LA-9530

UC-41
Issued: December 1982

LA--9530

DE83 007116

Lost-To-Follow-Up Bias in an Occupational Mortality Analysis: A Quantitative Consideration

J. F. Acquavella
G. L. Tietjen
G. S. Wilkinson

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

LOST-TO-FOLLOW-UP BIAS IN AN OCCUPATIONAL MORTALITY ANALYSIS: A QUANTITATIVE CONSIDERATION

by

J. F. Acquavella, G. L. Tietjen, and G. S. Wilkinson

ABSTRACT

A major problem in occupational cohort studies is how to treat study subjects who are lost to follow-up (LTF). The assumptions made concerning their vital status may affect the results of comparative mortality analyses. We have considered this problem within the context of an occupational follow-up study of white male employees at a nuclear facility in Colorado. In this analysis, 568 or 8% of cohort members were LTF. We estimated comparative mortality for the entire cohort by treating LTF workers as lost at employment termination date, as living at the end-of-study date, and with cumulative mortality simulated between 0% and 100%. Our results indicate that simulations of cumulative mortality among employees LTF can be useful in assessing the potential bias caused by LTF mortality assumptions. Further, we propose a general method for assessing LTF bias in occupational analyses.

I. INTRODUCTION

We recently reported a mortality follow-up study of white male workers at a nuclear facility in Colorado.[1] In this study, as in most historical prospective studies, there were many workers for whom vital status could not be ascertained. These study subjects are referred to as lost to follow-up (LTF). The assumptions made concerning their vital status may affect the results of occupational mortality analyses and thus motivated this study.

Workers LTF commonly are treated in one of two ways. First, they can be considered alive at the end of the study, as we did in our previous analysis.[11] This assumption underestimates mortality if any of the missing cohort members is dead. To determine the extent of this underestimate, we traced a 25% random sample (N = 187) of white males considered LTF to determine their vital status. A total of 159 workers (85%) were located, among whom 7 or 4.4% [95% confidence interval (CI) 2.1%-8.8%] were found to be dead. The upper confidence limit of 8.8% was considered the worst case estimate for potential underestimation of total mortality.

A second common technique for dealing with workers LTF is to drop their person-years from the analysis as soon as they become lost (typically, employment termination date). The effect of this treatment is less certain because the LTF group may have greater or lesser mortality than the remainder of the workforce. However, this technique does allow an unbiased estimate of comparative mortality among workers whose vital status is known. In this report, we consider the treatment of workers LTF in the context of our previous analysis, to determine the potential bias introduced by LTF mortality assumptions on estimates of comparative mortality for the total cohort.

II. METHODS

The cohort and the basic method of analysis have been described in detail elsewhere.[1] There has been additional follow-up of cohort members in the interim, so our numbers will differ slightly from those previously reported.[1] The main difference in the present investigation is that cumulative mortality among the LTF group has been simulated at several levels between 0% and 100%. These data were then combined with data for workers of known vital status to determine whether these simulations affect the epidemiologic estimates of overall comparative mortality [that is, the Standardized Mortality Ratio (SMR) and the associated 95% CI].

We created a study file of all LTF workers. Data for workers previously LTF but successfully traced as part of the sample were entered into the file of employees whose vital status was known. The time between the individual LTF employee termination date and the end-of-study date (January 1, 1978) was considered the period at risk of death. We randomly selected LTF individuals to be dead by generating a random number between 0.00 and 1.00. If the random number was less than the fraction of those LTF simulated to be dead, the individual worker was considered to have died. The specific cause of death was similarly randomly assigned according to the fractional distribution of deaths in the total cohort. For these cases, a death date was randomly assigned which was equally probable between the employment termination date and the end-of-study date. In this way, all LTF members were designated alive or dead, and those considered dead were assigned a specific cause of death and a death date. This file was then combined with the total cohort file.

The number of observed and expected deaths (based on National death rates) was tabulated by the Monson program.[2] We computed SMRs as the ratio of the observed to expected deaths (multiplied by 100) along with the associated 95% CIs.[3] The conditions imposed on the LTF group were varied from all living to all dying after their employment termination. However, we place greatest emphasis on simulations of LTF mortality consistent with the range suggested by our sample and follow-up (that is, 4.4%, 95% CI 2.1%-8.8%).

