Association of Carcinoembryonic Antigen with Mortality in an Insurance Applicant Population

Steven J. Rigatti, MD, DBIM, DABFM; Robert Stout, PhD

Objectives.—To quantify the mortality risks associated with elevated levels of carcinoembryonic antigen (CEA).

Background.—Carcinoembryonic antigen is cell surface glycoprotein and has been associated with the presence of high grade or metastatic cancers of the colon as well as other malignant and non-malignant disease. Prior publications have demonstrated the utility of CEA levels in the determination of mortality risk in life insurance applicants. The aim of this paper is to further characterize this risk with a larger set of data containing additional person-years of follow-up, more outcomes, and additional variables potentially associated with occult malignancy.

Methods.—By use of the Social Security Death Index, mortality was examined in 321,574 insurance applicants age 50 years and older, who submitted blood samples to Clinical Reference Laboratories for testing including CEA. Results were stratified by age group and by CEA level (<5 ng/mL, 5 to 9.9 ng/mL, 10+ ng/mL), though other thresholds were tested. Mortality comparisons were carried out using Cox models and tabular methods with the 2015 smoker-distinct Valuation Basic Tables as a comparator.

Results.—Relative mortality is increased at CEA levels above 4.0 ng/mL in both smokers and non-smokers. This association is persistent in Cox models when albumin, BMI and cholesterol are included as covariates. The strongest association with mortality risk occurred in the first 3-4 durations. The 3-year cumulative mortality ratio when using the 2015 VBT as baseline was 6.51 when comparing the group with CEA levels of 10+ ng/mL, compared to those with levels below 5.0 ng/mL.

Conclusion.—This study shows that CEA is strongly associated with the risk of early excess mortality in life insurance applicants, and this risk appears not to be mitigated by consideration of other markers thought to be associated with occult malignancy.

Address: Rigatti Risk Analytics, LLC, 157 Uconn Ave, Glastonbury, CT 06033, steven.rigatti@crlcorp.com

Correspondent: Steven J. Rigatti, MD.

Author Affiliations: Rigatti – Founder, Rigatti Risk Analytics, LLC, Consultant Medical Director, Clinical Reference Laboratories, Lenexa KS. Stout – Chief Scientific Officer/Laboratory Director, Clinical Reference Laboratories, Lenexa KS.

INTRODUCTION

Carcinoembryonic antigen belongs to a class of cell surface glycoproteins and is expressed at fairly high levels during the embryogenesis of the human GI tract. It was first discovered in 1965[1] in both fetal and cancerous colon tissue and was thought to be absent from adult non-malignant tissues. This gave rise to the name "carcino-embryonic" antigen (CEA). Later it was discovered that CEA is produced in normal adult cells,

generally confined to the glycocalyx of the mucosal epithelial cells of the colon and rectum.[2] In these normal cells, CEA is produced at much lower concentrations than in tumor cells. The CEA molecule is structurally similar to an immunoglobulin. The genes encoding CEA are part of the immunoglobulin complex of genes located on chromosome 19q. The native role of CEA in human physiology remains largely mysterious. In vitro experiments suggest that CEA is somehow involved in cellular adhesion,[3] or perhaps that it binds bacteria in the colonic lumen in order to prevent invasion and infection.[4] Because derangement of cell adhesion is necessary for metastasis, it has been proposed that CEA is actually a causal agent in cancer spread. Though there has been some evidence of this in animal models,[5] it is not yet conclusive.

Circulating levels of CEA have been used clinically to assess the prognosis and treatment effects in various malignancies. In particular, it is used as both a prognostic marker and as a means of monitoring after treatment for colorectal cancer.[6] Rising levels immediately after treatment or after successful induction of remission may lead to further imaging or diagnostic studies to rule out recurrence or metastases. Current guidelines suggest testing CEA every 3 to 6 months after surgery for colon cancer performed with curative intent.[7] It is not generally used as a screening test in the general population due to poor overall predictive value.[8] Interestingly, colonic neoplasms, which produce CEA, are generally low grade.[9] High grade colonic neoplasms tend to lose the ability to produce CEA in much the same way that high grade neoplasms in the breast lose the ability to produce hormone receptors.

Elevated levels of CEA can also be associated with other diseases. A recent study demonstrated its predictive utility for patient with pulmonary fibrosis awaiting transplant.[10] Another demonstrated an ability of elevated CEA levels (above 3.2 ng/mL) to predict shorter disease-free survival in early stage, operable breast cancer.[11] Historically, CEA was thought to be a marker of inflammatory bowel disease, and though levels may be elevated more commonly than in individuals without IBD, CEA is not currently regarded as a useful marker for disease severity, monitoring or prognosis in this context.[12] Other diseases which may cause an increase in CEA levels include pancreatitis, cirrhosis, COPD, and carcinomas of the stomach, pancreas, lung, liver and thyroid gland.[13,14]

The utility of CEA measurements in insurance testing has been evaluated in a prior publication in the *Journal of Insurance Medicine*.[15] Additional information was published subsequently in the industry journal, *On the Risk*.[16]

The purpose of this article is to further update the mortality analysis and to refine the previously published studies with additional data and to evaluate the possible contribution of other tests that may be associated with occult, advanced malignancy, such as build, cholesterol and albumin.

