4). The most striking feature is the two-peaked structure of all three graphs. The first peak in each distribution results from only recovery and death as the reason for program terminations because it is not possible to reach retirement age at that time. In all age groups, after the first peak, the percentage of terminations declines over the next period of time until the first opportunity for retirement occurs. At this time, retirement terminations are possible, creating another peak in the distribution.

Chart 5 presents the distribution of the mean number of years from program entitlement to termination for the total population and for the three age groups. The mean length of time in the DI program for the total population is 9.3 years—18.4 years for the youngest beneficiaries; 12.5 years for the middle age group; and 6.5 years for the oldest group. Note that although the youngest beneficiaries have the highest probability of recovery, they nevertheless spend the longest time, on average, in the DI program.

Diagnostic Group

The primary diagnosis of the disabling condition has a profound effect on the program termination process. Chart 6 shows the projected termination outcomes of the first entitlement period by diagnostic group.

The largest percentage of projected recoveries—43 percent—occurs for individuals in the accidents group. This group is 7 percent of the population. Persons with diseases of the respiratory system have the lowest percentage of recoveries—1 percent. Six percent of the beneficiaries were diagnosed with diseases of the respiratory system. It is clear that the percentage of recoveries varies considerably across diagnostic groups. In the diagnostic group—diseases of the circulatory system—which represents 29 percent of the beneficiaries, only 5 percent recover. In the next largest group—diseases of the the musculoskeletal system—which represents comprises about 16 percent of the beneficiary population, about 14 percent recover.

The largest proportion of projected deaths occurs for persons

Chart 4.—Distribution of length of time from entitlement to termination, by age at entitlement

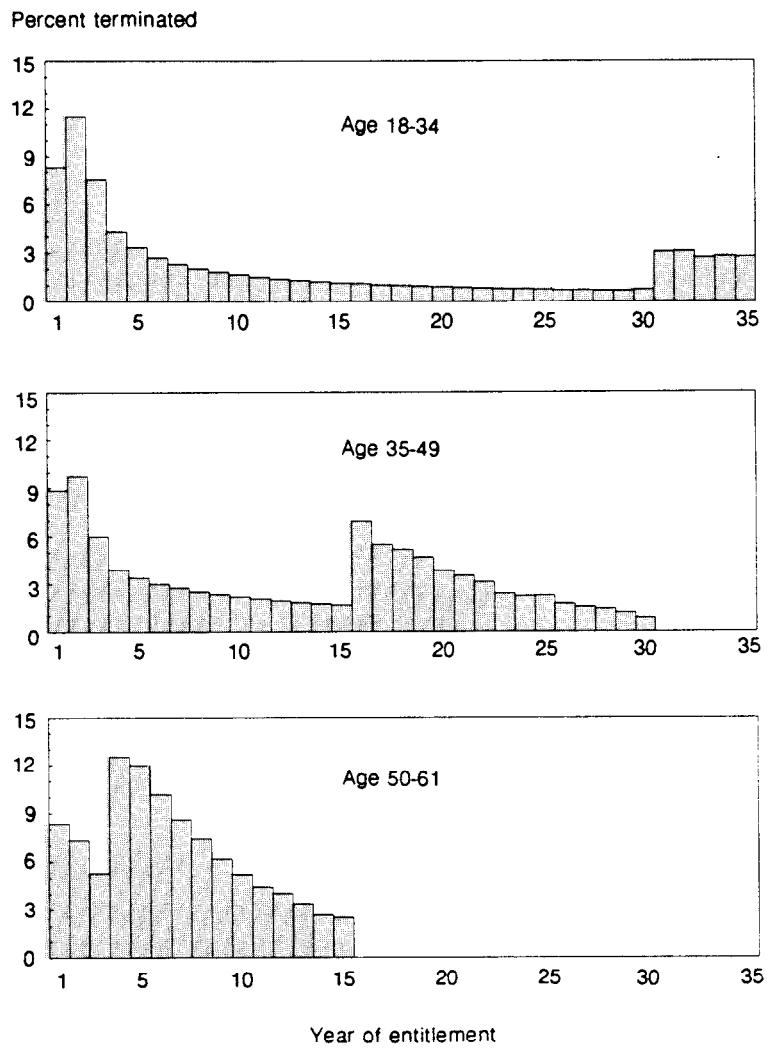
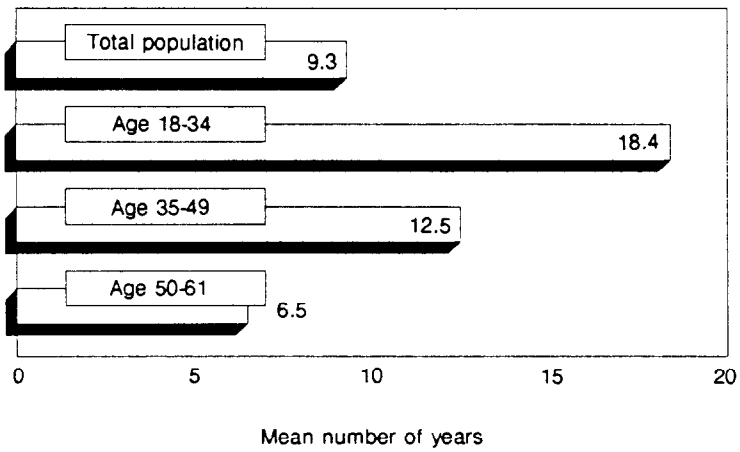


Chart 5.—Projected mean length of time of first entitlement period, by age at entitlement



in the neoplasms group—84 percent. Persons in the accidents group have the smallest proportion of deaths—14 percent—the same group that has the largest percentage of recoveries. In many diagnostic groups, the majority of beneficiaries would be transferred to the retirement program at age 65. However, the percentage retiring is less than 50 percent for five groups: Infective, neoplasms, digestive, genitourinary, and accidents.

The mean number of years in the program ranges from a low of 3.4 years for persons with neoplasms to a high of 15.6 years for persons with mental disorders (table 2, column 5). About 9 percent of the beneficiaries are in the mental disorders diagnostic group. The three groups with the largest mean number of years in the program are the mental disorders, congenital anomalies, and nervous system groups. Together they represent about 16 percent of the beneficiaries.

Age and Diagnostic Effect

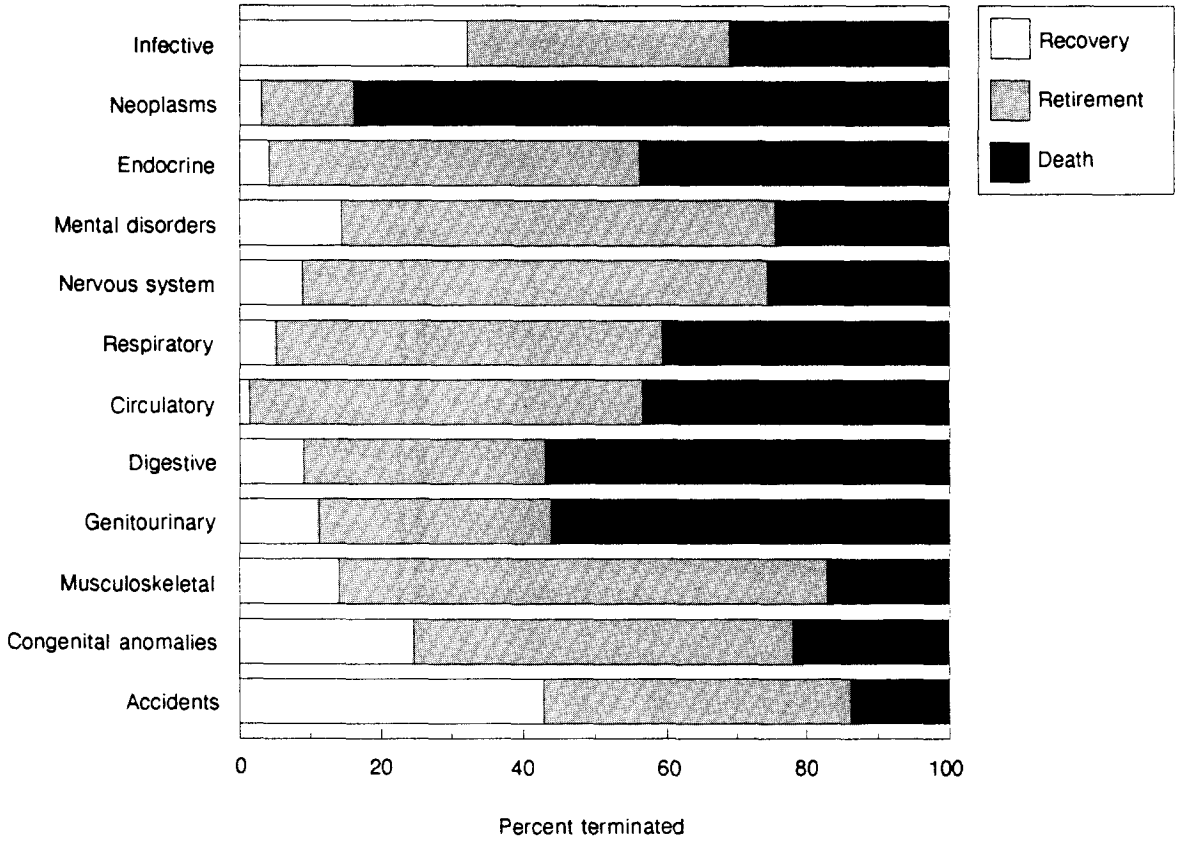
The differences found among diagnostic groups could be the result of different age distributions in the diagnostic groups. For this reason, the results are presented for age and diagnosis. The complex effect of age on the program termination process is shown in table 3. The recovery probability remains relatively high across all age groups for persons in the infective diseases and accidents groups. Beneficiaries with diseases of the respiratory system have a very low recovery probability in all age groups. The recovery probability declines very noticeably and the retirement probability increases for persons in the musculoskeletal group as the age at entitlement increases.

For the two youngest age groups, large differences were found in the mean number of years in the DI program by diagnostic group. In the oldest age group (with the exception of persons with neoplasms who have an especially low mean length of time in the program—3 years), the differences across diagnostic groups are

Table 2.—Projected outcomes of first entitlement period and average length of period, by primary diagnosis

Characteristic	Sample size	Percent of program terminations, by reason			Average years in program, by type of termination			
		Recovery	Death	Retirement	All types	Recovery	Death	Retirement
Total population	18,816	11	36	53	9.3	5.3	6.2	12.2
Primary diagnosis								
Infective	326	32	31	37	7.6	3.0	7.5	11.8
Neoplasms	1,525	3	84	13	3.4	3.1	2.4	10.4
Endocrine	620	4	44	52	8.3	6.4	6.8	9.7
Mental disorders	1,752	14	24	61	15.6	7.3	10.4	19.6
Nervous system	1,085	9	26	66	12.5	5.8	8.8	14.8
Circulatory	5,389	5	41	54	7.5	4.4	5.6	9.2
Respiratory	1,152	1	43	55	7.3	3.4	5.8	8.6
Digestive	545	9	57	34	7.0	4.5	5.6	10.0
Genitourinary	170	11	56	33	7.5	4.1	6.3	10.6
Musculosketal	2,917	14	17	69	10.0	5.2	8.0	11.5
Congenital anomalies	198	25	22	53	13.5	6.4	9.6	18.3
Accidents	1,276	43	14	43	9.9	4.8	9.0	15.2
Other	1,861	11	33	56	12.0	6.8	9.4	14.5

Chart 6.—Projected outcomes of first entitlement period, by diagnostic group



muted, in part because of the short period of time from entitlement to age 65. In all three age groups, persons with mental disorders—9 percent of the population—have the largest mean length of time in the program: 25.5 years for the youngest group, 16.4 years for the middle age group, and 7.8 for the oldest group. In the two youngest age groups, persons with diseases of the nervous system have the second largest mean number of years in the program: 23.4 years for those aged 18-34 and 16.2 years for those aged 35-49. Six percent of the population were diagnosed with diseases of the nervous system.

