

CPE WORKING PAPER SERIES

Diseases and Comorbidities for the Union Army:  
Survival Analysis and Impact on Death\*.

Paula Canavese\*\* and Claudia Linares\*\*\*

Working Paper xxxx-xx

CENTER FOR POPULATION ECONOMICS

1101 East 58<sup>th</sup> Street

Chicago, IL 60637

October, 2003

(First draft)

---

\* Prepared for the Center for Population Economics. The views expressed herein are those of the authors and not necessarily those of the Center for Population Economics.

\*\* CPE Graduate Research Assistant, Graduate School of Business, University of Chicago. Fourth year student in the Ph. D in Economics, Department of Economics, University of Chicago.

\*\*\* CPE Graduate Research Assistant, Graduate School of Business, University of Chicago. M.A. International Relations, San Francisco State University.

Diseases and Comorbidities for the Union Army: Survival Analysis and Impact on Death  
P. Canavese and C. Linares  
CPE Working Paper No. xxxx-xx  
October 2003

### **ABSTRACT**

Long-run historical comparisons of health status are crucial to illuminate the patterns of recent years. However, many conclusions in this area, especially with respect to the actual burden of diseases and comorbidities, are still provisional due to several lacunae in the analysis of data available. The goal of this paper is to bring broad conclusions to this kind of analysis by taking advantage of the information contained in the Union Army database.

Through the use of logistic regressions and OLS, we obtain results on the impact of different diseases and their combinations over cause and age of death. Later, we apply survival analysis to ascertain the determinants of waiting time to death and their relative importance. With these results, it is finally possible to make the large pool of evidence on diseases in the nineteenth century speak to the questions of health and aging in present times.

Paula Canavese  
Claudia Linares  
Center for Population Economics  
Graduate School of Business  
The University of Chicago  
1101 East 58<sup>th</sup> Street  
Chicago, IL 60637  
Tel: (773) 702-7709  
Fax: (773) 702-2901  
[paula@cpe.uchicago.edu](mailto:paula@cpe.uchicago.edu)  
[clinares@cpe.uchicago.edu](mailto:clinares@cpe.uchicago.edu)

## 1. INTRODUCTION

Because historical comparisons of health status over long periods of time are crucial to illuminating the patterns of recent years, it is unfortunate that many conclusions in this area, especially with respect to comorbidities and the burden of disease, are still provisional due to several lacunae in the available data. The Union Army database contains much information that can be gleaned about how with each other, by tapping into the information contained in. Such is the goal of the current paper.

Through the use of logistic regressions and OLS, we obtain results on the impact of different diseases and their combinations over cause and age of death. Later, we apply survival analysis to ascertain the determinants of waiting time to death and their relative importance. With these results, it is possible to make the large pool of evidence on diseases in the nineteenth century speak to the questions of health and aging in present times.

Some important work has already being done toward this goal, viz. Costa (2000, 2003) and Lee (2003). However, the present paper goes farther than the extant literature on several counts. First, we take a more comprehensive view of the Union Army database, considering all diseases. Second, our analysis covers both a longer period of time and more patients since it considers recruits who appear in the Surgeon's Certificates, not only a sub-sample. Importantly, our paper also deals systematically with combinations of diseases, introducing these combinations as explanatory variables in their own right. This innovation circumvents biases that may result from analyzing the diseases separately while at the same time it sheds light on the interaction among different health conditions in the sample.

The paper begins with a description of the data. It then proceeds to examine the relationship between diseases, comorbidity and death, how the cause of death and the age of death are explained by having been diagnosed with specific diseases or combination of them. This study also presents the survival analysis of the Union Army recruits with a specific impairment as well as the impact that age of onset of disabilities, the number of comorbidities at onset and the rates of deterioration of the specific chronic diseases have on the duration of the disease. Then we summarized our results and offer a conclusion. All the tables are contained in the Appendix.

## 2. DATA

The data used in this paper come from the Surgeons' Certificates as well as from the Military, Pension and Medical Records, both at the Center for Population Economics<sup>1</sup>.

### 2.1. The Surgeon's Certificates

The Surgeons' Certificates data contains 87,223 medical exam records for 17,721 pensioners. A medical examination was given when a Civil War veteran applied originally for a pension and every time he asked for an increase in the pension amount. As part of the pension process, a veteran had to visit a board of physicians, appointed by the Bureau of Pensions, who would perform a thorough medical examination. The physicians would assess his general health as well as diagnose any specific impairment and record the symptoms. A rate was given for each disability in terms of its effect on a veteran's ability to perform manual labor. In addition, a rate for total disability as a result of all his impairments was given in some cases.

The exams in the data are classified by disability groups. Those groups include cardiovascular, ear, eye, gastrointestinal, genito-urinary, respiratory, musculoskeletal, liver/spleen/gallbladder, as well as infectious diseases and fevers, injury, neoplasm/tumor, nervous disorders, rectum/hemorrhoids, varicose veins, hernia, and general appearance conditions involving mainly blood, nutrition and skin, gum and teeth, and muscles.

To obtain a comorbidity pattern per each recruit, first we need to determine whether he presents the disease. For each disease group we used several variables that indicate if the veteran was diagnosed with the disease. **Table 1** shows a complete list of the variables we used. We create a dummy variable coded as 1 when evidence of the disease was shown. All cases in which we have no record of the disease or in which doctors either explicitly imply a normal condition or indicated no disease was present were coded as 0. We include all diagnoses without considering severity. For example, a headache was considered a good evidence of a nervous disorder, as was the loss of memory. Our regression analysis includes a measure of severity derived from the disability rates.

---

<sup>1</sup> See Fogel (2000, 2001).

Since there are some cases in which there is evidence of the disease only in the disability ratings, and some of the ratings correspond to more than one disease recorded in related variables, we used these two variables as evidence of disease as well. We construct a dummy variable coded 1 if the veteran has a rate for any specific condition and we also count those diseases recorded in the related variables when the disability rate did not show any evidence of the presence of them. Finally, a recruit is said to have a disease when the dummy on diagnosis has a 1 or when it has a 0 but a value of 1 in the dummy based on rating and related variables.