METHODS

Data were obtained from insurance applicants undergoing testing at CRL between January 1, 2002, and December 31, 2013. The study end date was December 31, 2015. Deaths were assessed by reference to the Social Security Death Master File (SSDMF). Deaths occurring after the study end date were censored. Duration was determined as the time from the test date until the death date or study end date, whichever occurred first.

The study included those who had a CEA measurement at the time of their insurance testing, whether it was ordered by the testing company or performed as part of the original pilot study. All study subjects were at least 50, and no more than 90 years at the time of testing. A history of smoking was determined by either an admission of cigarette smoking on the laboratory consent form, or a urine cotinine level of >200 ng/dL. Note that this is slightly different from the previously

published studies, which used only the cotinine level as the indicator of smoking.

The association of CEA levels with mortality was investigated several different ways. To replicate the previously published data, tabular methods were utilized with the same age groupings (50-59 yrs, 60-69 yrs, 70 or more years), and the same CEA groupings (0-4.9 ng/mL, 5-9.9 ng/mL, and 10 ng/mL or higher) as used in prior studies. Additionally, Cox models were constructed using age, sex, smoking and, variously, body mass index (BMI), serum albumin level and serum total cholesterol level as covariates. These values were obtained on the same blood sample as CEA or, in the case of BMI, by a paramedical examiner at the time of testing. Because height and weight information were missing in approximately 25% of applicants, models involving BMI are reported separately.

In the prior publications, the tabular method was primarily utilized to calculate mortality ratios and excess death rates for a group with elevated levels of CEA when compared to those in the same age/smoker group with normal (0-4.9 ng/mL) levels of CEA. This method does not control for the fact that the CEA-based groups may not have identical age/sex distributions, and therefore, some of the mortality effect ascribed to CEA may be due to those differences. There are several ways to address this concern. One is to use Cox models which include age, sex and smoking status as covariates to control for those factors. This approach has the added advantage of allowing CEA to be treated as a continuous value, which enables evaluation for non-linear effects.[17] Another method is to compare the survival experience of each group to an external reference based on age, sex and smoker status if possible – resulting in a Standardized Mortality Ratio. This approach is also worthwhile but is hampered by the known incompleteness of the SSDMF with regard to mortality assessment.[18] This weakness can be overcome by calculating a ratio of SMRs. For instance, suppose a group of males in their 60s with CEA levels between 5.0 and 9.9 ng/dl had a 5-year SMR of 1.2 when compared to the US general population. This number is likely to be erroneously low due to incomplete death assessment in the study population. Suppose further that the study group of men with CEA levels below 5.0 ng/dL had a 5-year SMR of 0.4 when compared to the reference population – also biased downward by incomplete death assessment. A ratio of the 2 SMRs can be calculated ($1.2/0.4 = 3.0$), effectively removing the effect of incomplete assessment – assuming that the effect is nearly the same in each group. In this article, both the Cox model approach and the SMR approach outlined above are utilized; the SMR approach is termed the SMR ratio or "SMRR."

Note that in the Cox models, restricted cubic splines are used to account for the known J-shaped mortality curves associated with BMI and cholesterol. As has been noted in prior publications,[19] one should be cautious when interpreting the results of such models especially with regard to effects at the margins of the data – that is with extreme or unusual values of inputs where the data become sparse.

When comparing to an external population, both US life tables[20] and 2015 Valuation Basic Tables[21] (VBT) are used. Because the subjects of this study are insurance applicants, it should be considered that their survival would be expected to be somewhat better than the general population, but also somewhat worse than fully underwritten insured individuals. While this is true, these comparisons will only be used to calculate ratios of SMRs, effectively removing these biases. It should also be noted that the US life tables are not smoker-distinct. When using the US Life Tables, rates are matched by age, sex and calendar year. For VBT comparisons, rates are matched by age, sex, smoking status, and duration.