Sex

Chart 7 and table 4 show the projected outcomes for each of the covariates other than diagnosis. Considerably more women (62 percent) than men (49 percent) retire. The larger percentage of female retirees is due in part to the different age distributions of men and women: 68 percent of the women are in the oldest age group, compared with 64 percent of the men. Also, both the recovery and death rates are higher for men than for women. Women remain in the program somewhat longer than men, 9.9 years and 9.0 years, respectively.

Education

Although no differences are found in the percentage of terminations due to death, the percentage of recoveries increases as the educational level increases. However, the mean length of time in the program increases from 8.4 years for those whose highest level of education was eighth grade, to 9.6 years for those who completed high school. At the college level, the mean time in the DI program decreases slightly to 9.0 years.

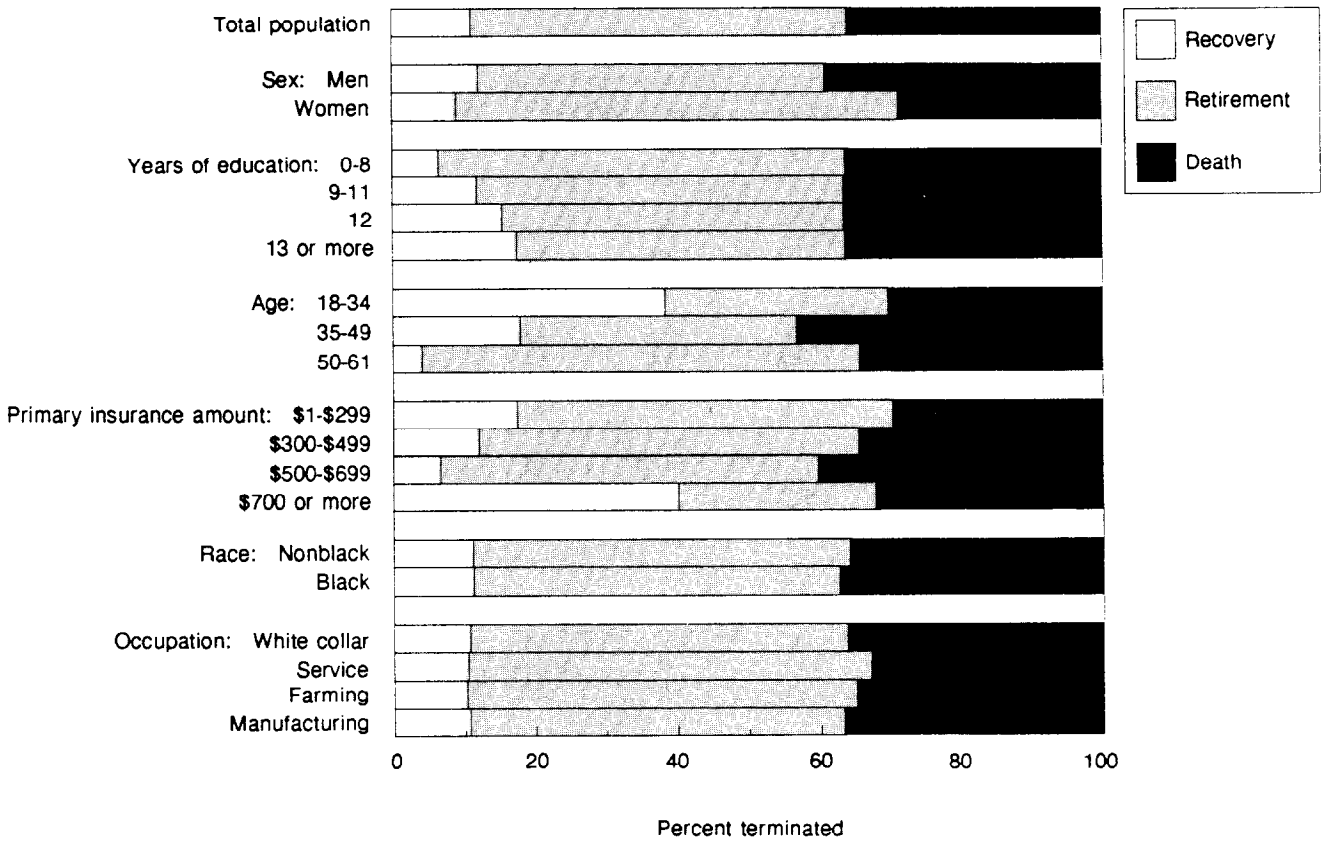
Primary Insurance Amount

For the first three PIA groups, the percentage of program terminations due to death rises from 30 percent to 40 percent. At the same time, the proportion of persons who recover and leave the program rolls decreases from 17 percent to 7 percent as the PIA increases. The percentage of retirement terminations remains constant for the first three PIA groups. For the last PIA group ($700 or more), the

Table 3.—Projected outcomes of first entitlement period and average length of period, by diagnostic group and age at entitlement

Characteristic	Sample size	Percent of program terminations, by reason			Average years in program, by type of termination			
		Recovery	Death	Retirement	All types	Recovery	Death	Retirement
	Aged 18-34 at entitlement							
Total	2,098	38	30	31	18.4	6.8	13.5	37.4
Infective	39	66	29	5	7.4	3.1	11.7	38.3
Neoplasms	105	6	89	5	5.1	3.2	3.5	36.9
Endocrine	53	6	90	4	11.1	10.7	9.9	35.8
Mental disorders	503	30	19	51	25.5	8.7	18.7	37.8
Nervous system	197	25	30	44	23.4	7.2	16.4	37.5
Circulatory	95	36	39	24	16.7	7.5	13.7	35.3
Respiratory	15	2	86	13	15.9	5.0	13.4	34.5
Digestive	34	27	68	5	9.0	3.7	9.2	36.8
Genitourinary	28	12	84	4	9.5	5.0	8.7	36.5
Musculoskeletal	260	59	15	26	15.4	6.3	16.6	35.7
Congenital anomalies	49	30	32	38	21.8	8.8	14.0	38.7
Accidents	411	73	10	16	11.8	5.6	15.0	37.8
Other	309	17	43	40	24.2	11.6	17.0	37.6
	Aged 35-49 at entitlement							
Total	4,446	18	43	39	12.5	5.0	8.1	20.9
Infective	127	44	37	20	8.8	3.3	8.8	21.1
Neoplasms	373	3	89	8	4.3	3.9	2.9	20.3
Endocrine	132	6	65	29	11.7	8.5	8.4	19.8
Mental disorders	548	15	27	57	16.4	5.6	11.2	21.7
Nervous system	272	11	30	59	16.2	5.0	10.2	21.3
Circulatory	984	13	56	32	11.6	5.1	8.4	19.6
Respiratory	160	2	63	35	12.4	4.3	9.1	18.9
Digestive	170	11	72	16	8.9	5.2	7.0	20.0
Genitourinary	56	13	67	20	9.6	4.8	7.1	20.7
Musculoskeletal	691	28	20	52	14.7	4.9	11.4	21.2
Congenital anomalies	70	27	24	49	15.0	6.6	9.0	22.7
Accidents	362	51	14	35	11.2	4.3	10.2	21.4
Other	501	13	43	44	14.5	7.0	10.0	21.0
	Aged 50-61 at entitlement							
Total	12,272	4	34	62	6.5	3.3	4.2	8.1
Infective	160	15	27	58	6.7	2.2	5.0	8.7
Neoplasms	1,047	3	82	16	3.0	2.7	2.1	7.8
Endocrine	435	3	32	65	6.9	4.2	4.7	8.2
Mental disorders	701	3	26	71	7.8	4.3	5.4	8.8
Nervous system	616	2	22	75	7.4	2.8	4.7	8.3
Circulatory	4,310	3	37	60	6.4	2.8	4.5	7.7
Respiratory	977	1	39	59	6.3	3.1	4.7	7.5
Digestive	341	6	48	46	5.8	4.2	4.0	7.9
Genitourinary	86	10	40	50	5.4	3.2	3.8	7.1
Musculoskeletal	1,966	3	16	80	7.6	3.3	5.5	8.2
Congenital anomalies	79	19	13	67	6.9	3.9	4.0	8.4
Accidents	503	12	17	71	7.4	2.8	5.2	8.7
Other	1,051	8	25	67	7.2	3.7	5.0	8.4

Chart 7.—Projected outcomes of first entitlement period, by covariates



percentage of recoveries dramatically increases to 40 percent while the percentage of deaths and retirements decreases. This may be due to the large percentage of young beneficiaries in this PIA group (table 1). The mean number of years in the program is largest for persons with a PIA of $700 or more—11.1 years—although the percentage of recoveries is highest for this group. The presence of a relatively large number of young beneficiaries in this group may be the reason behind this difference.

Race and Occupation

Very little difference in the projected termination outcomes was found by race. The percentages of recoveries are very close: 53 percent nonblack and 51 percent black. The percentage of deaths is slightly higher for black beneficiaries than for nonblack beneficiaries. Only small differences are found in the projected program termination outcomes among occupational groups.

Time to Recovery

Because the recovery process is of special interest, the next sections focus on the length of time in the program specifically for those with a recovery termination. Approximately 11 percent of the projected program terminations are due to recovery. Chart 8 shows that the largest percentage of recoveries occurs in the second year of entitlement. The mean length of time in the program for those who recover is 5.3 years. The 50th percentile is reached in the third year. Three-fourths of the recoveries occur by the sixth year.

Age and Diagnosis

Table 1 shows the number of beneficiaries in the following groups: Diagnoses, education, PIA, occupation, race, age, and sex. Table 3, column 6 presents the mean length of time from entitlement to recovery by age and diagnosis. In most cases, the mean length of time decreases with age. This decrease is, in part, a result of the retirement process. A person who becomes entitled to disabled-worker benefits at an older age has a shorter length of time to retirement age. Therefore, if a recovery occurs, it cannot be at

Table 4.—Projected outcomes of first entitlement period and average length of period, by selected characteristics

Characteristic	Sample size	Percent of program terminations, by reason			Average years in program, by type of termination			
		Recovery	Death	Retirement	All types	Recovery	Death	Retirement
Total population	18,816	11	36	53	9.3	5.3	6.2	12.2
Sex								
Men	13,139	12	39	49	9.0	5.1	6.2	12.2
Women	5,677	9	29	62	9.9	5.7	6.0	12.3
Years of education								
0-8	6,798	6	36	57	8.4	4.6	5.5	10.6
9-11	3,949	12	37	52	9.2	4.9	6.0	12.3
12	4,330	15	37	48	9.6	5.3	6.2	13.7
13 or more	1,495	18	36	46	9.0	5.4	5.8	12.9
Age at entitlement								
18-34	2,098	38	30	31	18.4	6.8	13.5	37.4
35-49	4,446	18	43	39	12.5	5.0	8.1	20.9
50-61	12,272	4	34	62	6.5	3.3	4.1	8.1
Primary insurance amount								
$1-$299	3,041	17	30	53	9.7	5.4	6.8	12.8
$300-$499	8,117	12	34	53	9.8	5.5	6.5	12.9
$500-$699	7,416	7	40	53	8.5	5.0	5.6	11.1
$700 or more	242	40	32	28	11.1	4.8	8.5	23.1
Race								
Nonblack	16,159	11	36	53	9.1	5.3	6.1	12.0
Black	2,657	11	37	51	10.1	5.3	6.7	13.5
Occupation								
White collar	4,092	11	36	53	8.7	5.2	5.5	11.7
Service	2,690	10	33	57	9.2	5.6	5.9	11.8
Farming	763	10	35	55	8.2	4.5	5.6	10.5
Manufacturing	5,613	11	37	53	9.0	5.0	6.0	11.9

some of the later times from entitlement available to the younger beneficiaries. Across all age groups, persons with infective and parasitic diseases have the lowest mean length of time from entitlement to recovery—approximately 2-3 years. In the first two age groups, persons with endocrine, nutritional, and metabolic diseases have the highest mean length of time to recovery—10.7 and 8.5 years, respectively. In the oldest age group, persons with endocrine, nutritional and metabolic diseases; mental disorders; and diseases of the digestive system have the highest mean length of time to recovery—approximately 4 years.