III. RESULTS

Table I characterizes the white male cohort over the period 1951-1978. The average age for the 6777 persons employed was 33 years at entry into the workforce. Their average year of entry of 1963 suggests that the cohort was still relatively young at the end-of-study date. This is an important consideration, because the "healthy worker effect" is greater for younger workers and gradually disappears as the workforce ages.[4,5] Accordingly, estimates of comparative mortality for this cohort might be lower than those for other groups followed over a longer period.

Table II summarizes comparative mortality for selected causes of death. In general, the number of deaths from most causes was less than expected. In particular, mortality from all causes of death, all cancers, and lung cancer was much less than expected. This is indicated by the upper boundary of the associated 95% CIs being less than the null value of 100 (the value of the SMR when the observed number of deaths is equal to the expected number). Death from brain cancer and the leukemias was not more frequent than expected. We selected these five cause-of-death categories to illustrate the effect of LTF mortality on the SMR and 95% CI based on many or few deaths.

The LTF subcohort is described in Table III. The 568 white males had an average age at hire of 29, which is 4 years younger than that of the remainder of the cohort. The mean age at employment termination was 35, compared with 42 for the total cohort. This indicates that many of those LTF worked at this facility for a short period early in their lives. Their young average age at termination suggests that this employment will not be their primary lifetime occupation. Their mean year of entry was the same as that for the total cohort, and they contribute 9429 person-years when all are considered living at the end-of-study date. This number of person-years is roughly 10% of total employee time.

Table IV presents an analysis of comparative mortality for the entire cohort when cumulative mortality among LTF employees is simulated between 0% and 100%. The causes of death considered are the same as those given in Table II. A comparison between the first two rows in Table IV illustrates the difference between treating LTF workers as living (top row) or as lost at employment termination. The respective SMRs will be lower in the former instance because those LTF are effectively "immortal" from their employment termination date to the end-of-study date. This difference is trivial if the SMR is based on many deaths or if those LTF are a small percentage of the total cohort. For example, the all-causes mortality SMR of 61 increases slightly to 64 when LTF workers are considered lost at employment termination. There is also an accompanying shift in the 95% CI. Small increments of the SMR and 95% CI also occur for all cancers and lung cancer. Somewhat larger changes result because of this treatment for the less frequent causes of death: brain cancer and leukemia. However, the extent of these differences would not change any conclusions that were based on the assumption that all those LTF were living at the end-of-study date.

The 5% row in Table IV corresponds roughly to our point estimate of LTF cumulative mortality (4.4%). The results of this simulation for all causes of death, all cancers, and lung cancer are consistent with the estimates generated by counting those LTF as lost at their employment termination. Accordingly, the estimates of comparative mortality in rows 1 and 2 appear relatively unbiased by those LTF for these three important cause-of-death categories. The 5% point estimates for brain cancer and the leukemias do not differ noticeably from the 0% simulation when LTF workers are considered lost at employment termination. Moreover, in epidemiologic studies, estimation is properly directed toward the range of the underlying risk ratio (particularly the upper limit), and in our data, the 95% CIs for all the 5% simulations correspond closely to both 0% analyses.

The data in Table IV are presented graphically in Figs. 1-5, which illustrate the effect of mortality simulated in excess of our worst case assumption (that is, 8.8% cumulative LTF mortality). There are two possible frameworks for considering these data. First, we can determine the approximate LTF mortality necessary to change our basic conclusions, derived from rows 1 or 2 (in Table IV), of either lesser or no greater mortality than expected. Alternatively, these data also indicate the extent of mortality in excess of our worst case assumption necessary to reverse our conclusions.

For the all-causes-of-death category, our conclusion of significantly fewer than expected deaths is supported until nearly 40% of those LTF are simulated to have died. This is more than 4 times our worst case estimate. LTF mortality of approximately 40% to 60% supports the conclusion that mortality within the total cohort is not different from that expected. More than 60% of those LTF need to be dead for our original conclusion to be untenable.