To get the comorbidities we converted dummies coded 1 into letters per disease group. We looked at the most frequent combinations of diseases. Analyzing those combinations that appear in the sample 100 times or more we found that the most frequent combinations of diseases generally include six of the twenty disease groups. These are cardiovascular, respiratory, rheumatism/musculoskeletal, injury/gunshot wound (GSW), general appearance, and nervous. It is important to take into account the influence of injury/GSW since this data is an analysis of war veterans and the presence of general appearance since an evaluation of general physical condition was performed for most of the pensioners.

When we analyze the comorbidities by cohort we found that there is not an important cohort effect. However, we can say that recruits were born between 1835 and 1845 present larger number of comorbidities usually as permutations of the same diseases: cardiovascular, respiratory, rheumatism, genito-urinary and injury/GSW. The most frequent combinations for all cohorts remain the same through cohorts, they basically includes injury/GSW and general appearance, the combination of cardiovascular, rheumatism and general appearance and also respiratory and general appearance. Hernias and eye disorders are also very frequent.

We also check the age effect of the comorbidities. We found that veterans younger than 35 years old were affected only by combinations that include injury/GSW usually combined with nervous, respiratory and general appearance. Combinations of diarrhea with gastrointestinal, rectum hemorrhoids or general appearance affect mostly people between 35 and 50 years old. Comorbidities that include cardiovascular conditions generally combined with rheumatism, respiratory and general appearance affect more frequently, people between 40 and 70 years old. Rheumatism/musculoskeletal in combination with cardiovascular, respiratory, genito-urinary and general appearance affect generally people at older ages, usually after 50 years old. Finally, respiratory problems in combination with cardiovascular, rheumatism and general appearance affect recruits at all ages.

Determination of the most frequent comorbidities can be biased due to those veterans who made more visits to the doctor and showed the same combination of diseases over time. We therefore decide also to look at the most frequent combinations per veteran. We summarize all the diseases per exam in a matrix in which each row shows the clinic history of a veteran from the Surgeon's Certificates and each column represents a year from 1860 to 1920. In each of these cells we have the comorbidities each recruit presented in any particular exam. Based on the more frequent combinations we create a dummy that tells us if the veteran ever had the combination or not. The frequencies of comorbidities per recruit are shown in **Table 2**.

Finally, we have added to each cell the disability rates per disease and the total rate when it is available. That means, that our final matrix has for each veteran one cell per year containing not only an indication of the diseases he had but also the rating for each of these conditions.

## **2.2. The Military Records**

This dataset was collected from military-related documents, comprising both military service records and applications made by veterans for pension support for a sample of 35,570 white volunteers from infantry regiments.

The military records give the military history of recruits, including length of service, battles fought and geographic locations where soldiers' companies were stationed. The pension records contain information relevant to meeting the legal requirements for pension eligibility. Pension records can include information relating to the military history of the veteran, as well as important health and demographic information from before, during, and after the war

We used several variables from this dataset and combined these with variables for the recruits in the Surgeons' Certificates. We obtained the year of birth for all recruits as the difference between the date at enlistment minus the age reported that same day. We also took the date of death for each recruit and the cause of death.

These variables allow us to calculate, per recruit, the years lived with a specific condition for all groups of diseases. We stipulate that all the conditions in the data are chronic: once the pensioner presented the disease, we assume he had it until he died. We determine the age at first exam, the age at first diagnosis of any specific condition, when a recruit presented a disease for the first time, the number of years he was sick from any condition and the number of

diseases he had when he was examined for the first time. The basic statistics of these variables are shown in **Table 3** and **Table 4**. We will use those variables later in our analysis.

For the purpose of this study we consider the first exam and the first diagnosis as the first time the recruit was examined or presented the disease in the Surgeon's Certificates. Since we do not know the individual's medical history before entering the sample, we have to rely on this date as an indicator of the onset of a disease. It is important to notice that this indicator may be biased because, in its initial stages, the pension system only admitted war-related diseases, a requirement that was abandoned in 1890's Disability Act. As a result some claimants may have had impairments long before they were examined for the first time. However, the health condition of recruits mitigates the bias. It is reasonable to assume that upon being enlisted, they were for the most part in good health. Furthermore, their average age was 23 at that time, meaning that the ages for which we have no direct evidence (between early 30s and early 50s) they were relatively healthy. Finally, some of the law's requirements were gradually relaxed even before 1890, so that the number of records that underestimate the onset of diseases is less than what may seem at first glance.

We also used the cause of death and the contributing cause of death variables. In order to make comparison with the disease data we coded the cause of death in terms of the same groups of diseases we have in the Surgeon's Certificates data<sup>2</sup>. Then we added the contributing cause of death since there are some diseases that are not lethal but certainly contributes to death. For example, for a veteran with a cancerous tumor who died for a heart attack, the cause of death is cardiac failure but cancer is a contributing cause of death. For our purposes it was worth keeping both the cause and the contributing cause. Most of the recruits have more than one disease as a cause of death, and so we assumed that the patient died as a result of all of them. We create dummies per disease group coded 1 if the cause of death was that particular condition.

**Table 5** shows how many observations we have in each of these 20 groups. Of the 16,245 recruits included in our sample, 78 % have causes of death that can be classified in one of the Surgeons' disease groups. The rest are classified as "other" cause of death.

**Table 6** describes the number of recruits sick from each particular disease who died from that specific condition.

---

<sup>2</sup> General appearance as a cause of death includes basically old age, debility, inanition and senility.

### 3. DISEASES, COMORBIDITY AND DEATH

#### 3.1. Diseases, Comorbidity and Cause of Death

The first question we pose is whether there is a relationship between conditions diagnosed in the Surgeon's Certificates and the cause of death. We also wonder if the severity of the diseases and the most frequent comorbidities have an effect on the cause of death of the recruits. In order to explore these relationships we run some logistic regressions and based on them we calculate the probabilities of dying from each specific condition.