Age and Sex

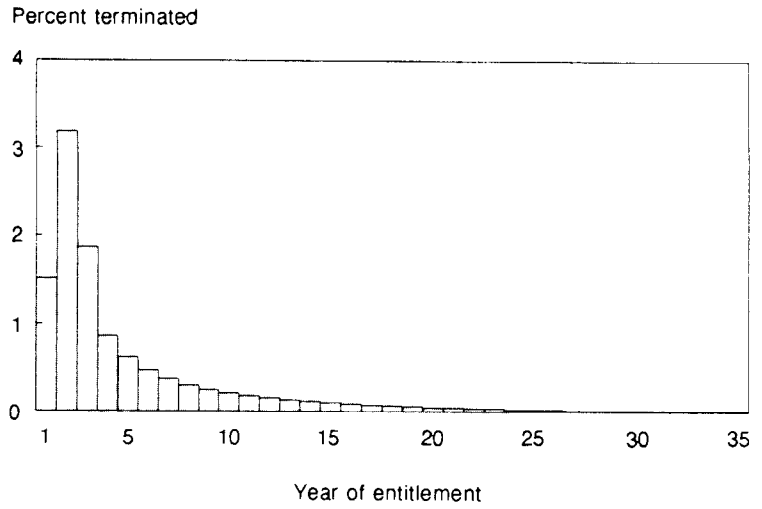
Table 5 presents the mean number of years in the DI program by age and sex for those who recover. In all age groups, the mean number of years in the program is lower for men than for women. The difference becomes smaller with age, as the time from entitlement to age 65 becomes shorter.

Other Covariates

Table 4 presents the mean number of years in the program from entitlement to recovery by other covariates. Across educational groups, a slight upward trend in the mean number of years to recovery is seen, although the probability of recovery increases.

Across PIA groups, the mean number of years ranges from 4.8 years to 5.5 years with no clear trend. The mean number of years in the program by race is identical—5.3 years—and across occupational groups it is 4.6 years to 5.5 years.

Chart 8.—Distribution of length of time from entitlement to recovery



Further Results

Tables 2-9 present more detailed information than the basic findings presented thus far. Tables 2-5 show the estimated percentage of ultimate recoveries, deaths, and retirements by selected characteristics. Average length of time in the DI program is shown for all terminations and by reason for program termination. These estimates are shown for each diagnostic group in table 2, and by the other covariates in table 4. Tables 3 and 5 show the same information for each age-diagnostic combination and for each age-sex combination, respectively. Tables 6-9 use the same covariate scheme to present the 25th, 50th, and 75th percentiles of the distribution of the length of time in the program for all terminations and separately for recoveries. The quartiles are presented to provide a sense of the shape of the distributions that cannot be obtained from the means alone.

Study Limitations

The sampling error for these projections was not estimated because of the large computational burden imposed by the calculations. The recovery and death termination models use about the first 9 years after of entitlement and have assumed that the same functional forms are appropriate for the later years. This assumption seems reasonable for those whose terminations were based on recovery because most recoveries appear to occur relatively soon after entitlement. The impact of changing death rates in the later years is lessened because a beneficiary is automatically converted from the DI program to the retirement program at age 65. It is believed these projections are reasonable approximations to the population figures.

Summary

From a 5-percent random sample of beneficiaries first entitled to disabled-worker benefits in 1972, with an event observation period up to 9 years, the postentitlement process was modeled and projections were made to the end of the first entitlement period. Basic information about the proportion of beneficiaries who ultimately recover, die, or retire, as well as the mean length of time to program termination, is presented as a demographic summary of the pre-1980's Disability Insurance program.

It is projected that 11 percent of the beneficiaries would ultimately recover, 36 percent would have benefits terminated at death, and 53 percent would remain in the program until retirement. The mean length of time in the program is estimated to be 9.3 years. The projected outcomes and mean number of years in the program vary considerably by diagnosis, education, PIA, age, and sex. These variations should provide useful baseline information to those interested in the effect of administrative changes in the disability determination process or in medical advances that would change the distribution of the covariates that are discussed in this article.

Table 5.—Projected outcomes of first entitlement period and average length of period, by age and sex

Characteristic	Sample size	Percent of program terminations, by reason			Average years in program, by type of termination			
		Recovery	Death	Retirement	All types	Recovery	Death	Retirement
Total population	18,816	11	36	53	9.3	5.3	6.2	12.2
Age 18-34	2,098	38	30	31	18.4	6.8	13.5	37.4
Men	1,596	41	31	28	17.4	6.5	13.7	37.4
Women	502	29	29	41	21.6	8.4	12.8	37.4
Age 35-49	4,446	18	43	39	12.5	5.0	8.1	20.9
Men	3,146	19	47	35	12.0	4.8	8.2	21.0
Women	1,300	16	35	49	13.8	5.5	7.8	20.7
Age 50-61	12,272	4	34	62	6.5	3.3	4.2	8.1
Men	8,397	4	38	58	6.3	3.1	4.2	7.9
Women	3,875	4	27	70	7.1	3.4	4.1	8.4

Table 6.—Year in which 1st, 2nd, and 3rd quartiles of program termination distributions occur, for all terminations and for recoveries only, by diagnostic group

Characteristic	Sample size	All type of terminations			Recovery terminations only		
		Year terminations occurred for—			Year terminations occurred for first—		
		25 percent	50 percent	75 percent	25 percent	50 percent [1]	75 percent
Total population	18,816	4	7	12	2	3	6
Diagnostic group							
Infective	326	2	6	11	2	2	3
Neoplasms	1,525	1	2	4	2	2	3
Endocrine	620	4	7	11	2	4	9
Mental disorders	1,752	6	12	23	2	4	9
Nervous system	1,085	5	9	17	2	3	
Circulatory	5,389	4	6	10	2	3	5
Respiratory	1,152	4	6	9	2	2	4
Diagestive	545	3	5	9	2	3	6
Genitourinary	170	3	5	10	2	3	5
Musculoskeletal	2,917	5	8	13	2	3	6
Congenital anomalies	198	5	9	20	2	4	8
Accidents	1,276	3	6	14	2	3	5
Other	1,861	5	9	16	2	4	9

[1] Median years.

Table 7.—Year in which 1st, 2nd, and 3rd quartiles of program termination distributions occur, for all terminations and for recoveries only, by selected characteristics

Characteristic	Sample size	All type of terminations			Recovery terminations only		
		Year terminations occurred for—			Year terminations occurred for first—		
		25 percent	50 percent	75 percent	25 percent	50 percent [1]	75 percent
Total population	18,816	4	7	12	2	3	6
Sex							
Men	13,139	4	6	12	2	3	6
Women	5,677	4	8	13	2	3	7
Years of education							
0-8	6,798	4	6	11	2	3	5
9-11	3,949	4	7	12	2	3	6
12	4,330	3	7	13	2	3	6
13 or more	1,495	3	6	12	2	3	6
Age at entitlement							
18-34	2,098	3	14	34	2	3	8
35-49	4,446	4	13	19	2	3	6
50-61	12,272	4	6	9	2	2	4
Primary insurance amount							
$1-$299	3,041	4	7	13	2	3	6
$300-$499	8,117	4	7	13	2	3	6
$500-$699	7,416	4	6	11	2	3	6
$700 or more	242	2	6	15	2	2	4
Race							
Nonblack	16,159	4	7	12	2	3	6
Black	2,657	4	7	13	2	3	6
Occupation							
White collar	4,092	4	6	11	2	3	6
Service	2,690	4	7	12	2	3	7
Farming	763	4	6	10	2	3	5
Manufacturing	5,613	4	7	12	2	3	6

[1] Median years.

Table 8.—Year in which 1st, 2nd, and 3rd quartiles of program termination distributions occur, for all terminations and for recoveries only, by age and diagnostic group

Diagnostic group	Sample size	All type of terminations: Year terminations occurred for— 25 percent	50 percent	75 percent	Recovery terminations only: Year terminations occurred for first— 25 percent	50 percent[1]	75 percent
	Aged 18-34 at entitlement						
Total	2,098	3	14	34	2	3	8
Infective	39	2	3	8	1	2	3
Neoplasms	105	1	2	3	2	3	3
Endocrine	53	3	8	16	2	8	17
Mental disorders	503	9	32	38	3	5	12
Nervous system	197	7	29	37	2	4	9
Circulatory	95	4	12	31	2	5	10
Respiratory	15	6	13	25	2	3	6
Digestive	34	2	4	12	2	2	3
Genitourinary	28	2	6	13	2	3	6
Musculoskeletal	260	3	8	31	2	3	7
Congenital anomalies	49	5	21	38	3	5	11
Accidents	411	2	4	18	2	3	6
Other	309	11	28	36	3	8	18
	Aged 35-49 at entitlement						
Total	4,446	4	13	19	2	3	6
Infective	127	2	5	16	2	2	3
Neoplasms	373	1	2	4	2	2	4
Endrocrine	132	5	11	17	2	7	13
Mental disorders	548	10	18	23	2	3	7
Nervous system	272	10	18	22	2	3	6
Circulatory	984	4	11	18	2	3	7
Respiratory	160	6	13	17	2	3	5
Digestive	170	3	7	15	2	3	7
Genitourinary	56	3	7	17	2	3	6
Musculoskeletal	691	6	17	21	2	3	6
Congenital anomalies	70	5	17	23	3	4	9
Accidents	362	2	8	19	2	2	5
Other	501	7	16	20	2	5	11
	Aged 50-61 at entitlement						
Total	12,272	4	6	9	2	2	4
Infective	160	4	6	10	1	2	2
Neoplasms	1,047	1	2	4	2	3	3
Endocrine	435	4	6	9	2	3	6
Mental disorders	701	5	7	11	2	4	6
Nervous system	616	5	7	10	2	2	3
Circulatory	4,310	4	6	9	2	2	3
Respiratory	977	4	6	8	2	2	3
Digestive	341	3	5	8	2	3	5
Genitourinary	86	3	5	7	2	3	4
Musculoskeletal	1,966	5	7	10	2	3	4
Congenital anomalies	79	4	6	10	2	3	5
Accidents	503	4	7	10	2	2	3
Other	1,051	4	7	10	2	3	5

[1] Median years.