Data gathering and analyses were performed with SPSS[22] version 24.0 and R version 3.4.3[23] using the following packages

Table 1. Characteristics of the Study Population by Age and CEA Group

Age Group (years)	CEA (ng/mL)	Population At Risk	Mean Age (years)	Male (%)	Smokers[1] (%)	Proportion of Smokers among Deaths (%)	Mean Albumin (g/dl)	Mean BMI[2] (kg/m2)	Mean Cholesterol (mg/dL)
50 to 69	0-4.9	183,324	54.7	67.5	13.9	26.9	4.48	27.6	203
	5-9.9	3,092	54.8	64.7	52.3	62.8	4.43	26.8	210
	10.0+	399	55.1	59.6	61.4	52.5	4.38	26.4	210
	Total	186,815	54.7	67.5	14.6	27.6	4.48	27.6	203
60 to 69	0-4.9	99,794	64.1	71.3	10.7	19.1	4.43	27.8	196
	5-9.9	1,899	64.3	66.4	39.8	52.0	4.37	27.2	203
	10.0+	256	64.3	60.2	43.4	45.5	4.28	26.3	200
	Total	101,949	64.1	71.2	11.3	19.8	4.43	27.8	196
70+	0-4.9	31,660	74.1	62.1	7.5	11.2	4.35	26.9	191
	5-9.9	997	74.9	53.8	23.2	29.1	4.30	25.8	200
	10.0+	153	74.9	48.4	19.0	20.5	4.27	25.9	196
	Total	32,810	74.112	61.8	8.0	11.8	4.35	26.9	191
All (50+)	0-4.9	314,778	59.7	68.2	12.2	18.8	4.45	27.6	200
	5-9.9	5,988	61.2	63.4	43.5	48.1	4.39	26.8	206
	10.0+	808	61.8	57.7	47.6	39.1	4.32	26.3	204
	Total	321,574	59.7	68.1	12.9	19.4	4.4	27.6	200

[1]Urine cotinine ≥ 200 ng/mL or admitted smoking at time of test
[2]BMI values could not be calculated for 25.2% of the study group

and their pre-requisites: rms,[24] tidyverse,[25] popEpi.[26]

RESULTS

The data contained 321,574 individuals, including 6084 who died during the study period. Total follow-up time was over 2.1 million person-years and averaged 6.55 years per subject. Table 1 gives the average age, albumin level, BMI and cholesterol level, as well as the proportion of males and proportion of smokers for each age/CEA group. Additionally, the proportion of smokers among decedents in each group is reported. This demonstrates an overall increased proportion of smokers among those with higher CEA levels, with further increases when only the deceased are considered. Table 1 also demonstrates a trend toward lower albumin and BMI in those with higher levels of CEA.

Since approximately 2009, CRL has collected information about applicants' history of medical illness including cancer. The data is collected as part of the routine paramedical examination process. The cancer question is a simple yes/no checkbox indicating the presence or absence of a history of cancer at any time in the past. The median CEA level was the same (1.6 ng/mL) in both those with and without an admitted history of cancer. There were some instances where the history of cancer seemed relevant. For instance, there were 3 cases after 2009 in which the CEA was over 1000 ng/mL, and there was an admitted history of cancer in 2 of them.

Standardized Mortality Ratio Analysis

The 2009 article presented a table stratified by CEA level with calculations of mortality rate (q) and a mortality ratio for those with CEA levels of 5.0 ng/mL and above. Using that same method, the 2-, 5- and 10-year cumulative mortality ratios for CEA values between 5 and 9.9 ng/mL are 545%, 429% and 330%, respectively. While for CEA values 10 ng/mL or higher, the corresponding ratios

Table 2. Standardized Mortality Ratios by Duration and CEA Group

Interval Years i	Observed Deaths d	Expected Deaths USLT[1] d'_{US}	Expected Deaths VBT[2] d'_{VBT}	Exposure Person-yrs E	SMR USLT d/d'_{US}	SMR VBT d/d'_{VBT}	Interval$_{US}$	Interval$_{VBT}$	Cumulative$_{US}$	Cumulative$_{VBT}$
CEA 0 - 4.9 mg/dl										
0 to 1	439	3686.3	403.5	314204.2	0.12*	1.09*				
1 to 2	677	3946.2	662.5	313622.8	0.17*	1.02*				
2 to 3	678	3937.1	901.9	292152.1	0.17*	0.75*				
3 to 4	666	3579.0	971.5	248628.2	0.19*	0.69*				
4 to 5	624	3284.5	991.8	213558.0	0.19*	0.63*				
5 to 6	531	3058.7	1001.5	185620.7	0.17*	0.53*				
6 to 7	507	2693.8	968.4	151832.3	0.19*	0.52*				
7 to 8	338	1930.7	773.6	103056.5	0.18*	0.44*				
8 to 9	307	1548.7	670.7	81428.2	0.20*	0.46*				
9 to 10	248	1253.4	586.1	61958.5	0.20*	0.42*				
10 to 11	310	1822.4	936.3	81795.9	0.17*	0.33*			**SMR Ratio[3]**	
CEA 5.0 - 9.9 mg/dl										
0 to 1	54	88.7	12.2	6302.0	0.61	4.41*	5.11	4.06	5.11	4.06
1 to 2	68	93.1	20.5	6235.9	0.73	3.32*	4.26	3.25	4.59	3.56
2 to 3	66	95.0	30.5	6011.9	0.69	2.16*	4.03	2.88	4.38	3.26
3 to 4	51	95.5	37.6	5672.7	0.53	1.36*	2.87	1.98	3.95	2.83
4 to 5	58	96.2	42.1	5349.3	0.60	1.38*	3.17	2.19	3.79	2.65
5 to 6	58	97.6	46.5	5051.5	0.59	1.25	3.42	2.35	3.73	2.56
6 to 7	40	97.1	50.2	4722.8	0.41	0.80	2.19	1.52	3.49	2.36
7 to 8	39	80.6	45.7	3757.9	0.48	0.85	2.76	1.95	3.42	2.28
8 to 9	22	66.0	40.4	3027.4	0.33	0.54*	1.68	1.19	3.27	2.16
9 to 10	26	53.5	34.6	2290.1	0.49	0.75	2.46	1.77	3.22	2.12
10 to 11	25	53.7	37.9	2120.7	0.47	0.66	2.74	1.99	3.19	2.12