Estimates of comparative mortality for the remaining cause-of-death categories are somewhat more sensitive to mortality among LTF employees. Our conclusion of fewer than expected deaths from all cancers appears valid in the range of our worst case estimate. LTF mortality simulated between 20% and 40% would suggest mortality not in excess of that expected. Mortality in excess of 40% would support the conclusion of more deaths than expected from this cause.

Similarly, our conclusion about lung cancer mortality would not change if up to 5% of those LTF were dead. A conclusion of mortality not different from that expected is supported by simulations between 5% and 85%. More than 85% of those LTF need to have died to reverse our original conclusion.

Our simulations for brain cancer and the leukemias support our earlier conclusion of mortality consistent with that expected. For brain cancer, LTF mortality almost 3 times our worst case assumption is necessary to change our conclusion. Our conclusion for the leukemias appears even more stable; mortality exceeding 7 times our worst case simulation would be required to reverse it.

Frequently, occupational mortality analyses make allowance for the "empirical latent period" (time from induction to detection) of cancers.[6] This is done by not counting person-years or deaths among workers from the date of initial employment for an interval that estimates the induction period of specific cancers. Because human cancer takes years to develop after a carcinogenic exposure (the number of years is the induction period), this technique ignores person-years that are not at increased risk of cancer development. In fact, if a cause-effect relationship does exist, varying the discounted person-time interval to maximize the epidemiologic risk measure should estimate the modal value for cancer induction.[6] However, the allowance for cancer induction ages the occupational cohort, which will increase SMRs slightly in the absence of a cause-effect relationship because the "healthy worker effect" decreases with age.[4,5] This bias is normally very small, but it should be remembered in the evaluation of an occupational latency analysis.

Simulations (in Tables V and VI) allowing 5 and 10 years for cancer induction differ from our general analysis (in Table IV) because deaths and person-years are not counted until 5 or 10 years after an individual's date of initial employment. Accordingly, fewer deaths and person-years are included in these analyses, and there is some loss of precision in the comparative mortality estimates (indicated by a widening of the 95% CIs). However, the same parameter should be estimated with and without latency considerations, when no association exists between occupational exposure and disease, which can be seen by comparing Table IV with Tables V and VI. In general, comparative mortality estimates increase slightly for the whole cohort, and the associated 95% CIs are slightly wider for common causes of death and noticeably wider for brain cancer and leukemia. These slight increases probably result from using an older segment of the population to derive the comparative mortality estimates.

The effect on SMRs for the total cohort of simulated LTF mortality, allowing 5 and 10 years for cancer induction, is illustrated in Tables V and VI. Our conclusions for all causes, brain cancer, and leukemia and aleukemia do not change, even when LTF mortality exceeds our worst case simulation. For all cancers and lung cancer, our conclusion of significantly fewer than expected deaths would change to no difference if LTF mortality exceeds our sample-based estimate (that is, 4.4%). However, the SMR and 95% CI for lung cancer would not suggest more deaths than expected even when 100% of those LTF are considered to have died. This stability may result from deleting lung cancer deaths that occurred soon after employment. Under the 5- and 10-year allowance for cancer induction, these deaths are not considered related to employment at the facility.

IV. DISCUSSION

These results suggest that LTF mortality is not a meaningful source of bias in our analysis of comparative mortality among these nuclear facility employees. As such, these results support our conclusion that this cohort does not manifest excess mortality related to employment. This conclusion appears justified even in the event of extremely high mortality among workers LTF.

These results must be taken in context. This analysis does not suggest that mortality among workers LTF is not an important potential bias in occupational mortality studies. Rather, it indicates for this cohort, characterized by young average age and by no indication of increased death due to employment, that even extreme mortality among those LTF has a minimal effect on SMRs for specific causes of death.

Further, this analysis demonstrates that potential LTF bias can be evaluated to determine its importance. This can be accomplished by tracing a sample of LTF workers to determine their vital status, developing a worst case estimate from this sample, and evaluating comparative mortality for the worst case situation. Concurrence between results for the worst case analysis and those based on the total cohort would suggest that mortality among workers LTF is not an important source of bias in the comparative mortality analysis.