Table 9.—Year in which 1st, 2nd, and 3rd quartiles of the termination distributions occur, for all terminations and for recoveries only, by age and sex

Characteristics	Sample size	All type of terminations			Recovery terminations only		
		Year terminations occurred for—			Year terminations occurred for first—		
		25 percent	50 percent	75 percent	25 percent	50 percent[1]	75 percent
Total population	18,816	4	7	12	2	3	6
Age 18-34	2,098	3	14	34	2	3	8
Men	1,596	3	11	33	2	3	8
Women	502	5	23	36	2	4	11
Age 35-49	4,446	4	13	19	2	3	6
Men	3,146	3	12	19	2	3	6
Women	1,300	5	16	20	2	3	7
Age 50-61	12,272	4	6	9	2	2	4
Men	8,397	4	6	9	2	2	4
Women	3,875	4	7	10	2	3	4

[1] Median years.

References

[1] Bye, Barry V. and Gerald F. Riley, "Eliminating the Medicare Waiting Period for Social Security Disabled-Worker Beneficiaries," **Social Security Bulletin**, May 1989, pages 2-15.

[2] Bye, Barry; Gerald Riley; and James Lubitz, "Medicare Utilization by Disabled-Worker Beneficiaries: A Longitudinal Analysis," **Social Security Bulletin**, December 1987, pages 13-28.

[3] Dykacz, Janice M. and John C. Hennessey, "Analysis of the Postrecovery Experience of Disabled-Worker Beneficiaries," **Social Security Bulletin**, September 1989, pages 44-66.

[4] World Health Organization, **Manual of the International Classification of Diseases, Injuries, and Causes of Death**, Eighth Revision, Geneva, 1967.

[5] Department of Labor, **Dictionary of Occupational Titles**, Washington, DC, 1965.

[6] Elandt-Johnson, Regina C. and Norman L. Johnson, **Survival Models and Data Analysis**, John Wiley & Sons, New York, 1980, page 185.

[7] Hennessey, John C., "Testing the Predictive Power of a Proportional Hazard Semi-Markov Population Model," **Mathematical Biosciences**, 1983, pages 193-212.

[8] Lawless, J.F., **Statistical Models and Methods for Lifetime Data**, John Wiley & Sons, New York, 1982.

[9] Social Security Administration, **Social Security Administration Disability Insurance State Manual (DISM)**, 1977, sections 347-349.

Technical Appendix

Data File

The data file used for this study is an extract from a larger file compiled from three data sources—two extracts from the Master Beneficiary Record (MBR) December 31, 1985, and June 1986 and the Duration & Termination Study (DTS) file. Sample selection was based on the two terminal digits of the Social Security Number from the MBR and resulted in a 5-percent random sample of 23,062 disabled-worker beneficiaries first entitled in 1972. Some 1,006 records were deleted from this larger file because they did not match the DTS records, which is the source of the data for the covariates. An additional 29 cases were rejected because of the absence of a sex/race code on both the DTS and the MBR extract. The remaining 22,027 cases constitute the data file.

A smaller number of cases is used in the analysis. Records of 2,824 persons whose age at entitlement is 62 or older are omitted because information distinguishing retired workers from disabled workers is not available in this data file. An additional 387 cases are omitted due to death in the month of entitlement or inconsistent information on key variables. The remaining 18,816 cases are used in the analysis.

Hazard Functions

The starting point in modeling the first entitlement period is the estimation of hazard functions for both terminations based on recovery and terminations due to death. A hazard function measures the instantaneous tendency of the given event (recovery or death) to occur at each point in time before retirement. Using formulas discussed below, the recovery and death hazard functions are combined with retirement age in a competing risk model to obtain the probabilities of termination of the first period of entitlement [6]. These probabilities are used to estimate the mean number of years that the individual was on the Disability Insurance (DI) program rolls.

To understand the hazard function concept, the recovery termination hazard function is presented as an example. To estimate $h_r(t)$, the value of the recovery hazard function at time t, a small interval of time, $(t, t+\Delta t)$ is considered. The ratio of the number of beneficiaries who recover during that time interval divided by the number who are still observed to be entitled to benefits at the beginning of the time interval is computed. This ratio provides an estimate of $P_r(t, t+\Delta t)$, the probability that a beneficiary will recover before time $t+\Delta t$ given that the individual is still in the program at time t. The ratio is then divided by the length of the time interval, Δt, to obtain an approximation of the hazard function value at time t. Because the size of the time interval is allowed to decrease toward zero $h_r(t)$ was obtained—the exact value of the hazard function at time t. The number, $h_r(t)$, measures the instantaneous tendency of a beneficiary to recover at time t. More precisely, the recovery hazard function is defined as

$$h_r(t) = \lim_{\Delta t \to 0} \frac{P_r(t, t+\Delta t)}{\Delta t} \qquad (1)$$

The death termination hazard function, $h_d(t)$, is defined the same way as the recovery hazard function except that death becomes the event of interest. Each hazard function is modeled separately, under the assumption that no common parameters exist. They are then combined in a competing risk model to compute the probabilities of recovering or dying in each year before retirement. Hereafter the recovery termination hazard function and the death termination hazard function will be referred to as the recovery hazard function and the death hazard function, respectively.

Competing Risk Model

The recovery and death hazards were estimated as functions of the covariates to obtain $h_r(t,\vec{x})$ and $h_d(t,\vec{x})$ where $\vec{x} = (x_1,\ldots,x_n)$ is the vector of covariates. From these estimated hazard functions, all quantities of interest can be computed using the following rules. (The formulas are derived from the ones appearing in [7].)

Outcome probabilities.—If the beneficiary attains age 65 after year t, then the probability, $P_r(t,\vec{x})$, of a recovery occuring in year t of entitlement is obtained from the formula

$$P_r(t,\vec{x}) = \int_{t-1}^{t} \left(h_r(u,\vec{x})\, e^{-\int_0^u [h_r(s,\vec{x})+h_d(s,\vec{x})]ds} \right) du. \qquad (2)$$

The corresponding probability of death in year t is

$$P_d(t,\vec{x}) = \int_{t-1}^{t} \left(h_d(u,\vec{x})\ e^{-\int_0^u [h_r(s,\vec{x})+h_d(s,\vec{x})]ds} \right) du. \quad (3)$$

The probability of retirement in year t is

$$P_{65}(t,\vec{x}) = 0. \quad (4)$$

If the beneficiary attains age 65 in year t, and t^* is the exact time of the birthday,

$$P_r(t,\vec{x}) = \int_{t-1}^{t^*} \left(h_r(u,\vec{x})\ e^{-\int_0^u [h_r(s,\vec{x})+h_d(s,\vec{x})]ds} \right) du, \quad (5)$$

$$P_d(t,\vec{x}) = \int_{t-1}^{t^*} \left(h_d(u,\vec{x})\ e^{-\int_0^u [h_r(s,\vec{x})+h_d(s,\vec{x})]ds} \right) du, \quad (6)$$

and

$$P_{65}(t,\vec{x}) = 1 - \sum_{k=1}^{t} \left[P_r(k,\vec{x}) + P_d(k,\vec{x}) \right]. \quad (7)$$

If the beneficiary attains age 65 before year t,

$$P_r(t,\vec{x}) = 0;\ P_d(t,\vec{x}) = 0;\ P_{65}(t,\vec{x}) = 0. \quad (8)$$

Because the integrals cannot be computed in closed form, Simpson's Rule was used.

Number of years in DI program.—The mean number of years in the program, $M(\vec{x})$, for a given covariate vector, $\vec{x}$, is given by

$$M(\vec{x}) = \int_0^{\infty} t\left[P_r(t,\vec{x}) + P_d(t,\vec{x}) + P_{65}(t,\vec{x}) \right] dt. \quad (9)$$

However, for computational convenience the mean number of years in the program is approximated by the discrete formula

$$\mathbb{M}(\vec{x}) = \sum_{t=1}^{50} t\left[P_r(t,\vec{x}) + P_d(t,\vec{x}) + P_{65}(t,\vec{x}) \right]. \quad (10)$$

To compute the mean length of time to recovery, $M_r(\vec{x})$, formula (10) is modified in the usual way:

$$\mathbb{M}_r(\vec{x}) = \frac{\sum_{t=1}^{50} t\left[P_r(t,\vec{x}) \right]}{\sum_{t=1}^{50} \left[P_r(t,\vec{x}) \right]} \quad (11)$$

The mean number of years to death or retirement is computed with similar formulas. Tables for subgroups of the population were produced by averaging the estimated probabilties over all sample cases in the subgroup.

Model Estimation

To construct the model, parametric forms were chosen for the recovery and death hazard functions. The details of these forms and the introduction of covariates will be discussed below. The recovery and death hazard functions are estimated by the method of maximum likelihood as described in [7]. The likelihood is the product of the probability densities of the duration of time in the program for each individual who recovers or dies. For those who remain in the program until they reach age 65 or who were in the program on January 1, 1981, the end of the observation period, the probability of remaining in the program for the given number of months is included in the likelihood function. Thus, the likelihood function is given by

$$L = \prod_{\left(\begin{smallmatrix}\text{all}\\ \text{recoveries}\end{smallmatrix}\right)} f_r(t,\vec{x}) \times \prod_{\left(\begin{smallmatrix}\text{all}\\ \text{deaths}\end{smallmatrix}\right)} f_d(t,\vec{x}) \times \prod_{\left(\begin{smallmatrix}\text{all retirements}\\ \text{and end of the}\\ \text{observation period}\end{smallmatrix}\right)} \left[1 - F_r(t,\vec{x}) - F_d(t,\vec{x}) \right], \quad (12)$$

where

t = time of the event (recovery, death, or retirement) or the time that the beneficiary reaches the end of the observation period, January 1, 1981,

$\vec{x} = (x_1,\ldots,x_n)$ is the vector of covariates,

$f_r(t,\vec{x})$ = the recovery density function,

$f_d(t,\vec{x})$ = the death density function,

$F_r(t,\vec{x})$ = the cumulative recovery termination distribution, and

$F_d(t,\vec{x})$ = the cumulative death termination distribution.

Both the densities and cumulative distributions are obtained from the hazard functions. The cumulative recovery distribution is given by

$$F_r(t,\vec{x}) = \int_0^t \left\{ h_r(u,\vec{x})\, e^{-\int_0^u [h_r(s,\vec{x})+h_d(s,\vec{x})]ds} \right\} du. \tag{13}$$

The cumulative distribution of terminations because of death is similar. The usual derivative computes the density function:

$$f_r(t,\vec{x}) = \frac{dF_r(t,\vec{x})}{dt} = h_r(t,\vec{x})\, e^{-\int_0^t [h_r(s,\vec{x})+h_d(s,\vec{x})]ds} \tag{14}$$

and

$$f_d(t,\vec{x}) = \frac{dF_d(t,\vec{x})}{dt} = h_d(t,\vec{x})\, e^{-\int_0^t [h_r(s,\vec{x})+h_d(s,\vec{x})]ds} \tag{15}$$

Taking the logarithm of the likelihood, (12), replacing the density and cumulative distribution functions with their formulas in terms of the hazard functions (13), (14) and (15) and rearranging by collecting terms involving the same hazard function, results in

$$\ln(L) = \left\{ \sum_{t,\vec{x}} \left[N_r(t,\vec{x})\ln(h_r(t,\vec{x})) - N(t,\vec{x})\int_0^t h_r(s,\vec{x})ds \right] \right\} + \left\{ \sum_{t,\vec{x}} \left[N_d(t,\vec{x})\ln(h_d(t,\vec{x})) - N(t,\vec{x})\int_0^t h_d(s,\vec{x})ds \right] \right\} \tag{16}$$

where

$N_r(t,\vec{x})$ = the number of recoveries at time t for those beneficiaries with covariate $\vec{x}$,

$N_d(t,\vec{x})$ = the number of deaths at time t for those beneficiaries with covariate $\vec{x}$, and

$N(t,\vec{x})$ = the number of events of any kind at time t for those beneficiaries with covariate $\vec{x}$.