Table 2. Continued

Interval Years i	Observed Deaths d	Expected Deaths USLT[1] d'_{US}	Expected Deaths VBT[2] d'_{VBT}	Exposure Person-yrs E	SMR USLT d/d'_{US}	SMR VBT d/d'_{VBT}				
CEA 10.0 mg/dl or higher										
0 to 1	27	11.5	1.6	808.5	2.34*	16.90*	19.64	15.53	19.64	15.53
1 to 2	34	11.8	2.6	778.7	2.89*	13.05*	16.85	12.77	17.90	13.86
2 to 3	24	11.7	3.8	728.2	2.05*	6.24*	11.91	8.30	15.66	11.59
3 to 4	9	11.2	4.7	667.1	0.80	1.93*	4.31	2.82	12.52	8.84
4 to 5	8	11.1	5.1	616.0	0.72	1.55	3.81	2.47	10.64	7.28
5 to 6	7	11.0	5.6	581.9	0.64	1.25	3.67	2.35	9.49	6.34
6 to 7	6	10.6	5.9	537.8	0.57	1.02	3.02	1.96	8.56	5.61
7 to 8	8	8.9	5.4	440.4	0.90	1.49	5.13	3.42	8.21	5.31
8 to 9	4	7.1	4.6	344.2	0.56	0.87	2.83	1.89	7.77	4.98
9 to 10	1	5.7	3.9	256.9	0.18	0.26	0.89	0.61	7.34	4.68
10 to 11	3	5.7	4.2	231.5	0.52	0.71	3.08	2.13	7.11	4.60

[1] Expected based on US Life Tables matched by age, sex and calendar year

[2] Expected based on 2015 VBT smoker distinct tables matched by age, sex, smoking status and duration

[3] All Ratios are vs. the SMR for the CEA 0-4.9 ng.mL group

* p-value < 0.01 for null hypothesis that SMR = 1

are 2077%, 1109% and 605%. These ratios are all compared to the group with CEA values below 5 ng/mL. A slightly different approach is utilized in this article. The change is primarily because the prior method does not account for possible differences in the age and sex distributions between the CEA-based groups. It is noted, however, that these differences are rather small as indicated in Table 1.

Table 2 displays the results of calculations of the Standardized Mortality Ratio using the US population life table and the 2015 VBT as a reference. This demonstrates that the group with low CEA levels has rates of mortality that are considerably lower than the general population. Again, this is likely due to two principal factors: the incomplete mortality assessment from the Social Security Death Master File, and the selection effect in this population which has elected to undergo testing for life insurance purposes. In the comparison to the VBT, the SMRs are much closer to 1 initially, but then decline as the expected rates increase and the observed rates decline. For those groups with higher levels of CEA, a ratio of the SMRs is calculated on both an interval and cumulative basis. At 10 years, these cumulative ratios show that those with CEA values between 5.0 and 9.9 ng/mL have an SMR approximately 3 times higher than the comparison group, while those with CEA values 10 ng/mL or more have approximately 7 times the SMR of the comparison group when the USLT is used for comparison. Note that statistical testing of the null hypothesis that the SMR is not different than 1 is not significant at the p<0.01 level when there are few observed deaths or when the SMR is very close to 1. This method demonstrates very high early interval mortality ratios for both groups with elevated CEA levels, which is consistent with prior publications. It is also noted that the interval and cumulative SMRRs are systematically lower when the VBT is used as the baseline. This is certainly due to the inclusion of smoker distinction in the VBT-based comparisons combined with the high proportion of smokers in the higher ranges of CEA. This

more properly accounts for the independent effect of smoking on mortality apart from the effect on CEA values.

Cox Model Analysis

Cox proportional hazards methods were utilized to generate hazard ratios for various levels of CEA. Because binning of continuous variables using arbitrary cut-offs may cause significant bias, models were fit using CEA in its native state as a continuous variable and with the previously utilized binning thresholds. When used as a continuous variable, CEA values above 15 ng/mL were imputed at 15 ng/mL to avoid issues with outliers and heavily right-skewed variables. There were 303 such values in the data.