**Table 7**<sup>3</sup> shows the results of our first regression where we use a dummy for each cause of death as dependent variable. The explanatory variables are the dummies of having the condition and, an approximation of how severe the condition is, the mean of disability ratings given to veterans for each disease over 60 years. The results show that the dummy variable indicating whether a veteran had the condition diagnosed at any time in his life is not significant for all the causes of death. But the specific disability rate is significant in all the cases, except for the eyes disorders, and has a positive sign, which suggests that the probability of dying from any particular condition is greater the more severe the disease is.

**Table 8** shows the probabilities of dying of each particular disease calculated on the basis of the results of the regression in **Table 7**. In the analysis per disease the greater average probabilities are for cardio, respiratory and genito-urinary conditions, and the lower average probabilities are for eye disorders and ear conditions.

Given the fact that being diagnosed with the disease from which he died seems to have no effect on veteran's probability of dying, the question then is which diseases could have had an impact on the probability of dying of any specific condition. We use the forward procedure<sup>4</sup> to run a new logistic regression with the cause of death as dependent variable and dummies for having a disease from each condition group as independent variables (**Table 9**). For most of the diseases the probability of dying increases when the veteran was diagnosed with the same disease except in the cases of ears, general appearance, and spleen where the probability of dying as a result of being sick of that conditions is completely independent.

---

<sup>3</sup> In the table we have excluded some diseases for which the number of observations did not allow to get confident results.

<sup>4</sup> Starts by including the best single regressor, then finds the best one to add to what exists; the next best, and so on.



We can also see the effect of having diseases other than that of which the veteran died on the veteran's death of any specific condition. For example, the second column of **Table 9** shows that the probability of dying from cardio is positively affected not only by having cardio but also by having a nervous condition. Additionally, having infectious disease influenced the probability of dying of respiratory conditions and presenting cardiovascular condition positively affected the probability of dying of rheumatism.

A third question arises here. If the veteran did not die from a condition he was diagnosed then what was his cause of death. To answer this question we reverse the causality by using as dependent variable a dummy indicating whether the veteran was diagnosed with each specific disease and as independent variables causes of death other than that condition used as dependent variable<sup>5</sup>. **Table 10** shows what other causes would more likely explain the death given that the patient had any specific condition. Cardiovascular diseases are a very common cause of death when a veteran had other diseases. In the contrast, people suffering of respiratory deficiencies and varicose vein problems only died from those causes. Most of the diseases show one or two related causes of death. For example, if a veteran had diarrhea but he did not die of that disease he more likely would die of gastrointestinal diseases or rectum/hemorrhoids. (**Table 10**, column 3).

Based on the results from the regression in **Table 10**, we have computed the probabilities of having each particular condition and reported these in **Table 11**. Here we see that the greater probabilities were for having general appearance problems, rheumatism or cardiovascular conditions and the lower average probabilities were for having gallbladder and endocrine disorders.

Finally, we want to see the effect of comorbidity in the cause of death. Using again the forward procedure, we ran a regression where the dependent variable is a dummy for the cause of death and the independent variables were dummies indicating whether the veteran had the specific combination and the disease, no matter how many times he was examined for the same combination. In this case we used the most frequent comorbidities, obtained in the previous section. According to these results for combinations that include diseases such as cardiovascular, diarrhea, eye, injury, respiratory and rheumatism, we usually observe a positive effect in the probability of dying of the same diseases. For example, having a combination of cardiac problems and rheumatism or one of heart and injury increase the probability of dying of cardiovascular disease (**Table 12**, column 2). For the rest of the diseases the impact of comorbidity in the probability of dying of any specific condition is not very conclusive. For example, being sick

from heart disease and rheumatism have a negative effect on the probability of dying on genitourinary conditions but a combination of cardiac, general appearance and injury problems have a positive effect on that probability (**Table 12**, column 9).

### 3.2. Diseases, Comorbidity and Age of Death

In this section we want to analyze whether there is any relationship between the time of death and the diseases the veterans had over his lifetime, as well as whether having a particular disease and dying from a specific condition had an effect on the time of the death. To measure the time of death we considered two different variables: the year and the age at death. However, in this paper we report the results only based on the age of death, since using year of death leads to similar conclusions. Then, using least squares with the forward procedure we performed a set of regressions using age of death as dependent variable and controlling by age at the first exam.

The first thing we want to evaluate is what kind of diseases precipitated dying at younger ages. **Table 13** displays the results of the regressions using a dummy for each condition the veteran could have had during his lifetime as independent variables. In almost all cases the sign of the coefficients is positive implying that having a condition increases the age at death. This result may seem strange, since it says that people who are sick live longer. However, we have to remember that we are considering chronic conditions and not acute ones from which a person dies quickly. On the other hand, it is important to understand that the benchmark against which these cases are compared is not that of perfectly healthy patients but that of patients who may be suffering an alternative, possibly more deadly, combination of diseases. The only significant condition that causes the death at younger ages is diarrhea.

Besides the impact of specific conditions on the age of death we also wished to explore which combination of diseases produces death at younger ages. In **Table 14** the independent variables are the most frequent comorbidities we mentioned in Section 2.1. In this case again, generally the signs of the coefficients are positive reflecting the fact that we are dealing with chronic diseases. Only a very few combinations had a negative impact on the age of death: general appearance and respiratory conditions; general appearance, injury and nervous; and general appearance, injury and respiratory. Removing the effect of general appearance, since most of the recruits had a general physical evaluation, we can say that combinations of

---

<sup>5</sup> We use a forward procedure.

injury/GSW, nervous ailments, and respiratory conditions are good indicators of dying at younger ages.

Until now we have only evaluated the impact of diseases and comorbidities on age of death. It is also interesting to consider the joint effect of other variables that could have an impact on the age of death. Therefore, we ran regressions for each of the 20 conditions in the Surgeon's Certificates using the age of death as dependent variable. In each regression, we used as independent variables a dummy for having the condition, a dummy when the cause of death is the same condition with which the veteran was diagnosed, the number of years the veteran had the disease, the age at which he first was diagnosed with the condition and the disability rate he received for the disease at first diagnosis.