Each hazard function appears in only one of the braces in (16), but each brace contains all of the data. Therefore, if no common parameters are shared by the two hazard functions, each one can be estimated separately by maximizing the corresponding brace from (16).

Weibull Hazard Function

As a first step, a standard mathematical form used in survival analysis—the Weibull hazard function—is estimated. With no covariates, its equation is

$$h(t) = \frac{\eta\, t^{\eta-1}}{\beta^{\eta}}, \qquad \eta > 0,\ \beta > 0. \tag{17}$$

The Weibull form provides a wide variety of unimodal hazard functions and has been found appropriate in a number of studies mentioned in [8]. The two constants in this equation, η and β, are called the shape and scale parameters, respectively. If $\eta = 1$, the hazard

function is constant—that is, the tendency for the given event to occur is the same over time. If $\eta > 1$, the hazard function is increasing, and, if $\eta < 1$, the hazard function is decreasing. The constant β affects the scale. Because β appears in the denominator, a larger β will result in lower hazard values and a smaller β will result in higher hazard values.

Introduction of Covariates

Covariates are normally introduced into the Weibull model as follows:

$$h_r(t,x_1,\ldots,x_n) = \frac{e^{\gamma}\, t^{e^{\gamma}-1}}{\left[e^{\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n}\right]^{e^{\gamma}}}, \quad (18)$$

where each x_i is a 0-1 dummy variable representing a particular covariate group. This approach has been used successfully in several other studies mentioned in [8]. Also, it produces a proportional hazards model that was tested in related work in [7].

In this equation, the expression e^{γ} has been used instead of η to ensure that this constant is estimated as a positive number. The expression

$$e^{\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n} \quad (19)$$

replaces the constant β in equation (17). In this way, the shape of the hazard function described by e^{γ} is the same for all covariate values. The scale described by β is adjusted to a different value for each set of covariate values. The parameters β_0 through β_n describe how the covariates affect the tendency for the event to occur. Each x_i represents an element of a particular covariate group. Because (19) appears in the denominator of (18), a positive β_i indicates that the group represented by $x_i = 1$ will have a lower value for the hazard function—that is, a weaker tendency for the event to occur than the reference group for the given covariate. A negative β_i implies that the group will have a stronger tendency than the reference group.

Certain covariates used in the model, such as age at entitlement, educational level, and PIA were all discretized into groups because there was no reason to assume that the relationships were linear and monotone. The age classification groups of 18-34, 35-49, and 50-61 seem like a natural grouping of beneficiaries into young, middle aged, and near to retirement. Educational groups were formed according to highest level attained: grade school, some high school, high school graduate, and some education beyond high school. The PIA groups were formed as \$1-\$299, \$300-\$499, \$500-\$699, and \$700 or more. The remaining variables (sex, race, and occupation) were already nominal in nature.

Goodness-of-Fit

The principal use of the model is to project program terminations due to recovery, death, and retirement beyond the observation period. Therefore, the assessment of goodness-of-fit centers on whether or not the model of the hazard function is able to reproduce the occurrence of events especially near the end of the observation period. A visual comparison is used between the Kaplan-Meier (KM) estimate and the model estimate of the survival curve. The value of the survival curve, S(t), is the proportion of the population that would still be in the program after t months of entitlement if the only type of termination possible is the type being estimated. Although this function is not of direct interest in the analysis, it nonetheless isolates the given termination process from the others so that the goodness-of-fit can be studied for each hazard function separately.

The Kaplan-Meier estimate of the survival curve is a nonparametric step function approximation to the survival curve. As such, it makes no assumptions about a functional form and is computed directly from the data. If, when the two graphs appear overlayed, it seems that the model curve is approximating the Kaplan-Meier step function well, especially near the end, the model is considered an appropriate representation of the process.

Recovery Hazard Function

The graph of the Kaplan-Meier estimate of the recovery survival curve for those in the infective diagnostic group is presented in chart I. The steep initial decline indicates that the recovery hazard is higher during the first few years. After this critical period, the survival curve levels off fairly quickly, indicating that the hazard function has become very small. Thus, a curve of a simple mathematical form might not fit the data very well; and indeed, this lack of fit is borne out in chart I for the infective diagnostic group. Because of this lack of fit, it was decided to try a two-part Weibull function to model the data. This approach, described in [6], consists of using one curve for the first part and a separate curve for the second part. The composite is described by letting $\vec{\beta}_1 = (\beta_{11},\ldots,\beta_{1n})$ be the vector of covariate

parameters for the recovery hazard function, $h_{r_1}(t,\vec{x})$, in (18), for the first process. Let γ_1 be the corresponding shape parameter for the first process. Let the respective parameters for the second process be given by $\vec{\beta}_2$ and γ_2. Further, let t_0 be the point at which the two parts are joined, the cutoff point.

Then, the survival curve is given by

$$S_r(t,\vec{x}) = \begin{cases} S_1(t,\vec{x}); & \text{if } t \leq t_0 \\ S_1(t_0,\vec{x}) \times \dfrac{S_2(t,\vec{x})}{S_2(t_0,\vec{x})}; & \text{if } t > t_0 \end{cases}, \quad (20)$$

where

$$S_1(t,\vec{x}) = e^{-\left[\frac{t}{e^{\beta_{01}+\vec{\beta}_1'\vec{x}}}\right]^{e^{\gamma_1}}} \quad (21)$$

and

$$S_2(t,\vec{x}) = e^{-\left[\frac{t}{e^{\beta_{02}+\vec{\beta}_2'\vec{x}}}\right]^{e^{\gamma_2}}} \quad (22)$$

Determining Cutoff Point

The cutoff point was determined by examination of the data and consideration of the medical reexamination procedures. A lower bound of 18 months and an upper bound of 30 months appeared appropriate because some time must elapse between the medical reexamination diary and program termination. An attempt was made to estimate the cutoff point from the data in a model with no covariates, but the computer program would not converge. Instead, for the infective diagnostic group, models were estimated separately with cutoff points at 20 months, 23 months, 25 months, 27 months, and 30 months after the date of entitlement. On examination of corresponding graphs in comparison with the Kaplan-Meier estimates, the model with a cutoff point at 27 months was deemed the most appropriate. This model was estimated for all diagnostic groups. Chart II presents a comparison of the full model estimate, the reduced model estimate, and the Kaplan-Meier estimate of the recovery survival curve for several of these groups. The full model compared with the Kaplan-Meier estimate shows that the fit is quite satisfactory and is typical of the fit for all diagnostic groups. The reduced model is discussed below.

Two-Part Hazard Function

Chart III presents the two-part recovery hazard function that was estimated for nonblack males aged 18-34 at the time of entitlement, with 0-8 years of education, who had a PIA of $1-$299, and had a white-collar job before onset of the disability, and whose disabling condition falls in the infective diagnostic group. Graphically, the two parts of the model are quite evident. During the first 27 months, the curve rises, indicating an increasing tendency to recover. A large jump occurs at 27 months, the point at which the second part of the model starts. The hazard function then decreases slowly, indicating a gradual decrease in the tendency to recover over time after 27 months.

The recovery hazard function for the endocrine group with the same characteristics (chart IV), displays a similar pattern, except that after 27 months the tendency toward increased recovery is somewhat muted. A comparison of the scales on the vertical axes indicates that the recovery hazard for this group is considerably lower in value than that of the infective group.

Coefficient Estimates

Table I presents the coefficient estimates of the recovery hazard functions for all diagnostic groups. For six groups, covariates are included in both parts of the model (infective, mental disorders, nervous system, circulatory, musculoskeletal, and accidents). For the digestive and neoplasms groups, only the first part of the model (recoveries in 27 months or less)

Chart I.—Recovery survival curve for infective and parasitic diseases

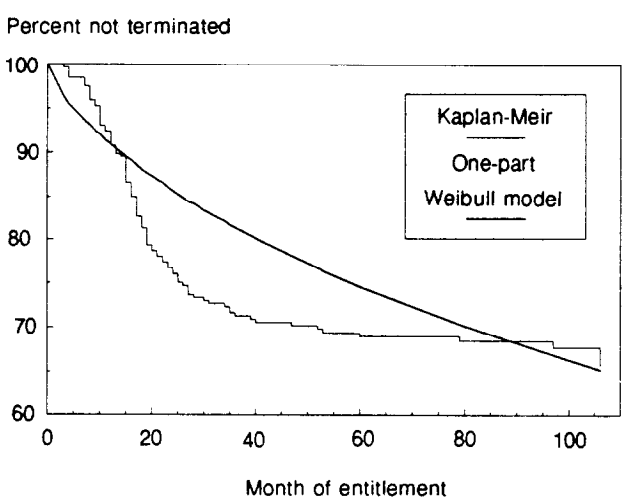


Chart II.—Two-part recovery survival curves, by selected diagnostic groups

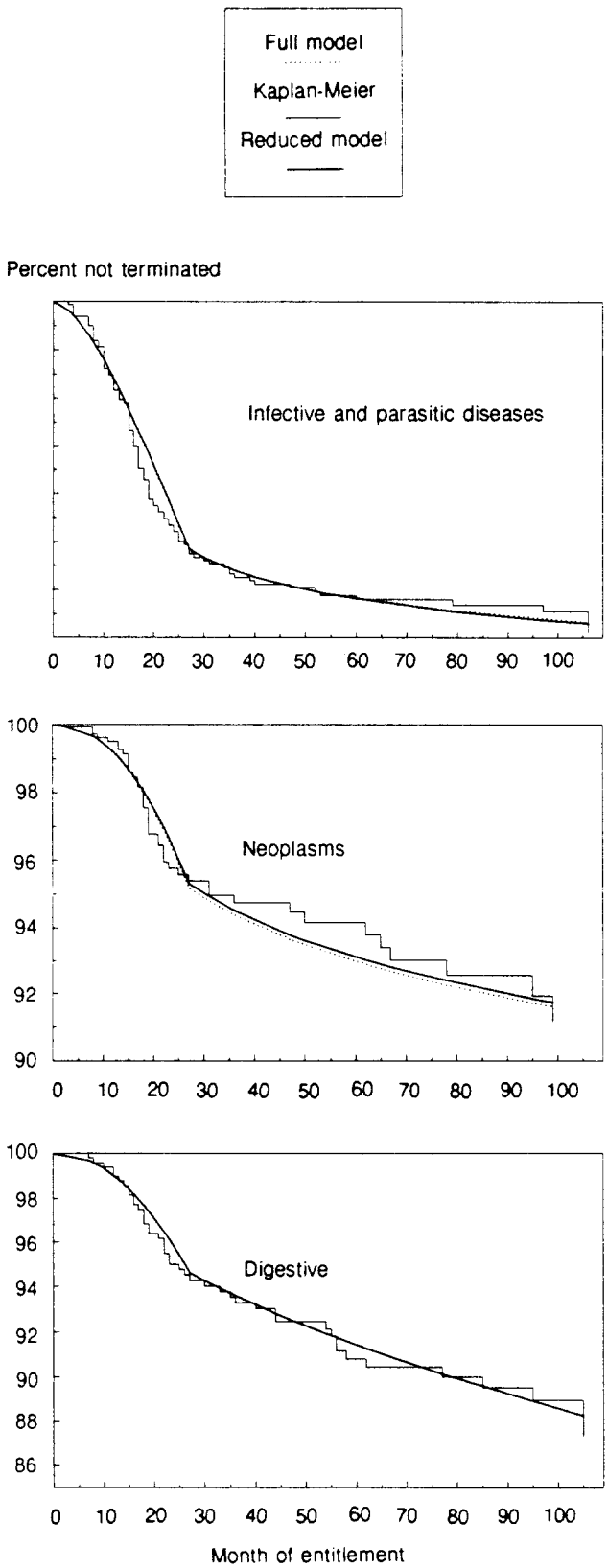


Chart III.—Recovery hazard function for those in infective and parasitic group, by selected characteristics (nonblack males, aged 18-34, 0-8 years of education, PIA $1-$299, white-collar job)