Models with continuous CEA demonstrated a hazard ratio of 1.19 per unit increase, when controlled for age, sex and smoking status and 1.17 when additionally controlled for cholesterol and albumin. When treated as a categorical variable in a model with age, sex, smoking status, cholesterol and albumin, the hazard ratio was 2.36 for CEA values of 5.0-9.9 ng/mL and 4.08 for values of 10 ng/mL or higher. Examination of Schoenfeld residuals demonstrated that these models violate the proportional hazards assumption (PHA) for CEA. That is, there is evidence of time-dependency. This is consistent with the tabular data which demonstrated much higher mortality ratios in the early durations after testing. Because of this issue, Models 3 and 4 were fit, replicating Models 1 and 2 but limiting the data to 3 years of follow-up. Note that nearly the entire data set has at least 3 years of follow up because the latest allowed entry date and the study end date are 2 years apart. This modification resulted in models that did not violate the PHA but which had many fewer deaths and showed a greater effect of CEA on mortality. Model 5 included a 4-knot restricted cubic spline term for BMI. Since BMI data was missing in approximately one quarter of study subjects, this further limited the total number of deaths. BMI was
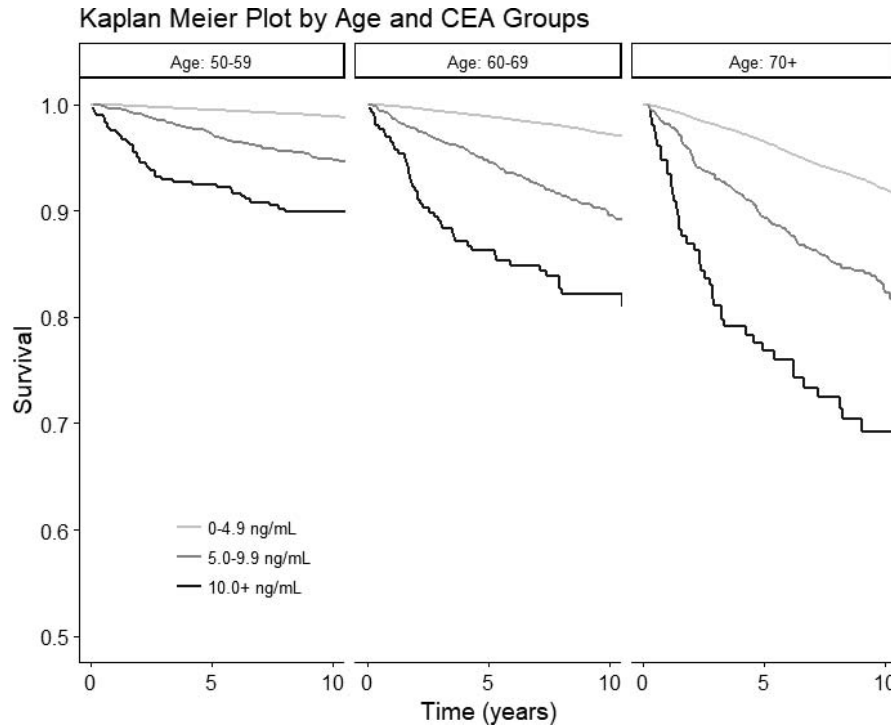
Figure 1. *Survival time in years by CEA group and age group.*

a significant predictor of mortality though its modification effect on CEA was modest, suggesting that BMI and CEA were mostly independent predictors of mortality.

For Models 6, 7 and 8, the data were split by age category but otherwise included the same variable structure as Model 4. BMI was not included since it would have further reduced the age-restricted data sets, and because it was not especially influential on the hazard ratios for CEA. These age-restricted models demonstrated fairly consistent hazard ratios of approximately 8 for the highest category of CEA levels and 3-4 for the middle category. Analysis of variance (ANOVA) was carried out on Models 4 and 5 (Table 4), and this demonstrates that the included variables were significantly associated with improved model fits. Figure 1 displays a Kaplan-Meier plot of survival split by CEA category and by age group.

Threshold Testing

In order to evaluate the efficacy of various cutoffs of CEA in the selection of life risks, the SMR method was utilized to test thresholds

between 4 and 9 ng/ml. In these comparisons, values of CEA below the test threshold were considered the "low" (reference) group, those with CEA values between the threshold and 10 ng/ml were considered the "mid" group, and those with values 10 ng/mL or higher were considered the "high" group. When considering a strategy for underwriting a laboratory value such as CEA, it is useful to consider the proportion of applicants who will fit into the constructed categories and the mortality implications of each group. Hence for each tested threshold, Table 5 reports the proportion of applicants fitting into the high, mid and low groups as well as the 5-year cumulative SMRR for the mid and high groups vs the low group. For this analysis, smokers and non-smokers are considered separately, and the USLT is the baseline comparator.