ACKNOWLEDGMENTS

We wish to acknowledge the advice of the Epidemiology Group Advisory Committee, which includes Dr. Saxon Graham (chairman), Dr. Roy Shore, Dr. Strother Walker, Dr. Robert Miller, and Dr. Warren Winkelstein. Dr. Richard Waxweiler contributed many useful suggestions for the preparation of our final draft. We also thank Ms. Annabelle Salazar for preparing our many revisions.

REFERENCES

1. G. L. Voelz, G. S. Wilkinson, J. F. Acquavella, G. L. Tietjen, R. M. Brackbill, M. Reyes, and L. D. Wiggs, "An Update of Epidemiologic Studies of Plutonium Workers," Health Phys. (in press).

2. R. R. Monson, "Analysis of Relative Survival and Proportionate Mortality," Comput. Biomed. Res. 7, 325-332 (1974).

3. K. J. Rothman and J. D. Boice, "Epidemiologic Analysis with a Programmable Calculator," Department of Health, Education, and Welfare, National Institutes of Health publication NIH-79-1649 (1979).

4. J. R. Goldsmith, "What We Expect from an Occupational Cohort," J. Occup. Med. 17, 126-127 (1975).

5. A. I. McMichael, S. G. Haynes, and M. A. Tyroler, "Observations on the Evaluation of Occupational Mortality Data," J. Occup. Med. 17, 128-131 (1975).

6. K. J. Rothman, "Induction and Latent Periods," Am. J. Epidemiol. 114, 253-259 (1981).

TABLE I. Selected Characteristics of White Males, 1951-1978

Employees	6 777
Deaths	425
LTF	568
Mean Age at Entry[a]	33
Mean Year of Entry[a]	1963
Mean Age at Death[a]	57
Mean Age at Employment Termination[a]	42
Total Person-Years[b]	91 277
Person-Years (5 years since initial employment)[b]	70 570
Person-Years (10 years since initial employment)[b]	42 376

[a]Excluding those LTF.

[b]Those LTF dropped from the analysis at their employment termination date.

TABLE II. Comparative Mortality for Selected Causes of Death, 1951-1978

Cause of Death (ICD)[a]	Observed	Expected	SMR[b]	95% CI
All Causes (1-998)	425	666.56	64	58—70
All Cancers (140-209)	101	132.24	77	63—94
Lung Cancer (161, 162)	30	44.28	68	46—97
Brain Cancer (191, 192)	8	5.26	152	65—299
Leukemia, Aleukemia (204-207)	6	5.54	108	40—236

[a]According to 8th Revision of the International Classification of Disease Codes.

[b]Those LTF dropped from the analysis at their employment termination date.

TABLE III. Selected Characteristics of LTF Employees, 1951-1978

LTF	568
Mean Age at Entry	29
Mean Year of Entry	1963
Mean Age at Employment Termination	35
Total Person-Years	9 429
Person-Years (5 years after initial employment)	6 599
Person-Years (10 years after initial employment)	3 769

TABLE IV. Comparative Mortality for Selected Causes, 1951-1978, Assuming Various Percent LTF Dead

Cause of Death (ICD)	% LTF[a]	Observed	Expected	SMR	95% CI
All Causes (1-998)	0[b]	425	692.94	61	56—67
	0[c]	425	666.56	64	58—70
	5	449	667.00	67	61—74
	10	483	667.69	72	66—79
	25	549	669.06	82	75—89
	50	695	671.44	104	96—111
	75	838	673.97	124	116—133
	100	987	677.28	146	137—155
All Cancers (140-209)	0[b]	101	136.05	74	60—90
	0[c]	101	131.24	77	63—94
	5	107	131.31	81	67—98
	10	116	131.42	88	73—106
	25	133	131.65	101	85—120
	50	169	132.03	128	109—149
	75	204	132.42	154	134—177
	100	234	133.00	176	154—200
Lung Cancer (161,162)	0[b]	30	45.90	65	44—93
	0[c]	30	44.28	68	46—97
	5	32	44.30	72	49—102
	10	35	44.33	79	55—110
	25	38	44.40	86	61—117
	50	43	44.51	97	70—130
	75	56	44.63	125	95—163
	100	61	44.81	136	104—175
Brain Cancer (191,192)	0[b]	8	5.53	145	62—285
	0[c]	8	5.26	152	65—299
	5	8	5.27	152	65—299
	10	8	5.28	152	65—299
	25	11	5.29	208	104—372
	50	15	5.31	282	158—466
	75	17	5.34	318	185—510
	100	20	5.37	372	227—575
Leukemia, Aleukemia (204-207)	0[b]	6	5.78	104	38—226
	0[c]	6	5.54	108	40—236
	5	6	5.54	108	40—236
	10	7	5.55	126	51—260
	25	7	5.56	126	50—259
	50	10	5.59	179	86—329
	75	12	5.61	214	110—374
	100	18	5.64	319	189—504