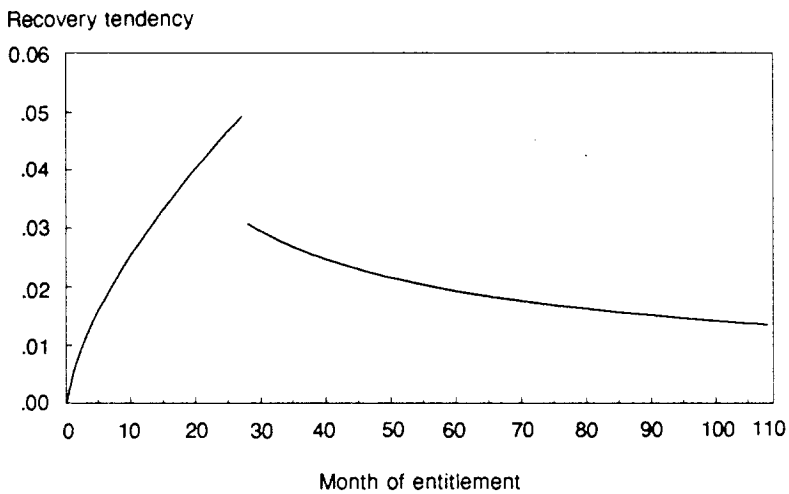
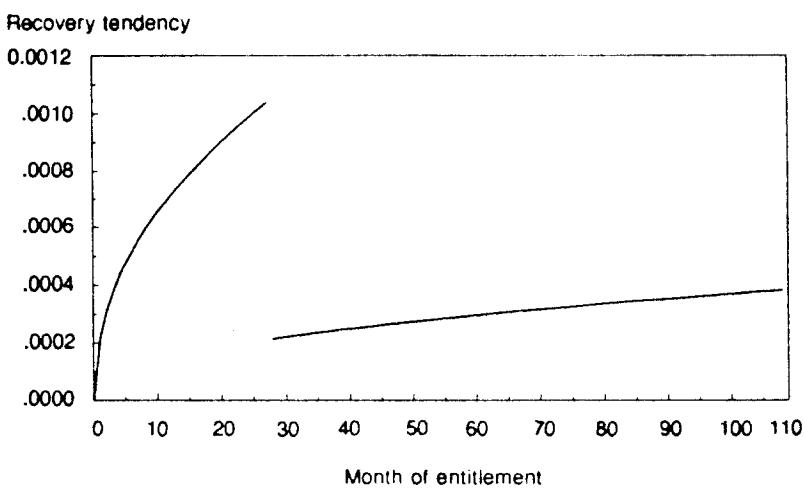


Chart IV.—Recovery hazard function for those in the endocrine group, by selected characteristics (nonblack males, aged 18-34, 0-8 years of education, PIA $1-$299, white-collar job)



incorporates the covariates because of the small number of recoveries after 27 months in the program. The remaining five diagnostic groups contain no covariates in either part of the model for a variety of reasons. Total sample sizes for the genitourinary and congenital anomalies groups are 170 and 198. No covariates are incorporated into these models because of the small sample sizes. Because only 15 of the 1,152 persons in the respiratory group recovered, model estimation using the covariates was not possible. Although the sample size of the endocrine group is relatively large—620—the computer program, which involves an iterative procedure, would not converge when the covariates were introduced, despite several attempts at different starting points. Covariates are not incorporated into the model for the "other" diagnostic group because the mix of disabling conditions within the group makes interpretation of model coefficients of little interest and would not greatly improve prediction.

Some patterns in the tendency to recover as a function of the covariates are described next. Statistical significance of a covariate is assessed with t-tests at the 0.05 level.

A summary of the association of the covariates with the tendency to recover, as described by the coefficients in table I, is described. A more detailed discussion follows the summary.

Within a diagnostic group, covariates are significant more often in the second part of the model than in the first part. Except for one case, a covariate that is significant in the first part of the model is also significant in the second part of the model. The only exception is the occupation group covariate in the accidents diagnostic group. There are numerous examples where a covariate is significant in the second part of the model and not in the first part. The constants and estimated model coefficients tend to be much larger in the second part of the model, indicating that the tendency to recover after 27 months is considerably weaker than in the earlier months.

Sex.—Sex is a significant covariate in the second parts of the model (recovery after 27 months in the program) for the infective, nervous system, circulatory, musculoskeletal, and accidents groups. In addition, it is significant for the first parts of the model for the circulatory, musculoskeletal, and accidents groups. In all cases, the estimated model coefficient is positive, indicating that women have a lower tendency to recover than do men.

Race.—Race does not appear to be an important variable. It is significant in only the second parts of the models for the nervous system and accidents groups. In both these cases, the estimated model coefficients are positive, indicating that blacks have a lower tendency to recover than nonblacks. However, there does not appear to be any pattern of association over diagnostic groups.

Age at entitlement.—Age at entitlement appears to be a key variable for many of the diagnostic groups. It is significant in both parts of the models for all groups for which covariates are included in both parts—the infective, mental disorders, nervous system, circulatory, musculoskeletal, and accidents groups. For the two groups that incorporate covariates in only the first parts of the model, it is significant for the first part of the model for the digestive group, but is not significant for the neoplasms group. In almost all cases, the trend with age is that older workers have a weaker tendency to recover than younger workers, both before and after the first 27 months after entitlement.

Education.—Education is another variable exhibiting a consistent pattern over diagnostic groups. For five of the six groups containing all covariates in both parts, education is a significant variable. For the infective diagnostic group, education is significant in only the second part of the model. For almost all diagnostic groups, as educational level increases, the tendency to recover increases.

Occupation.—Occupation did not show as clear a pattern as education across diagnostic groups. It is significant in the second parts of the model for the infective, mental disorders, nervous system, and circulatory groups, and in the first parts of the model for the mental disorders, nervous system, and accidents groups. The general trend is that workers in the white-collar occupations have a stronger tendency to recover than do workers in the other occupation groups.

Primary insurance amount.—The PIA is another variable that appears to have a more complex interpretation. It is significant in all parts of models containing all covariates. In most cases, individuals in the lowest PIA group have a greater tendency to recover, compared with individuals in the three highest PIA groups. Among the PIA groups, a "quadratic" trend or an inverted U-shaped effect on the tendency to recover seems to apply. For example, the coefficients for the PIA groups in the first part of the recovery hazard function for the musculoskeletal group are 0.41, 0.9, and 0.046. The same up-down pattern is present in the coefficients for the second part—10.13, 23.01, and 15.57. This pattern means that as PIA level increases, workers have less of a tendency to recover, with the pattern starting to reverse for the highest PIA level.

Reduced model

For projection purposes, it was decided to eliminate a variable from a part of the model if the coefficients for all categories of that variable were not significant at the 0.05 level, using the t-test. This decision was made to minimize the variance in the projections. There was one exception to this rule. In the circulatory group, the only significant coefficient in the occupation variable for the first part of the model corresponds to the category "other" and it was decided to eliminate the occupation variable from the first part of the model. A likelihood-ratio test indicated that the reduced model without the occupation group variable in the first part was acceptable. These reduced models are used in all predictions. Table II contains the coefficient estimates for the reduced models. Chart II displays the recovery survival curves for the reduced model, compared with the Kaplan-Meier and the full model, for three diagnostic groups. The finding, typical of all diagnostic groups, is that the survival curve for the reduced model is very close to that for the full model and, if anything, is slightly closer to the Kaplan-Meier estimate. This finding indicates that no slippage in fit is caused by omitting the nonsignificant variables.

Death Hazard Function

Similar to the analysis plan completed for the recovery hazard function, a death hazard function is computed for each diagnostic group. This death hazard function describes the tendency for a beneficiary to terminate the first entitlement period by death at a particular point in time. The larger the value of the hazard function, the greater the tendency or risk. Unlike the analysis of recovery hazards, death hazards were amenable to simpler, one-part Weibull models for almost all diagnostic groups. Therefore, a different computer program is used to estimate the parameters of almost all death hazard functions. The formula for the parameters, although equivalent to (18), is given by

$$h_d(t,x_1,\ldots,x_n) = \frac{\gamma\, t^{\gamma-1}}{\left[\alpha\, e^{\beta_1 x_1 + \ldots + \beta_n x_n}\right]^{\gamma}}. \qquad (23)$$

The only exception was the neoplasms group, for which a one-part Weibull model did not fit. Comparison of the Kaplan-Meier curve obtained from the raw data and the one-part Weibull survival curve indicated a lack of fit for most of the sex-age covariate subgroups. The Kaplan-Meier curve dropped more sharply at the beginning of the time period than the model would allow. Because this difficulty was similar to the problems experienced with the recovery hazard function, the death hazard was modeled for neoplasms with the same model used for the recovery hazard function, a two-part Weibull model with time intervals of less than 27 months and greater than or equal to 27 months. This model, without covariates, had an acceptable fit. When covariates were introduced into the model, the computer program estimating model coefficients failed to converge, so that the model without covariates is used.

Covariate Effects

Table III shows the coefficient estimates for the death hazard functions for the full model with age and sex covariates included and for the reduced model. In the initial model estimations that included all of the covariates, only in rare instances were other coefficients statistically significant. However, covariates other than age and sex exhibited no pattern over diagnostic groups. Therefore, covariates other than age and sex were not used in the final models.

For seven groups—accidents, endocrine, mental disorders, nervous system, circulatory, musculoskeletal, and other—both sex and age were significant covariates as determined by two-sided t-tests at the 0.05 significance level. In all these groups, with the exception of the endocrine group, the effect of the covariate is as expected: Men and older beneficiaries exhibit a greater death tendency. For example, in the mental disorders group, the model coefficient for the sex covariate is positive indicating that the death hazard for women is lower than that for men. The coefficients for the variables describing the age covariate are -0.84 and -1.61, which means that beneficiaries in the groups aged 35-49 and 50-61 have a higher death hazard than beneficiaries in the youngest age group. The coefficient for the oldest age group is the most negative, indicating the death hazard is highest for this age group.

The endocrine group exhibits a somewhat different pattern. The coefficients for age are positive, indicating that older beneficiaries have a lower death hazard than younger beneficiaries. This result is consistent with the raw data. It is not clear why this diagnostic group differs from the others.