**Table 15** shows the results of these regressions. The most significant effects for all diseases were the effects produced by the number of years the veteran lived with the disease and the age at first diagnosis. The coefficients for these variables are positive, meaning that when the number of years sick increases or when the veteran was diagnosed at an older age, he would probably die at an older age. This is again a pattern characterizing chronic diseases. However, when significant, the disability rate had a negative impact on the age of death, which implies that the more severe the disease, the younger a veteran would die. Since the dummies indicating the veteran had the condition and the one for the cause of death are not significant, the facts solely of having a disease or of dying of it had no effect on how old a veteran would be at the time of death. According to these regressions, the severity of the disease and the duration of it are the important factors in determining the age at death.

#### **4. SURVIVAL ANALYSIS**

A better way to explain the effect of different factors on the duration of a disease is by applying the statistical techniques of survival analysis<sup>6</sup>. The goal in our survival analysis is to characterize the distribution of the survival time for each specific condition for the Civil War veterans, in order to study the relationship between survival time and some concomitant variables such as age, number of diseases and severity of diseases at first exam.

Survival data consist of a response variable that measures the duration of time until a specified event occurs (event time, failure time, or survival time) and a set of independent

variables thought to be associated with the failure time variable. These independent variables can be either discrete, such as gender or race, or continuous, such as age or temperature. The purpose of survival analysis is to model the underlying distribution of the failure time variable and to assess the dependence of the failure time variable on the independent variables. In our case this specific event is death, so we try to explain the time until death after being diagnosed with a disease.

An intrinsic characteristic of survival data is the possibility of censoring observations because the actual time until the event is not observed. Such censoring can arise from withdrawal from the experiment or termination of the experiment. Because the response is usually the duration, some of the possible events may not yet have occurred when the period for data collection has terminated.

In our case, some veterans did not have a registered date of death. For these veterans we cannot analyze the time until death, and consequently the time to the event may not be ascertained for all veterans in the sample. These observations are said to be right censored. Thus, an additional variable is incorporated into the analysis indicating which responses are observed event times and which are censored times.

The association between covariates and the survival time variable can be investigated by computing estimates of the survival distribution function within strata defined by the covariates.

Our analysis proceeded in two steps. First, we used a lifetest procedure to compute non-parametric estimates of the survival distribution function. The second step was then to examine the relationship between the concomitant variables and survival time through a regression model in which the survival time has a distribution that depends on the concomitant variables. The regression coefficients may be interpreted as describing the direction and strength of the relationship of each explanatory variable to the effect of the survival time.

Following Cox we calculate the hazard rate of dying from a particular type of disease<sup>7</sup>. The regression coefficients are interpreted as the increase of the log hazard ratio resulting in the increase of one unit in the covariate. However, the underlying hazard function is left unspecified

---

<sup>6</sup> Also known as Event History Analysis, Failure Time Analysis or Hazard Analysis.

<sup>7</sup> The Cox model is a semiparametric model in which the hazard function of the survival time is given by:  $h(t | x) = h_0(t) \exp(\beta'x(t))$ , where  $h_0(t)$  is an unspecified baseline hazard function,  $x(t)$  is a vectors of covariate values, possibly time-dependent, and  $\beta$  is a vector of unknown regression parameters. The model is referred to as a semiparametric model since part of the model involves the unspecified baseline function over time and the other part involves a finite number of regression parameters.

and, as in any other model, the results can be misleading if the proportional hazard assumption does not hold.

#### 4.1. Lifetest Analysis

The first thing we would like to explain is the survival time of veterans who are sick from any condition. The dependent variable is the years a recruit was sick after the first exam<sup>8</sup>. We censored by whether the veteran had not died until after the end of our observing period. **Table 16** shows the results. We can see here that after being sick for 20 years the probability of dying for Union Army veterans decreases as the number of years living with a disease increases. For example, for a veteran who was sick between 20 and 30 years, the probability of dying during that interval is almost 40%, while a veteran sick between 40 and 50 years the probability is about 30%. Therefore, the survival probability possesses an inverse pattern: it decreases as the years sick increase. The same thing happens with survival time, that is, a person who was sick at most 10 years but did not die would survive approximately another 25 years. But a person sick for 40 years had the possibility of living only an extra 19 years.

We used the same method to analyze the survival time for each specific disease, but censoring by whether the veteran had died of the specific condition analyzed.

Life tables results are in **Table 17**. We included only the results for those conditions we had more observations to perform the estimation. We also added a graph (see **Graph 1**) of the survival functions for those conditions. The highest probabilities of dying are for cardiovascular diseases, no matter which interval of years sick we looked at. The lower probabilities are those computed for rectum hemorrhoids. In **Graph 1**, we see that for diseases like diarrhea, gastrointestinal, and rectum/hemorrhoids the probability of dying is scarcely affected by the number of years the person was sick. Injury/GSW has a very similar curve to these diseases. It is interesting to note, however, that it does not show the initial increase in the probability of dying from an injury, accident or gunshot. This is probably explained by the fact that veterans asking for a pension and examined for this cause already passed the time with the highest probability of dying as a result of the cause. The opposite is demonstrated in the cases of people having cardiovascular conditions, genito-urinary problems and respiratory diseases, who were more

---

<sup>8</sup> It was calculated as the difference between the date of death and the year the veteran had the first exam with a positive diagnosis of having any disease.

likely to die as the number of years living with the disease increased. The survival function for nervous and rheumatism is affected more when the veteran had lived with the disease for a long time.

## 4.2. Proportional Hazard Estimation

In this section, instead of analyzing the survival time of each particular chronic condition, we determine what factors explain the duration of the conditions.

The first step was to study the influence of different covariates on the time until death if the veteran was sick of any of the 20 conditions. **Table 18** shows the estimated effects of age at first exam, number of diseases at first exam and the severity of the diseases diagnosed at first exam<sup>9</sup>. We right censored those veterans who were not observed to have died before the end of our sample period.