[a]Percent of those LTF simulated to have died.

[b]Those LTF considered alive at the end-of-study date.

[c]Those LTF dropped from the analysis at their employment termination date.

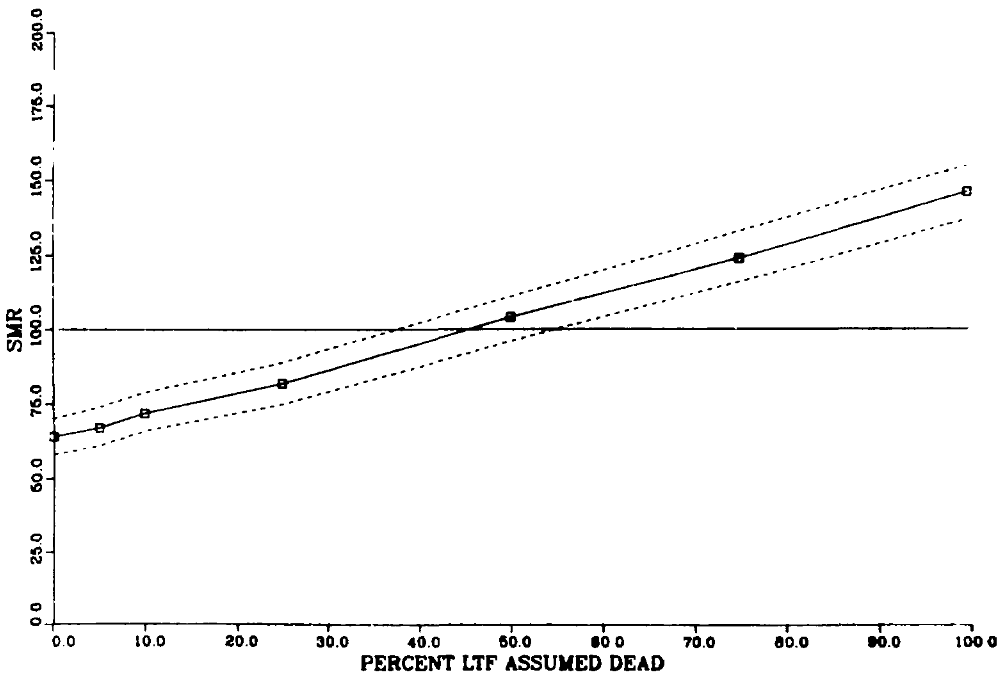

Fig. 1. SMRs and 95% CIs for all causes of death.