This analysis demonstrates that for the category with CEA values of 10 ng/mL or higher, the SMRR for non-smokers is approximately 11 vs the low group with CEA values below the tested threshold. For smokers, the comparable value is approximately 5 to 6. For the mid group vs the low group (see
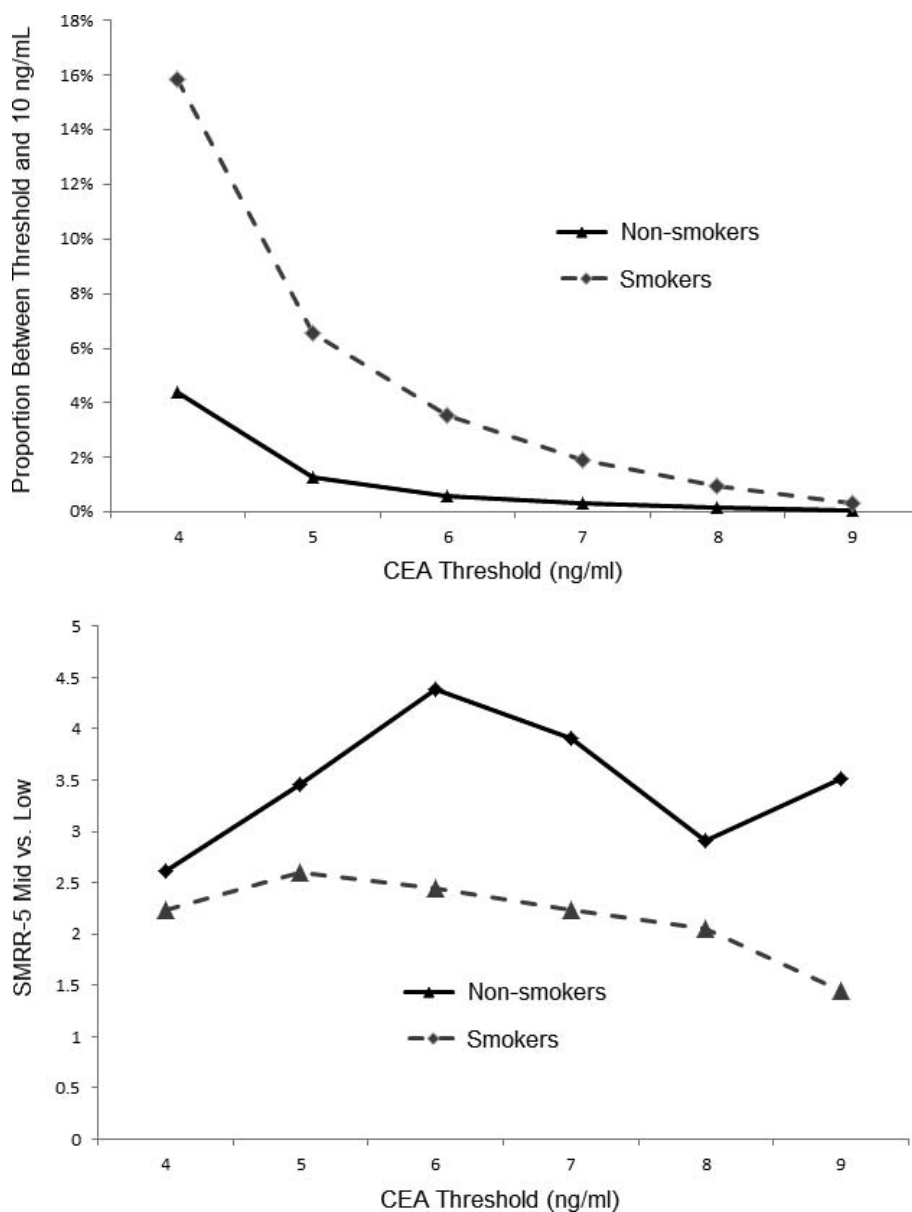
Figure 2. *Threshold Effects of CEA. Upper panel: For each whole-number threshold between 4 and 9 ng/mL, points represent the proportion of study subjects with CEA values between that threshold and 10 ng/mL. Lower panel: For each whole-number threshold between 4 and 9 ng/mL, points represent SMRR-5 for the group with values between that threshold and 10 ng/mL vs the group with values below that threshold. The SMRR-5 displayed is calculated based on the US-population comparison group. Smoker (dashed line) and non-smokers (solid line) are displayed separately.*

Figure 2), the SMRR varies more strongly with the threshold. For non-smokers, the ratio peaks at a threshold of 6 ng/mL, while for smokers it peaks at 5 ng/mL but remains rather steady up to a threshold of 8 ng/mL.