It seems counterintuitive that first diagnosis at older ages and greater severity of the conditions are associated with higher hazard of duration of diseases while the number of diagnosed diseases is related with lower hazard of time until death. We would expect that being diagnosed with a disease later in life and having more severe conditions will kill more quickly, but again we need to remember that we are studying chronic conditions with which a person can live a long time.

The third column in **Table 18** shows the effect on the years sick of increasing each explanatory variable in one unit, holding the remaining variables constant. For each one-year increase in the age at the first exam or one unit increase in the disability rate the probability of being sick longer increased, while it decreased when one additional disease was diagnosed.

If we run a least squares regression using as dependent variable the number of years sick after first exam and the same independent variables as the hazard regression in **Table 18**, the coefficients of age and of the total rate are negative. That is, the effect of increasing the age of diagnosis or the severity of the disease reduced the remaining years of life. This seems a more logical result than that of the hazard regression but we have to be careful, since the meaning of the coefficients is very different in each case. In a hazard regression the coefficients tell what variation in the probability is while the least square coefficients are in levels. More important is

the fact that when using OLS we can get biased estimates because it does not consider the censored observations. But censoring is not the only reason for preferring survival analysis; a better reason is that the distribution of event times are often far from normal, making survival model superior to OLS regression.

We would like to extend our analysis to the duration of each specific disease. We produce these estimates in **Table 19**. Here the variable that is most significant in explaining the duration for all diseases is the age at first diagnosis, followed by severity of the disease<sup>10</sup>, which also has a positive impact when significant. For almost all conditions, the number of diseases (including the one under study) is not statistically significant.

We can see in **Table 20** that for most diseases the effect of one year increase in age at first diagnosis represents an increase of almost 9% in the probability of being sick longer with that condition. Again, we obtain the opposite result if we run the regression using least squares because of the reasons explained above.

Another way of studying the influence of different factors on the time until death after a disease is diagnosed is to use a competing risk analysis. To apply this analysis we divided the recruits in two groups: those who died of any of the 20 conditions (group one) and those veterans who died of different causes (group two). Then, by using proportional hazard estimation, we run a regression using the years the recruits were sick from any condition. As independent variables we include age at first exam, total number of diseases during a veteran's life and the difference in the total rate as a proxy for the severity of the disease<sup>11</sup>. We censored those who died of causes different from those specified (group two).

The results of this analysis are presented **Table 21**. The variable with less power to explain the duration of a disease is found to be the disability rate, while the age at first exam and the total number of diseases a veteran was diagnosed with produce very significant results. An increase in one year in the age at first exam resulted in an increase of almost 7% in the probability of being sick longer. Number of conditions<sup>12</sup> and severity of a disease had a negative effect on the duration

---

<sup>9</sup> To measure severity we have used the total rate. Since for many recruits the total rate was not available we approximated it by adding the specific rates (see Canavese and Linares, 2003).

<sup>10</sup> In this case the severity is measured by the specific rate for each condition that was given by the Surgeon at the first diagnosis.

<sup>11</sup> The difference in the total rate is the total rate at last exam minus the total rate at first exam. Again we approximate the total rate as the sum of specific ratings per conditions in that exam. (See Canavese and Linares, 2003).

<sup>12</sup> Calculated as the sum of the diseases the veteran had in the sample.

of the illness. Therefore, we will expect a veteran to live shorter with a disease if its severity was greater or if he had larger number of conditions.

## **5. SUMMARY AND IMPLICATIONS**

This study has explored the impact of diseases and comorbidities on death and duration of illnesses for the Union Army veterans.

We found that the comorbidity pattern for the Union Army veterans was principally a combination of six groups of illness according to the Surgeons' Certificates classification. These are: injury, cardiovascular, rheumatism, respiratory, nervous and general appearance.

We analyzed the impact of diseases, comorbidities and their severity on the cause of death. We found that the effect solely of having a disease is not significant in explaining the cause of death from the same condition. However, the joint effect of having several diseases is positive and significant in explaining that variable. Our results show that the probability of dying from cardiovascular conditions is affected not only by having cardiac problems but also by the presence of nervous conditions. Further, having a combination of infectious and respiratory conditions explains death from respiratory troubles, as having a combination of cardiovascular and rheumatism/musculoskeletal problems explains dying of rheumatism.

Then we looked at cases where the cause of death was different from the disease the veteran presented in his exams. Here we found that, in general, dying from cardiac problems is very common when a veteran had other diseases. Outside of cardiovascular deaths, other very related diseases could explain a cause of death. For example, if a veteran was diagnosed with diarrhea and did not die from it, his more probable cause of death would be gastrointestinal disease or rectum/hemorrhoids. For respiratory and varicose veins, no other cause but the disease itself explains the death.

From the analysis of the impact of specific combinations on the cause of death we can conclude that those with cardiovascular, diarrhea, injury, respiratory and rheumatism tend to die of these same conditions.

The severity of diseases is significant and shows that the probability of dying of any condition is greater the more severe the disease is.



We also explored the impact of diseases and combinations of diseases on dying at younger ages. Our findings show that the most lethal conditions are diarrhea, infectious, respiratory and endocrine diseases. The rest of the disease groups are very chronic and do not cause death at younger age. We also found that the comorbidities that include injury and respiratory conditions are a good indicator of dying younger. The fact that most of the diseases are chronic is also reinforced when we analyzed the effects of other factors, such as the number of years living with the disease and the age at first diagnosis. When the number of years sick increases or when the veteran was diagnosed at an older age, he would probably die at an older age.

The results of the survival analysis show that the probability of dying for the Union Army veterans decreased as the number of years with the illness increased. The highest probability of dying was for cardiovascular conditions. For conditions like genito-urinary and respiratory ailments, the probability increased as the number of years with the disease increased. However, for diseases like diarrhea, gastrointestinal disorders and rectum/hemorrhoids the probability of dying was scarcely affected by the duration of the disease.