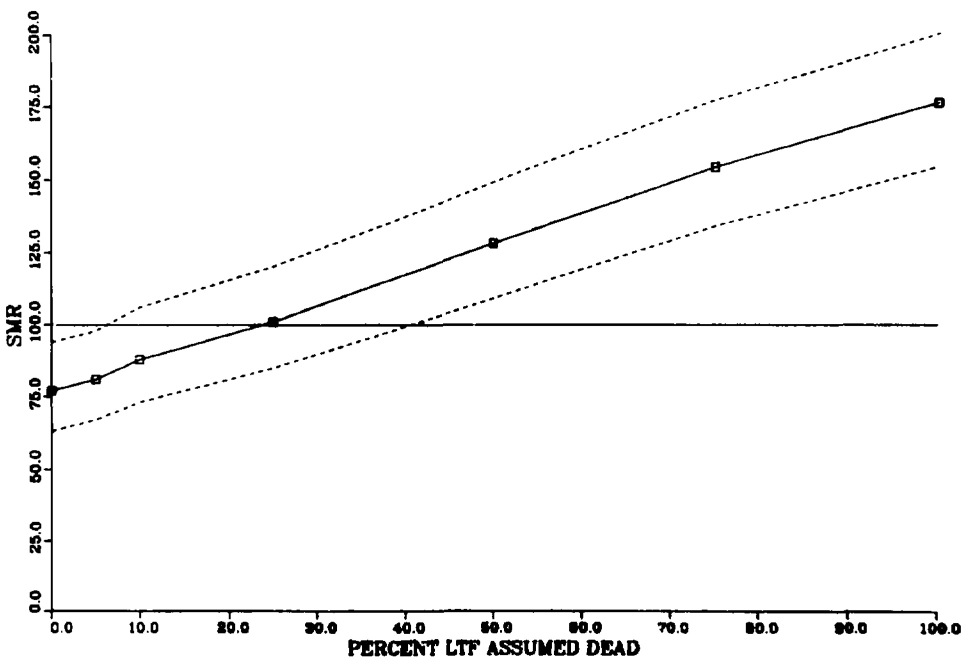

Fig. 2. SMRs and 95% CIs for all cancers.

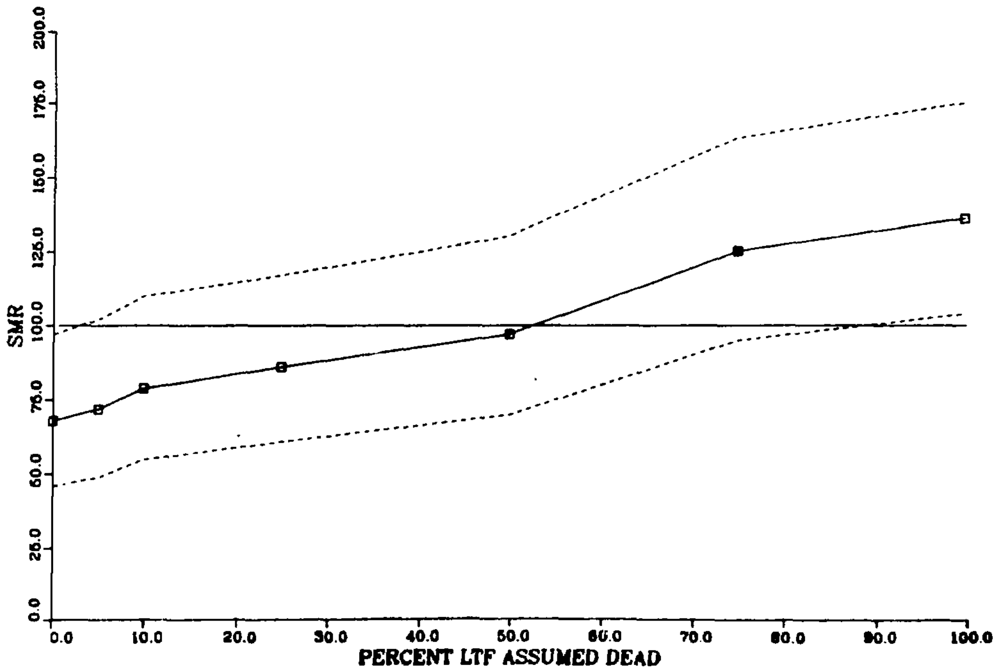

Fig. 3. SMRs and 95% CIs for lung cancer.