DISCUSSION

This study of the mortality implications of CEA in a population applying for life insur-ance is an expansion of prior studies with considerably longer follow up, and a greater number of study subjects and outcomes. Accordingly, the present study permits a more reliable estimate of the mortality effects of elevated CEA.

Further, this study utilizes additional methods by which the mortality effects may be quantified. These included the ratio of SMRs method in Table 2, and the Cox proportional

Table 3. Cox Model Results

Descriptives		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Included Ages:		50-84	50-84	50-84	50-84	50-84	50-59	60-69	70-84
Duration, Mean(Max), yrs.:		7.9(15.0)	7.9(15.0)	3.0(3.0)	3.0(3.0)	3.0(3.0)	3.0(3.0)	3.0(3.0)	3.0(3.0)
Number of Deaths:		6084	6084	2076	2076	1323	622	735	719
PHA Violation, CEA[1]:		Yes	Yes	No	No	No	No	No	No
Variables									
Age:	(years)	1.1*	1.1*	1.09*	1.09*	1.09*	1.08*	1.11*	1.09*
Sex:	Male vs. Female	1.37*	1.39*	1.52*	1.55*	1.73*	1.97*	1.70*	1.26*
Smoker:	Yes vs. No	1.88*	2.03*	1.77*	1.92*	1.92*	1.93*	1.96*	1.77*
CEA[2]	(ng/mL)	1.17*		1.24*					
CEA	0-4.9 ng/mL		1 (ref.)		1(ref.)				
	5.0-9.9 ng/mL		2.36*		3.24*	3.26*	3.15*	3.78*	2.87*
	10.0+ ng/mL		4.08*		8.70*	7.84*	9.97*	8.26*	8.09*
Cholesterol[3]	main	0.997	0.997	0.997	0.997	0.996	0.997	0.996	0.998
	first spline	1.002	1.002	1.00	1.00	0.997	0.998	1.01	1.00
	second spline	1.01	1.01	1.02	1.02	1.007	1.03	0.99	1.01
Albumin[3]	main	0.2*	0.018*	0.145*	0.14*	0.14*	0.155*	0.13*	0.13*
	first spline	4.02*	4.511*	2.41	2.51	2.72	0.989	3.07	3.48
	second spline	0.01	0.009	5.72	5.47	3.76	954.6	3.59	0.48
BMI[3]	main					0.87*			
	first spline					1.64*			
	second spline					0.28*			

[1] Test for violation of the proportional hazards assumption of the CEA variable at 99% confidence level.

[2] CEA is treated as a continuous variable in models 1 and 3 and as categorical in all others. CEA values above 15 are imputed at 15 ng/m

[3] Cholesterol, BMI and Albumin are all included as restricted cubic splines with 4 knots

* Hazard ratio significant at 99% confidence level

hazards modeling approach in Table 3. It is heartening that all these methods seem to converge on a similar answer to the overall risk ratios experienced by those with modest (5.0 – 9.9 ng/mL) and severe (10.0+ ng/mL) elevations in CEA level. All of these methods have different strengths and weaknesses – the tabular methods are not able to control for other factors such as smoking, or levels of cholesterol, albumin or BMI, while the Cox models, which can control for these factors, demonstrate violation of the proportional hazards assumption. When the Cox model is limited to consider only early durations, the mortality ratios for CEA increase. Additional methodological adjustments could be made, such as using Cox models with time-varying covariates or using a parametric survival model, such as a Poisson model, to examine how death rates vary with duration and other included covariates.

Because CEA is generally regarded as a tumor marker, it is likely that the excess deaths seen in those with high levels are due to occult malignancy. It should be noted that this study does not contain cause of death information, and therefore, cannot make this assertion. It is apparent, however, that the mortality implications of elevated CEA are strongest in the first few years after testing, which strongly implies a relationship to some factor which is associated with short-term mortality. It should also be noted that the hazard ratios for CEA in the Cox models fell only slightly or rose when additional covariates thought to be related to occult malignancy were included. This

Table 4. ANOVA results

	d.f.	Chi-Square[1] Model 4	Chi-Square[1] Model 5
Age	1	1161.18	665.48
Sex	1	72.25	65.55
Smoker	1	143.85	93.99
CEA	2	487.3	251.96
Cholesterol	3	26.87	22.44
-nonlinear	2	23.28	17.24
Albumin	3	855.74	527.64
-nonlinear	2	85.82	60.19
BMI	3		80.66
-nonlinear	2		74.01

[1] All Chi-square p-values $<10^{-4}$

suggests that the association between mortality risk and CEA level is unique amongst the variables considered in this study. This study did not have access to information that would usually be collected at the time of underwriting such as medical questionnaires, family history, prescription records, or medical files, so it is not possible to further isolate the mortality implications of CEA.