We conducted a hazard analysis of time until death after diagnosis with a disease. We found that being examined at older ages or diagnosed with a more severe disease implied a higher hazard of duration of the disease while the number of conditions represented a lower hazard. For each one-year increase in the age at first exam or one unit increase in the disability rate the probability of being sick longer went up by 6.9% and 95% respectively.

Although these results allow us to gain better understanding of the pattern of diseases and comorbidities for the Union Army veterans and the impact on death and waiting time until death, our study will be incomplete until we compare its results with twentieth-century data. Therefore, our future research will focus on analyzing the occurrence of similar patterns in current data.

## 6. REFERENCES

Canavese, Paula and Linares, Claudia, “*Alternative Measures to the Total Disability Rates. Surgeon’s Certificates Data*”. CPE Working Paper Series #2003-2, 2003

Costa Dora, “*Understanding Mid-Life and Older Age Mortality Declines: Evidence from Union Army*”, Journal of Econometrics, January 112 (1) 175-92, 2003.

Costa Dora, “*Understanding the Twentieth Century Decline in Chronic Conditions Among Older Men*”, Demography 37(1) 53-72, 2000

Fogel, R.W, *Public Use tape on the Aging of Veterans of the Union Army: Surgeon’s Certificates, 1862-1940, Version S-1 Standardized*. Center For Population Economics, University of Chicago, Graduate School of Business; and Department of Economics, Brigham Young University, 2001.

Fogel, R.W, *Public Use tape on the Aging of Veterans of the Union Army: Military, Pension, and Medical Records, 1860-1940, Version M-5*. Center For Population Economics, University of Chicago, Graduate School of Business; and Department of Economics, Brigham Young University, 2000.

Lee Chulhee, “*Prior Exposure to Disease and Later Health and Mortality. Evidence from Civil War Medical Records*”, in Health and Labor Force Participation over the Life Cycle, Dora Costa, ed., Chicago: University of Chicago Press for NBER, 2003.

## 7. APPENDIX

**Table 1**

This table describes the variables per disease used from the Surgeon's Certificates, to determine whether the veterans have the disease. The first column gives a description of the condition for each disease group while the second one gives the name of the variable that identifies that condition in the data set. The numbers between brackets represents the number of instances for each condition. The variables indicated with ST are standardized and the ones indicated with Y/N are binary variables.

<b>A. CARDIOVASCULAR</b>	
1.If the claimant had palpitations	c_palp (Y/N)
2.The presence of murmur	c_murmur (Y/N)
3.Dilation enlargement Hypertrophy	c_enalrg (Y/N)
4.Oedema (Oedema and Anasarca)	c_oedema (Y/N)
5.Any indication of Cyanos or Bluing	c_cyanos (Y/N)
6.Difficult Breathing and Dyspnea	c_dyspne (Y/N)
7. Arteriosclerosis, hardening of the arteries, atheroma and Bruit.	c_artscl (Y/N)
8.Poor or impaired circulation	c_impeir (Y/N)
<b>B. DIARRHEA</b>	
1.If the patient had Diarrhea (flux), Disentery (Bloody flux) or Constipation	d_dia[1-2] (ST)
<b>C. EAR DISEASE</b>	
1.If the claimant was totally deaf	e_def[1-3] (Y/N)
2.If there was Otitis Media/Inflammation of the middle ear (Aural catarrh, discharge, drainage, injection, pus, supuration, congestion and inflammation )	e_inf[1-3] (Y/N)
3.Was/Were the eustaquian tube(s) blocked	e_eus[1-3] (Y/N)
4.Was/Were the tympanic membrane or eardrum(s) mentioned?	e_drm[1-3] (Y/N)
5.Was there disease of Mastoid	e_mas[1-3] (Y/N)
6.Other causes of hearing loss mentioned: Upper respiratory infections, noise, trauma, infections, measles, small pox, etc	e_cau[1-2][01-06] (ST)
<b>D. ENDOCRINE DISEASE</b>	
1.Was the thyroid enlarged (goiter?)	k_goiter (Y/N)
2.Were there any symptoms of diabetes(Sweet urine): Diabetes insipidus, excessive thirst, polydipsia, polyuria, etc.	k_dbetes (Y/N)
<b>E. EYE DISORDERS</b>	
1.Conjunctiva description: Anemia, atrophy, conjunctivitis, discharge, hyperemia, icteric, scar trachoma, ulcer.	y_cnj[1-3][01-12] (ST)
2.Cornea description:atrophy, inflammation, opacification, purulent, scar, staphyloma, synechia, ulcer.	y_cor[1-3][1-7] (ST)
3.Other eye inflammations and infections: Abscess, Anemia, atrophy, blefaritis, Choroiditis, congestion, detachment, discharge, hemorrhage,infection, inflammation(eyelid, iris, retina), iritis, ptosis, retinitis, scar, ulcer.	y_inf[1-3][01-10] (ST)
4.If the claimant had a cataract.	y_ctr[1-3] (Y/N)
5.If the claimant had Arcus senilis.	y_arc[1-3] (Y/N)
6.If Pterygium, Chalazion, Coloboma or Pannus.	y_ptr[1-3][1-5] (ST)
7.Ectropium or entropium designation.	y_ect[1-3][1-7] (ST)
8.Vision defects: Amaurosis, amblyopia, ametropia, astenopia,astigmatism, diplopia, emmetropia, glaucoma, hypermetropia, hyperopia, myopia, paralysis, photophobia, presbyopia, ptosis, strabismus.	y_def[1-4] (ST)