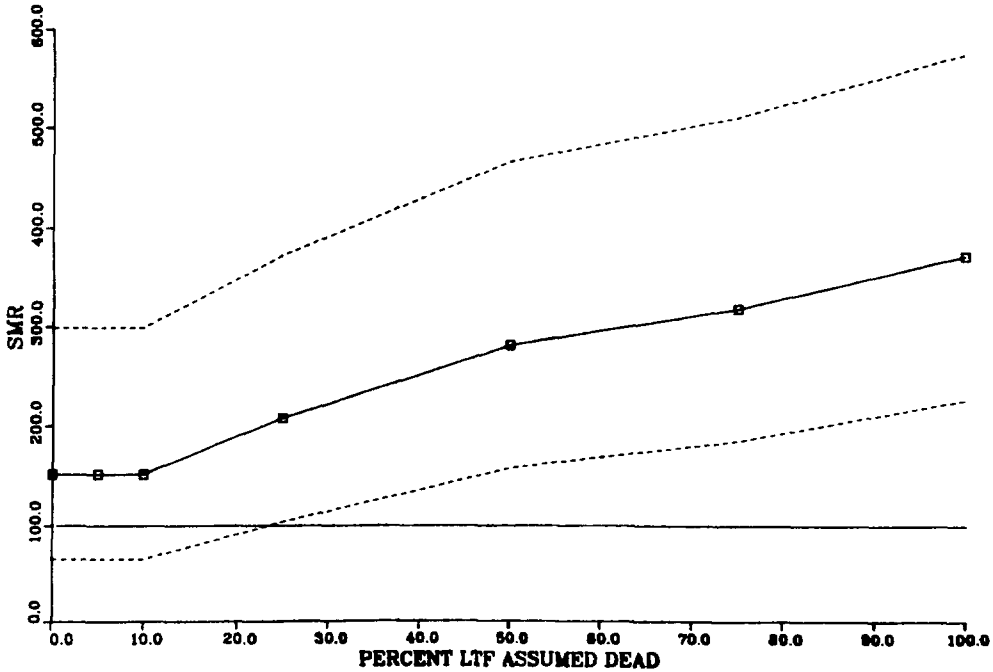

Fig. 4. SMRs and 95% CIs for brain cancer.

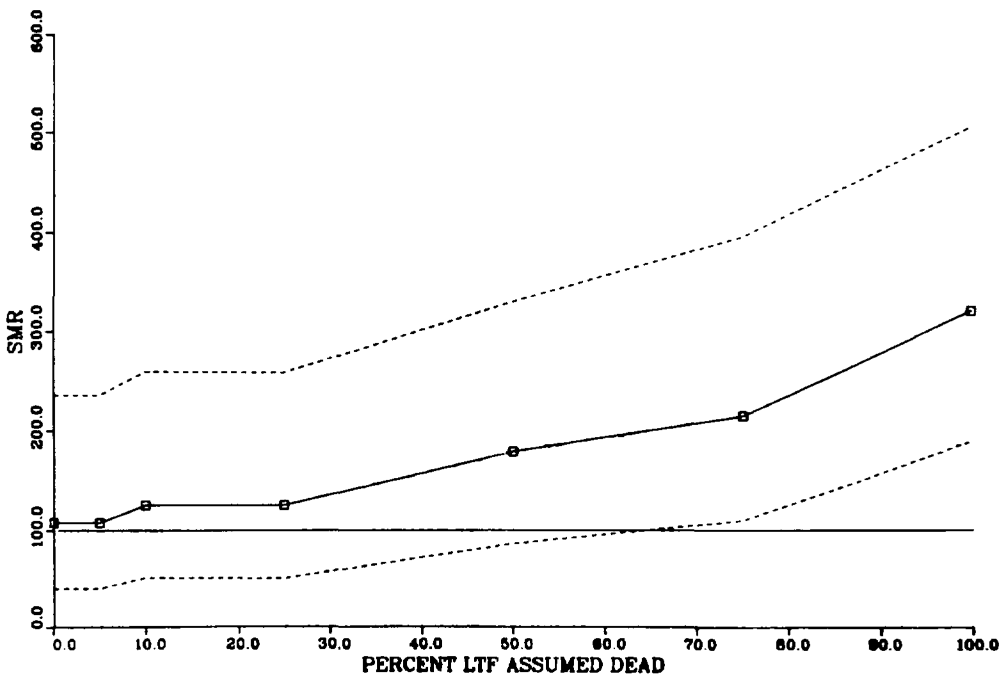

Fig. 5. SMRs and 95% CIs for leukemia, aleukemia.