In summary, the effect of age and sex on the death hazards was for the most part predictable. Women have

Table I.—Coefficient estimates of the recovery hazard functions

Variable [1]	Infective		Neoplasms		Endocrine	
	Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic
Part I: Less than 27 months						
Constant	3.43	[2]11.06	4.8	[2] 9.06	6	[2]7.23
Sex						
Women	.21	1.11	.28	1.22	---	---
Occupation						
Service	.0048	.02	−.13	−.52	---	---
Farming	.84	1.29	−.17	−.35	---	---
Manufacturing	−024	−.10	−.095	−.43	---	---
Unknown and miscellaneous	.011	.04	.049	.19	---	---
Years of education						
9-11	−.038	−.21	−.21	−.88	---	---
12	−.088	−.42	−.27	−1.17	---	---
13 or more	−.62	−1.88	−.056	−.17	---	---
Unknown	−.38	−1.36	3.34	−.21	---	---
Primary insurance amount						
$300-499	.26	1.53	.41	1.52	---	---
$500-$699	.45	[2]2.05	.27	.90	---	---
$700 or more	−.71	−1.78	−.92	[2] −2.19	---	---
Age at entitlement						
35-49	.32	1.78	−.5	−1.35	---	---
50 or older	−84	[2] 3.82	−.21	−.62	---	---
Race						
Black	−.21	-1.31	−.13	−.50	---	---
Constant: Gamma	.51	[2]5.10	.82	[2]5.47	.38	1.27

See footnotes at end of table.

a lower death hazard than men, and younger beneficiaries have a lower death hazard than older beneficiaries.

Table III also contains the coefficients for the reduced model of the death hazard functions. For three of the diagnostic groups (infective, digestive, and genitourinary) age was not a significant covariate and was dropped from the model. Sex is an important covariate in these groups, with women exhibiting a lower death hazard than men.

The respiratory group is modeled using only age as a covariate to ensure that the constants in the Weibull model did not have confidence intervals with negative values. In the remaining diagnostic group—congenital anomalies—sex and age are not significant covariates, so that a one-part Weibull model without covariates is used.

Mental disorders		Nervous system		Circulatory		
Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic	Variable [1]
						Part I: Less than 27 months
4.31	[2]15.39	5.28	[2]10.78	5.1	[2]20.4	Constant
						Sex
.16	1.00	.29	1.32	.24	[2]2.4	Women
						Occupation
.5	[2]2.17	−.97	[2] −2.94	−.059	−.49	Service
1.01	1.66	−.65	−1.23	−.04	−.17	Farming
.58	[2]2.64	−.33	−1.06	−.13	−1.43	Manufacturing
.13	.76	−.66	[2] −2.20	−.2	[2] −2.06	Unknown and miscellaneous
						Years of education
−.4	−1.82	−.39	−1.30	−.35	[2] −3.80	9-11
−.19	−0.86	−.6	[2] −2.14	−.23	[2] −2.35	12
−.62	[2] −2.70	−.99	[2] −3.00	−.44	[2] −3.38	13 or more
−.12	−.43	.24	.38	.023	.08	Unknown
						Primary insurance amount
.52	[2]3.25	.26	1.18	.16	1.33	$300-$499
.59	[2]2.95	.64	[2]2.21	.18	1.38	$500-$699
.48	1.12	−.054	−.13	−1.32	[2] −7.33	$700 or more
						Age at entitlement
.07	.50	.19	.86	−.34	[2] −2.00	35-49
1.11	[2]4.11	.83	[2]3.07	.095	.56	50 or older
						Race
.075	.42	.0083	.03	.18	1.50	Black
.56	[2]5.09	0.52	[2]3.47	.77	[2]10.13	Constant: Gamma

Table I.—Coefficient estimates of the recovery hazard function—**Continued**

Variable [1]	Respiratory		Digestive		Genitourinary	
	Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic
Part I: Less than 27 months						
Constant	5.82	[2]7.19	3.73	[2]9.10	4.48	[2]11.49
Sex						
Women	---	---	.23	.92	---	---
Occupation						
Service	---	---	−.011	−.04	---	---
Farming	---	---	.15	.30	---	---
Manufacturing	---	---	−.003	−.01	---	---
Unknown and miscellaneous	---	---	−.29	−1.00	---	---
Years of education						
9-11	---	---	−.11	−.39	---	---
12	---	---	−.08	−.30	---	---
13 or more	---	---	−.5	−1.72	---	---
Unknown	---	---	4.19	.11	---	---
Primary insurance amount						
$300-$499	---	---	.31	1.24	---	---
$500-$699	---	---	.69	[2]2.09	---	---
$700 or more	---	---	−.31	−0.77	---	---
Age at entitlement						
35-49	---	---	.46	1.92	---	---
50 or older	---	---	.81	[2] 2.89	---	---
Race						
Black	---	---	.69	1.47	---	---
Constant: Gamma	.61	[2]1.97	.82	[2]4.32	.79	---

See footnotes at end of table.

Musculoskeletal		Congenital anomalies		Accidents		Other		
Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic	Variable [1]
								Part I: Less than 27 months
3.89	[2]21.61	4.54	[2]13.76	3.55	[2]25.36	5.4	[2]22.50	Constant
								Sex
.48	[2]4.36	- - -	- - -	.34	[2]3.09	- - -	- - -	Women
								Occupation
.25	1.56	- - -	- - -	.31	[2]2.38	- - -	- - -	Service
.2	.80	- - -	- - -	.096	.56	- - -	- - -	Farming
−.14	−1.17	- - -	- - -	−.0094	−.09	- - -	- - -	Manufacturing
−.12	−.92	- - -	- - -	.069	−.63	- - -	- - -	Unknown and miscellaneous
								Years of education
−.27	[2] −2.25	- - -	- - -	−.26	[2] −2.60	- - -	- - -	9-11
−.59	[2] −4.92	- - -	- - -	−.31	[2] −3.16	- - -	- - -	12
−.39	[2] −2.17	- - -	- - -	−.3	[2] −2.31	- - -	- - -	13 or more
7.25	[2]7.25	- - -	- - -	−.2	.59	- - -	- - -	Unknown
								Primary insurance amount
.41	4.10	- - -	- - -	.3	[2]3.53	- - -	- - -	$300-$499
.9	[2]6.43	- - -	- - -	.77	[2]6.42	- - -	- - -	$500-$699
.046	.23	- - -	- - -	.45	[2]2.50	- - -	- - -	$700 or more
								Age at entitlement
.14	1.44	- - -	- - -	−.017	−.22	- - -	- - -	35-49
1.19	[2]8.50	- - -	- - -	−63	[2]5.73	- - -	- - -	50 or older
								Race
.15	1.07	- - -	- - -	−.22	1.83	- - -	- - -	Black
.56	[2]8.48	.67	[2]2.79	.57	[2]10.56	.36	[2]3.27	Constant: Gamma

Table I.—Coefficient estimates of the recovery hazard functions—**Continued**

Variable [1]	Infective		Neoplasms		Endocrine	
	Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic
Part II: Time greater than or equal to 27 months						
Constant	−69.92	[2] −6.32	−8.95	[2] −8.69	7.21	[2] 4.27
Sex						
Women	44.02	[2] 3.50	---	---	---	---
Occupation						
Serve	24.09	[2] 7.70	---	---	---	---
Farming	99.24	[2] 7.04	---	---	---	---
Manufacturing	−16.36	−1.62	---	---	---	---
Unknown and miscellanous	−20.52	−1.62	---	---	---	---
Years of education						
9-11	−53.75	[2] −3.32	---	---	---	---
12	−41.41	[2] −2.32	---	---	---	---
13 or more	−45.65	[2] −4.44	---	---	---	---
Unknown	66.66	[2] 4.81	---	---	---	---
Primary insurance amount						
$300-$499	63.99	[2] 4.08	---	---	---	---
$500-$699	47.56	[2] 4.92	---	---	---	---
$700 or more	66.57	[2] 3.83	---	---	---	---
Age at entitlement						
35-49	24.67	1.68	---	---	---	---
50 or older	90.29	[2] 10.12	---	---	---	---
Race						
Black	−19.7	−1.66	---	---	---	---
Constant: Gamma	−3.49	[2] 21.81	−3.81	[2] 17.32	.34	.49

See footnotes at end of table.

Mental disorders		Nervous system		Circulatory		
Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic	Variable [1]
						Part II: Time greater than or equal to 27 months
-25.27	[2] -4.43	-225	[2] -20.22	-20.34	[2] -10.48	Constant
						Sex
2.08	.78	75.11	[2]9.86	17.23	[2]6.18	Women
						Occupation
7.29	[2]3.05	4.16	.54	6.28	[2]2.42	Service
45.01	[2]6.22	15.97	1.15	13.96	[2]2.08	Farming
18.44	[2]6.47	-31.06	-1.72	4.17	[2]2.61	Manufacturing
11.08	[2]3.38	61.25	[2]2.86	6.86	[2]4.18	Unknown and miscellaneous
						Years of education
-22.1	[2] -2.06	126.66	[2]42.50	-5.21	[2] -2.26	9-11
-27.19	[2] -2.27	-15.85	-.83	-11.24	[2] -4.34	12
-42.86	[2] -3.00	-49.76	[2] -3.35	-11.19	[2] -3.54	13 or more
-31.63	[2] -2.85	-65.44	[2] -13.86	-14.35	[2] -3.24	Unknown
						Primary insurance amount
34.06	[2]4.51	76.72	[2]5.14	13.38	[2]5.29	$300-$499
44.15	[2]4.32	157.57	[2]8.92	21.86	[2]9.63	$500-$699
13.42	1.30	18.77	[2]4.98	7.71	1.34	$700 or more
						Age at entitlement
14.73	[2]4.55	51.16	[2]2.82	12.49	[2]5.43	35-49
36.75	[2]5.38	161.21	[2]10.03	36.18	[2]12.83	50 or older
						Race
-1.47	-.63	220.86	[2]8.05	-2.1	-1.00	Black
-3.16	[2] -14.36	-4.33	[2] -43.30	-2.61	[2] -23.73	Constant: Gamma

Table I.—Coefficient estimates of the recovery hazard functions—Continued

Variable [1]	Infective		Neoplasms		Endocrine	
	Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic
Part II: Time greater than or equal to 27 months						
Constant	33.92	[2]33.92	8.38	[2] 8.91	−6.34	−.48
Sex						
Women	---	---	---	---	---	---
Occupation						
Service	---	---	---	---	---	---
Farming	---	---	---	---	---	---
Manufacturing	---	---	---	---	---	---
Unknown and miscellaneous	---	---	---	---	---	---
Years of education						
9-11	---	---	---	---	---	---
12	---	---	---	---	---	---
13 or more	---	---	---	---	---	---
Unknown	---	---	---	---	---	---
Primary insurance amount						
$300-$499	---	---	---	---	---	---
$500-$699	---	---	---	---	---	---
$700 or more	---	---	---	---	---	---
Age at entitlement						
35-49	---	---	---	---	---	---
50 or older	---	---	---	---	---	---
Race						
Black	---	---	---	---	---	---
Constant: Gamma	−5.07	[2] −9.22	−.61	−1.61	−3.22	[2] −7.00

[1] Reference groups: Sex, men; occupation, white collar; education, 0-8 years; PIA, $1-$299; age, 18-34; race; nonblack; and diagnostic group, circulatory system.

[2] Significant at the 0.05 level.