---

---

**F. GASTROINTESTINAL**

---

---

1. Gastrointestinal disorders: Diarrhea, Dyspepsia, Gastroenteritis, Pain, Ulcer.	q_ind[1-4] (ST)
2. Malassimilation Designation: Cannot digest meat, impaired digestion and poor digestion.	q_malass (Y/N)
3. Diet restriction designation: If the claimant had to follow or observe diet restrictions.	q_diet (Y/N)
4. Nausea designation: If the claimant experienced nausea after eating.	q_nausea (Y/N)
5. Claimant Vomit food designation: If the claimant vomited food after eating.	q_vmtfd (Y/N)
6. Claimants vomit blood designation: If the claimant vomited blood.	q_vmtbld (Y/N)
7. Distension of abdomen designation: Any indication of distension or enlargement of the stomach or intestines.	q_abddsc (Y/N)
8. Where tenderness is evident: Abdomen, colon epigastrium, hypocondrium, iliac, inguinal, intestine, liver, small intestine, etc.	q_tend[1-6](ST)

---

---

---

---

**G. GENERAL APPEARANCE**

---

---

1. Description of General appearance: aged, anemic, emaciated, obese, pale, etc	g_app[01-19] (ST)
2. Skin condition: Rashes, abscesses, ulcers, Eczema, etc.	g_skn[01-26] (ST)
3. Scurvy in teeth/Gums.	g_scurvy (Y/N)
4. Gums description: inflammation, bleeding, atrophy, ulceration, etc	g_gum[1-9] (ST)
5. Posture description: Stooped, cannot stand, etc	g_pst[1-6] (ST)
6. Ability to walk: Lame, uses a cane, unsteady.	g_gai[1-8] (ST)
7. general Muscle description: Atrophy, weak, etc.	g_mus[1-13] (ST).

---

---

---

---

**H. GENITO-URINARY**

---

---

1. Urinalysis: Results of urinalysis indicating the presence of acid, albumin, casts, glucose, Mucus, nitrate, phosphate stones, etc.	u_urn[01-20] (ST)
2. Tenderness: This variable indicate the presence of pain or tenderness in the genito-urinary organs: Bladder, kydneyes, prostate and urethra.	u_tnd[1-6] (ST)
3. If the patient had nephritis or inflammation of the kydneyes.	u_nphrts (Y/N)
4. If the patient had Cystitis or inflammation of the bladder.	u_cysts (Y/N)
5. Calculus designation. If the claimant suffered from Bladder Stones or Calculi.	u_calcul (Y/N)
6. Uremia Designation.	u_uremia (Y/N)
7. Stricture designation: If the patient had leakage of urine or obstruction.	u_stretr (Y/N)
8. Prostate: if the claimant presented enlarged, hard inflammation, pain or prostatectomy.	u_prs [1-4] (ST)
9. Testes Description: Abscess, absent, atrophy, cyst, discharge, epididymitis, fistula, inflamed, pain, scar, tumor	u_tst[01-18] (ST)

---

---

---

---

**I. HERNIA**

---

---

1. Kind of hernia: The type of hernia as inguinal, scrotal, oblique, direct and complete.	h_knd[1-4][1-5] (ST)
2. Hydrocele or varicocele.	h_var[1-3] (Y/N)
3. Hernia involved the scrotum: If the certificate mentioned that the hernia descended into scrotum or the claimant scrotal hernia.	h_scr[1-3] (Y/N)

---

---

---

---

**J. INFECTIOUS DISEASES AND FEVERS**

---

---

1. If the claimant suffered from infectious diseases by the time of examination.	i_cur[1-4] (Y/N)
--	------------------

---

---

---

---

**K. INJURY/GSW**

---

---

1. Body part injured.	w_bdy[1-8][1-9] (ST)
-----------------------	----------------------

---

---

---

---

**L. LIVER DISEASES**

---

---

1. If there was any enlargement of the liver.	l_enlarg (Y/N)
2. If the liver was painful or tender.	l_pain (Y/N)
3. Any indication of atrophy or induration.	l_atrphy (Y/N)
4. The presence of ascites (An effusion and accumulation of serous fluid in the abdomen)	l_ascite (Y/N)
5. Any indication of Jaundice such as yellowness of the skin or eyes.	l_yellow (Y/N)
6. If dark urine was found by urinalysis	l_dkurin (Y/N)

---

---

<b>M. SPLEEN DISORDERS</b>	
1.If the spleen was enlarged or palpable	s_enlarg (Y/N)
2. If the spleen was tender or painful	s_pain (Y/N)
<b>N. GALLBLADDER DISORDER</b>	
1. If the gallbladder was painful or tender.	z_pain (Y/N)
<b>O. NEOPLASM/TUMOR</b>	
1. Type of tumor as neoplasm, lipoma, carcinoma	t_tmrf[1-4] (ST)
<b>P. NERVOUS SYSTEM</b>	
1.If the claimant had trouble with balance at rest.	n_balnce (Y/N)
2. If the claimant had trouble with balance at movement.	n_incoor (Y/N)
3.If tere was Aphasia or difficulty speaking.	n_aphsia (Y/N)
4. Any indication of paralysis, paresis or Palsy.	n_parlys (Y/N)
5. Which reflex was affected like abdominal, brachial, patellar, pupillary, triceps, etc.	n_rfl[1-8] (ST)
6.Any indication of paralysis Agitans, tremors, trembling or tremulous tongue.	n_tremor (Y/N)
7. The existence of Hyperaesthesia or neuralgia.	n_nrg[01-18]
8. Sensory paralysis: Anaesthesia. Hypaesthesia, Numbness or feels like he is walking on cushions.	n_sense (Y/N)
9.Mention of vertigo or dizziness.	n_vrtigo (Y/N)
10.If the claimant suffered from headaches.	n_hedach (Y/N)
11.Nervousness: When the surgeon indicated the claimant was usually anxious, nervous, irritable or excitable.	n_anxous (Y/N)
12.Any indication of epilepsy or Fits.	n_eplpsy (Y/N)
13.Memory Loss: Includes forgetfulness or difficulty remembering past events as well.	n_memlos (Y/N)
14. Indication of Mental Illness like Amnesia, Anxiety, Dementia, Depression, Hypochondriasis, Homicidal, Inson	n_mil[01-23]
<u>Mania, mental retardation, Poor memory, suicidal, Violent, etc.</u>	
<b>Q. RECTUM / HEMORROIDS</b>	
1.The presence of hemorrhoids (Piles engorged or enlarged rectal veins).	m_hmrrhd (Y/N)
2.Any indication of rectal trouble or irritation.	m_rectum (Y/N)
3.Any indication of scarring of the sphincter, stricture of rectum and contracted rectum.	m_spscarr (Y/N)
4. The presence of rectal fistula.	m_fstula (Y/N)
5. If the claimant had pruritus ani.	m_prurts (Y/N)
<b>R. RESPIRATORY</b>	
1. Upper respiratory tract diseases: Adenoids, Larynx, Lungs, Mouth, Nasopharynx, Nose, Tonsil.	p_urs[01-26] (ST)
2. Dullness on percussion, consolidation or hepatization of lungs.	p_dull (Y/N)
3.Breathing sounds: Crepitation, Hyperresonance, Rhonchi, Sibilant, Wheeze.	p_snd[01-14] (ST)
Hyperresonance, Rhonchi, Sibilant, Wheeze.	
4. The presence of cough or expectoration.	p_cough (Y/N)
5. Breathing signs of respiratory diseases: Apnea,Coarse, Dyspnea, obstructive, Irregular prolonged or reterictiv	p_bre[1-8] (ST)
inspiration/expiration, tachypnea, weak.	
6.Lower respiratory diseases:Abscess, asthma, bronchitis,effussion, edema, emphysema, pneumonia,hemptysis,	p_lrs[1-7] (ST)
<u>pleuritis, tuberculosis.</u>	
<b>S. RHEUMATISM / MUSCULOSKELETAL</b>	
1. Where rheumatism was located: Back, arms, legs, knees, shoulders.	r_whr[01-21] (ST)
2. Pain of tenderness of the sciatic nerve.	r_sctica (Y/N)
3. Spinal curvature: Kyphosis or scoliosis.	r_spn[1-8] (Y/N)
<b>T. VARICOSE VEINS</b>	
1. If the claimant had varicose veins in the lower extremities.	v_veins (Y/N)