TABLE V. Comparative Mortality for Selected Causes, 1951-1978, Assuming Various Percent LTF Dead; 5 Years After Initial Employment

Cause of Death (ICD)	% LTF[a]	Observed	Expected	SMR	95% CI
All Causes (1-998)	0[b]	367	569.92	64	58—71
	0[c]	367	546.86	67	60—74
	5	384	547.20	70	63—78
	10	413	547.77	75	68—83
	25	459	548.64	84	76—92
	50	570	550.38	104	95—112
	75	672	552.34	122	113—131
	100	778	555.02	142	132—152
All Cancers (140-209)	0[b]	88	116.75	75	60—93
	0[c]	88	112.39	78	63—96
	5	93	112.44	83	67—101
	10	101	112.55	90	73—109
	25	115	112.69	102	84—122
	50	142	112.99	126	106—148
	75	166	113.32	146	125—171
	100	189	113.80	166	143—192
Lung Cancer (161,162)	0[b]	25	40.64	62	40—91
	0[c]	25	39.13	64	41—94
	5	27	39.15	69	45—100
	10	29	39.18	74	50—106
	25	32	39.23	82	56—115
	50	36	39.32	92	64—127
	75	46	39.42	117	85—156
	100	50	39.58	126	94—167
Brain Cancer (191,192)	0[b]	7	4.37	160	64—330
	0[c]	7	4.14	169	68—348
	5	7	4.14	169	68—348
	10	7	4.15	169	68—348
	25	10	4.16	241	115—443
	50	13	4.17	311	166—533
	75	15	4.19	358	200—590
	100	16	4.22	379	216—616
Leukemia, Aleukemia (204-207)	0[b]	6	4.58	131	48—285
	0[c]	6	4.38	137	50—298
	5	6	4.38	137	50—298
	10	7	4.39	159	64—329
	25	7	4.40	159	64—328
	50	9	4.41	204	93—387
	75	10	4.43	226	108—415
	100	15	4.46	336	188—555

[a]Percent of those LTF simulated to have died.

[b]Those LTF considered alive at the end-of-study date.

[c]Those LTF dropped from the analysis at their employment termination date.

TABLE VI. Comparative Mortality for Selected Causes, 1951-1978, Assuming Various Percent LTF Dead; 10 Years After Initial Employment

Cause of Death (ICD)	% LTF[a]	Observed	Expected	SMR	95% CI
All Causes (1-998)	0[b]	291	416.62	70	62—78
	0[c]	291	399.91	73	65—82
	5	305	400.14	76	68—85
	10	319	400.47	80	71—89
	25	347	400.94	87	78—96
	50	405	401.81	101	91—111
	75	464	402.97	115	105—126
	100	525	404.70	130	119—141
All Cancers (140-209)	0[b]	72	88.72	81	63—102
	0[c]	72	85.40	84	66—106
	5	77	85.44	90	71—113
	10	80	85.50	94	74—116
	25	89	85.59	104	84—128
	50	105	85.75	122	100—148
	75	115	85.96	134	110—161
	100	127	86.30	147	123—175
Lung Cancer (161,162)	0[b]	19	31.72	60	36—94
	0[c]	19	30.52	62	37—97
	5	21	30.53	69	43—105
	10	22	30.55	72	45—109
	25	24	30.58	78	50—117
	50	28	30.63	91	61—132
	75	30	30.70	98	66—139
	100	32	30.82	104	71—147
Brain Cancer (191,192)	0[b]	5	3.00	167	54—389
	0[c]	5	2.83	177	57—412
	5	5	2.84	176	57—412
	10	5	2.84	176	57—412
	25	6	2.84	211	77—459
	50	8	2.85	280	121—553
	75	10	2.86	349	167—642
	100	11	2.88	382	190—683
Leukemia, Aleukemia (204-207)	0[b]	5	3.25	154	50—359
	0[c]	5	3.11	161	52—375
	5	5	3.11	161	52—375
	10	6	3.11	193	70—419
	25	6	3.12	192	70—419
	50	7	3.13	224	90—461
	75	8	3.14	255	110—503
	100	12	3.15	381	197—665

[a]Percent of those LTF simulated to have died.

[b]Those LTF considered alive at the end-of-study date.

[c]Those LTF dropped from the analysis at their employment termination date.