Musculoskeletal		Congenital anomalies		Accidents		Other		
Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic	Coefficient	T-statistic	Variable [1]
								Part II: Time greater than or equal to 27 months
-15.95	-1.10	-37.57	[2] -2.43	-51	[2] -3.12	8.51	[2] 5.32	Constant
								Sex
11.6	[2] 2.32	---	---	17.33	[2] 4.68	---	---	Women
								Occupation
6.51	1.69	---	---	2.52	.56	---	---	Service
2.21	1.44	---	---	-.68	-.09	---	---	Farming
1.52	.60	---	---	-3.26	-.76	---	---	Manufacturing
2.69	.91	---	---	1.92	.44	---	---	Unknown and miscellaneous
								Years of education
-9.59	[2] -2.29	---	---	-4.91	[2] -2.48	---	---	9-11
-12.65	[2] -2.38	---	---	-9.02	[2] -4.85	---	---	12
-18.61	[2] -2.39	---	---	-16.77	[2] -4.74	---	---	13 or more
-2.14	-.23	---	---	.48	.06	---	---	Unknown
								Primary insurance amount
10.13	[2] 2.13	---	---	11.62	[2] 4.70	---	---	$300-$499
23.01	[2] 2.27	---	---	37.59	[2] 4.70	---	---	$500-$699
15.57	[2] 2.06	---	---	34.06	[2] 4.38	---	---	$700 or more
								Age at entitlement
11.11	[2] 2.15	---	---	12.29	[2] 2.95	---	---	35-49
30.63	[2] 2.27	---	---	43.27	[2] 3.82	---	---	50 or older
								Race
-.87	-.42	---	---	13.38	[2] 3.08	---	---	Black
-2.46	[2] -5.35	-3.45	[2] -12.78	-3.02	[2] -13.13	-.5	-.93	Constant: Gamma

Table II.—Coefficient estimates of recovery hazard functions, reduced model

Variable [1]	Infective	Neoplasms	Endocrine	Mental disorders	Nervous disorders	Circulatory
Part I: Time less than 27 months						
Constant	3.36	4.57	6	4.39	5.43	5.03
Sex:						
Women	---	---	---	---	---	.28
Occuptation:						
Service	---	---	---	.5	-1	---
Farming	---	---	---	.93	−.81	---
Manufacturing	---	---	---	.57	−.41	---
Unknown and miscellaneous	---	---	.09	−.76	---	
Years of education:						
9-11	---	---	---	−.39	−.35	−.35
12	---	---	---	−.17	−.54	−.2
13 or more	---	---	---	−.62	−.94	−.37
Unknown	---	---	---	−.1	.3	.03
Primary insurance amount:						
$300-$499	.21	.28	---	.5	.24	.13
$500-$699	.33	.06	---	.53	.52	.13
$700 or more	−.75	−.97	---	.45	−.13	-1.38
Age at entitlement:						
35-49	.36	---	---	.09	.22	−.34
50 or older	.9	---	---	1.14	.87	.1
Race:						
Black	---	---	---	---	---	---
Constant: Gamma	.5	.81	.38	.56	.51	.77

See footnotes at end of table.

Respiratory	Digestive	Genitourinary	Musculoskeletal	Congenital anomalies	Accidents	Other	Variable [1]
							Part I: Time less than 27 months
5.82	3.74	4.48	3.86	4.54	3.58	5.4	Constant
							Sex:
---	---	---	.55	---	.34	---	Women
							Occupation:
---	---	---	---	---	.32	---	Service
---	---	---	---	---	09	---	Farming
---	---	---	---	---	-.01	---	Manufacturing
---	---	---	---	---	.09	---	Unknown and miscellaneous
							Years of education:
---	---	---	-.27	---	-.27	---	9-11
---	---	---	-.59	---	-.33	---	12
---	---	---	-.35	---	-.34	---	13 or more
---	---	---	7.25	---	.17	---	Unknown
							Primary insurance amount
---	.14	---	.36	---	.3	---	$300-$499
---	.47	---	.83	---	.76	---	$500-$699
---	-.56	---	-.009	---	.43	---	$700 or more
							Age at entitlement:
---	.57	---	.17	---	-.01	---	35-49
---	.94	---	1.25	---	.63	---	50 or older
							Race:
---	---	---	---	---	---	---	Black
.62	.79	.79	.56	.67	.57	.36	Constant: Gamma

Table II.—Coefficient estimates of recovery hazard functions, reduced model—**Continued**

Variable [1]	Infective	Neoplasms	Endocrine	Mental disorders	Nervous system	Circulatory
Part II: Time greater than or equal to 27 months						
Constant	-66.6	-8.95	7.21	-27.26	-219.09	-20.63
Sex:						
Women	39.17	- - -	- - -	73.18	17.12	- - -
Occupation:						
Service	28.44	- - -	7.13	5.75	5.71	- - -
Farming	46.92	- - -	45.97	18.69	13.93	- - -
Manufacturing	-22.2	- - -	19.14	-30.18	4.1	- - -
Unknown and miscellaneous	-23.48	- - -	10.73	60.53	6.52	- - -
Years of education:						
9-11	-51.42	- - -	-22.35	124.77	-5.06	- - -
12	-34.23	- - -	-27.60	-14.46	-10.91	- - -
13 or more	-44.57	- - -	- - -	-44.12	-45.48	-10.88
Unknown	34.14	- - -	-32.39	-61.77	-14.3	- - -
Primary insurance amount:						
$300-$499	61.81	- - -	35.28	75.34	13.38	- - -
$500-$699	47.96	- - -	45.41	152.62	21.93	- - -
$700 or more	23.3	- - -	13.25	15.94	7.91	- - -
Age at entitlement:						
35-49	24.98	- - -	15.77	50.91	12.41	- - -
50 or older	84.55	- - -	39.31	157.04	36.05	- - -
Race:						
Black	- - -	- - -	- - -	238.38	- - -	- - -
Constant: Gamma	-3.42	-3.81	.34	-3.20	-4.3	-2.61

[1] Reference group: Sex, men; occupation, white collar; education, 0-8 years; PIA, $1-$299; age at entitlement, 18-34; race, nonblack.
[2] Significant at the 0.05 level.

Respiratory	Digestive	Genitourinary	Musculoskeletal	Congenital anomalies	Accidents	Other	Variable [1]
							Part II: Time greater than or equal to 27 months
33.92	8.38	-6.34	-14.67	-37.57	-53.03	8.51	Constant
							Sex:
---	---	---	12.26	---	18.47	---	Women
							Occupation:
---	---	---	---	---	---	---	Service
---	---	---	---	---	---	---	Farming
---	---	---	---	---	---	---	Manufacturing
---	---	---	---	---	---	---	Unknown and miscellaneous
							Years of education:
---	---	---	-10.18	---	-4.55	---	9-11
---	---	---	-13.71	---	-8.45	---	12
---	---	---	20.72	---	-15.78	---	13 or more
---	---	---	-4.01	---	2.18	---	Unknown
							Primary insurance amount:
---	---	---	10.11	---	11.36	---	$300-$499
---	---	---	23.38	---	37.54	---	$500-$699
---	---	---	15.01	---	33.91	---	$700 or more
							Age at entitlement:
---	---	---	11.65	---	11.93	---	35-49
---	---	---	32.36	---	44.1	---	50 or older
							Race:
---	---	---	---	---	14.56	---	Black
-5.07	-.61	-3.22	-2.51	-3.45	-3.05	-.5	Constant: Gamma

Table III.—Coefficient estimates of death hazard functions, one-part model

Variable [1] and diagnostic group	Beta	T-statistic	New Beta
Infective			
Sex:			
Women	0.79	[2]2.08	0.79
Age at entitlement:			
35-49	[3]0	---	---
50 or older	.39	1.62	---
Constant: Gamma	1.05	[2]8.44	1.05
Constant: Alpha	188.31	[2]5.34	229.12
Endocrine			
Sex:			
Women	.35	[2]2.22	2.22
Age at entitlement:			
35-49	.45	1.84	1.84
50 or older	.62	[2]2.81	2.81
Constant: Gamma	.97	[2]13.40	13.40
Constant: Alpha	119.07	[2]4.69	4.69
Mental disorders			
Sex:			
Women	.68	[2]4.47	4.47
Age at entitlement:			
35-49	.84	[2] -3.76	-3.76
50 or older	-1.61	[2] -7.18	-7.18
Constant: Gamma	1.11	[2]14.76	14.76
Constant: Alpha	1226.82	[2]3.87	3.87
Nervous system			
Sex:			
Women	.63	[2]3.35	3.35
Age at entitlement:			
35-49	.32	-1.12	-1.12
50 or older	.85	[2] - 3.33	-3.33
Constant: Gamma	.96	[2]12.85	12.85
Constant: Alpha	754.37	[2]3.56	3.56
Circulatory			
Sex:			
Women	.70	[2]10.89	10.89
Age at entitlement:			
35-49	-.65	[2] -2.48	-2.48
50 or older	-.89	[2] -3.44	-3.44
Constant: Gamma	1.00	[2]42.90	42.90
Constant: Alpha	391.46	[2]3.84	3.84
Respiratory			
Sex:			
Women	.16	1.33	---
Age at entitlement:			
35-49	-.29	[2] -2.30	-.30
50 or older	---	---	---
Constant: Gamma	1.13	[2]20.26	1.13
Constant: Alpha	211.61	[2]8.09	219.86
Digestive			
Sex:			
Women	.59	[2]3.27	.59
Age at entitlement:			
35-49	-.53	-1.42	---
50 or older	-.44	-1.20	---
Constant: Gamma	.89	[2]14.54	0.89
Constant: Alpha	187.77	[2]2.77	119.61
Genitourinary			
Sex:			
Women	1.09	[2]3.17	1.03
Age at entitlement:			
35-49	.06	.15	---
50 or older	-.59	-1.57	---
Constant: Gamma	.91	[2]7.37	.88
Constant: Alpha	137.36	[2]3.02	108.97
Musculoskeletal			
Sex:			
Women	.46	[2]4.60	4.60
Age at entitlement:			
35-49	-.32	-1.20	-1.20
50 or older	-.67	[2] -3.46	-3.46
Constant: Gamma	1.20	[2]18.05	18.05
Constant: Alpha	805.16	[2]3.66	3.66
Congenital anomalies			
Sex:			
Women	.99	1.56	---
Age at entitlement:			
35-49	-.58	-.60	---
50 or older	-1.63	-1.79	---
Constant: Gamma	.77	[2]3.60	.76
Constant: Alpha	2025.86	.84	1054.57

See footnotes at end of table.

Table III.—Coefficient estimates of death hazard functions, one-part model—**Continued**

Variable [1] and diagnostic group	Beta	T-statistic	New Beta	Variable [1] and diagnostic group	Beta	T-statistic	New Beta
Accidents				**Other**			
Sex:				Sex:			
Women	.60	[2] 2.38	2.38	Women	.64	[2] 5.35	5.35
Age at entitlement:				Age at entitlement:			
35-49	−.35	−1.16	−1.16	35-49	−.59	[2] −3.19	−3.19
50 or older	−1.05	[2] −3.80	−3.80	50 or older	−.82	[2] −4.76	−4.76
Constant: Gamma	1.05	[2] 11.15	11.15	Constant: Gamma	1.06	[2] 19.06	19.06
Constant: Alpha	1144.03	[2] 3.03	3.03	Constant: Alpha	553.40	[2] 5.35	5.35

[1] Reference group: Sex, men and age at entitlement, 18-34.
[2] Significant at 0.05 level.
[3] Defined to be 0.

Table IV.—Coefficient estimates of death hazard functions for neoplasms diagnostic group, two-part model

Model	Beta	Gamma
Part I		
Coefficient	3.16	−.08
T-statistic	90.54	−2.78
New coefficient	90.54	−2.78
Part II		
Coefficient	−9.58	−2.14
T-statistic	−2.01	−15.16
New coefficient	−2.01	−15.16