Table 5. CEA Thresholds vs. Distribution and Mortality

Threshold ng/mL	CEA Group Low	CEA Group Mid	CEA Group High	5-Year SMRR$_{USLT}$ Mid vs. Low	5-Year SMRR$_{USLT}$ High vs. Low
		Non-smokers			
4	95.4%	4.4%	0.2%	2.62	12.43
5	98.6%	1.3%	0.2%	3.46	11.91
6	99.2%	0.6%	0.2%	4.39	11.71
7	99.6%	0.3%	0.2%	3.90	11.51
8	99.7%	0.1%	0.2%	2.91	11.41
9	99.8%	0.1%	0.2%	3.51	11.38
		Smokers			
4	83.2%	15.9%	0.9%	2.23	5.99
5	92.5%	6.5%	0.9%	2.59	5.53
6	95.5%	3.5%	0.9%	2.44	5.21
7	97.1%	1.9%	0.9%	2.23	5.05
8	98.1%	0.9%	0.9%	2.05	4.97
9	98.7%	0.3%	0.9%	1.45	4.92

CONCLUSIONS

Elevated levels of CEA in a life insurance testing context are associated with significant mortality risk independent of other identifiable factors contained in the study data.

REFERENCES

1. Gold P, Freedman SO. Demonstration of tumor-specific antigens in human colonic carcinomata by immunological tolerance and absorption techniques. *J Exp Med*. 1965:121;439-462.
2. Macdonald JF. Carcinoembryonic Antigen Screening: pros and Cons. *Semin Oncol*. 1999:25;556-560.
3. Thompson JA, Grunert F, Zimmermann W. Carcinoembryonic antigen gene family: molecular biology and clinical perspectives. *J Clin Lab Anal*. 1991;5:344-366.
4. Hammarstrom S. The carcinoembryonic antigen (CEA) family: structure, suggested functions and expression in normal and malignant tissue. *Semin Cancer Biol*. 1999;9:67-81.
5. Hostetter RB, Augustus LB, Mankarious R, et al. Carcinoembryonic antigen as a selective enhancer of colorectal cancer metastasis. *J Natl Cancer Inst*. 1990;82:380-385.
6. Benson AB, Venook AP, Cederquist L, et al. NCCN Clinical Practice Guidelines in Oncology: Colon Cancer. NCCN.org March 14, 2018.
7. Burgers K, Moor C, Bednash L. Care of the Colorectal Cancer Survivor. *Am Fam Physician*. 2018;97(5):331-336.
8. Duffy M, Dalen AV, Haglund C, et al. Clinical utility of biochemical markers in colorectal cancer. *Eur J Cancer*. 2003:39;718-727. doi:10.1016/s0959-8049(02)00811-0
9. O'Brien MJ, Steele G. Correlation of plasma CEA and CEA tissue staining in poorly differentiated colorectal cancer. *Am J Med*. 1981;71:45-249.
10. Yu WS, Lee JG, Paik HC, et al. Carcinoembryonic antigen predicts waitlist mortality in lung transplant patients with idiopathic pulmonary fibrosis. *Eur J Cardiothorac Surg*. 2018 Apr 17. Doe: 10.1093/ejcts/ezy170.
11. Imamura M, Morimoto T, Nomura T, et al. Independent prognostic impact of preoperative serum carcinoembryonic antigen and cancer antigen 15-3 levels for early breast cancer subtypes. *World J Surg Oncol*. 2018:16;26. Doi: 10.1186/s12957-018-1325-6.
12. Gardner RC, Feinerman AE, Kantrowitz PA, et al. Serial carcinoembryonic antigen (CEA) blood

levels in patients with ulcerative colitis. *Am J Dig Dis*. 1978:23;129-133.

13. Maestranzi S, Przemioslo R, Mitchell H, Sherwood RA. The effect of benign and malignant liver disease on the tumour markers CA 19-9 and CEA. *Ann Clin Biochem*. 1998:35(Pt 1); 99-103.

14. Duffy MJ. Carcinoembryonic antigen as a marker for colorectal cancer: Is it clinically useful?" *Clinical Chemistry*. 2001;47(4):624-630.

15. Stout RL, Fulks MW, Dolan VF, Magee ME, Suarez L. Increased Mortality Associated with Elevated Carcinoembryonic Antigen in Insurance Applicants. *J Insur Med*. 2007;39:251-258.

16. Stout RL, Fulks MW, Dolan VF. Underwriting Implications of Elevated Carcinoembryonic Antigen. *On The Risk*. 2009;25:48-54.

17. Harrell FE. *Regression Modeling Strategies: With Applications, to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2nd ed. Heidelberg: Springer. Biom. 1984;72:18-26.

18. Ashley T, Cheung L, Wokanovicz R. Accuracy of Vital Status Ascertainment Using the Social Security Death Master File in a Deceased Population. *J Insur Med*. 2012;43:135-144.

19. Wesley D, Cox HF. Modeling Total Cholesterol as Predictor of Mortality: The Low-Cholesterol Paradox. *J Insur Med*. 2011;42:62-75.