**Table 2**

This table shows the number of recruits with the most frequent diseases and comorbidities. These most frequent diseases and comorbidities are calculated by sorting the sample by exam and taking combinations that present 100 or more observations.

	Combination	Frequency	Percent
1	w	2,613	12.75
2	gw	1,239	6.05
4	r	726	3.54
5	gr	668	3.26
8	wr	600	2.93
6	wn	527	2.57
3	y	476	2.32
9	gp	474	2.31
11	g	454	2.22
13	cgr	450	2.20
12	gwr	444	2.17
10	p	419	2.04
7	h	415	2.02
16	cr	362	1.77
15	wp	283	1.38
17	gwn	279	1.36
14	yw	259	1.26
20	gpr	254	1.24
24	cgwr	245	1.20
26	cgpr	241	1.18
23	cgp	228	1.11
22	gwp	222	1.08
27	gnr	216	1.05
29	cgur	213	1.04
25	m	205	1.00
18	wv	202	0.99
32	cgnr	197	0.96
33	cwr	195	0.95
34	pr	194	0.95
36	dqgm	192	0.94
21	hw	184	0.90
31	cp	182	0.89
37	gur	180	0.88
19	yg	177	0.86
39	dqg	175	0.85
40	dg	173	0.84
30	c	168	0.82
44	ur	162	0.79
41	ep	155	0.76

	Combination	Frequency	Percent
42	cg	155	0.76
35	n	154	0.75
38	gn	154	0.75
43	cw	148	0.72
52	cpr	141	0.69
53	gwnr	133	0.65
51	yr	130	0.63
49	cgw	129	0.63
56	cur	128	0.62
62	guwr	127	0.62
61	mr	125	0.61
28	v	123	0.60
59	ygr	123	0.60
66	gmr	123	0.60
54	uw	122	0.60
63	gwpr	122	0.60
46	e	119	0.58
48	gh	117	0.57
70	dq	117	0.57
47	gww	114	0.56
50	ew	114	0.56
68	gm	114	0.56
55	ygw	113	0.55
65	d	111	0.54
69	dgm	111	0.54
60	wm	110	0.54
64	hr	108	0.53
75	gupr	107	0.52
67	nr	105	0.51
71	egp	103	0.50
79	wnr	94	0.46
85	cygr	94	0.46
77	egw	92	0.45
92	uwr	91	0.44
84	guw	90	0.44
96	cgmr	89	0.43
91	cnr	88	0.43
74	u	84	0.41
82	ewp	83	0.40
87	ghr	83	0.40
97	cmr	83	0.40
73	gv	81	0.40
80	ygn	80	0.39
95	cyr	80	0.39
93	cgwp	79	0.39
94	cuwr	79	0.39
89	qg	78	0.38
81	yp	75	0.37
86	yn	75	0.37
78	cgn	74	0.36
88	ghw	74	0.36
45	cgupr	NA	NA
57	dqglm	NA	NA
58	cguwr	NA	NA
72	dqgmr	NA	NA
76	dqglsm	NA	NA
83	cgwpr	NA	NA
90	cygur	NA	NA
	Total	20,494	100.00
	Total Observations	16,610	

c=cardio, d=diarrhea, e=ear, k=endocrine, y=eyes, q=gastro,  
g=g. appearance, u=genito u., h=hernia, i=infectious, w=injury, l=liver,  
s=spleen, z=gallbladder, t=neoplasm, n=nervous, m=rectum/hemorr.,  
p=respiratory, r=rheumatism, v=varicose veins

**Table 3**

This table shows by group of diseases the basic statistics for the following variables: 1) number of years a veteran was sick of each particular disease after first diagnosis, 2) age of a veteran at time of first diagnosis, and 3) number of diseases a veteran was diagnosed with at the first diagnosis of each particular disease.

	Years sick after first diagnosis ( <i>vsickdiag</i> )				Age at first diagnosis ( <i>agediag</i> )				Num Diseases at first diagnosis (including the diagnosed condition) ( <i>disdiag</i> )			
	Num. Observations	Mean	Min	Max	Num. Observations	Mean	Min	Max	Num. Observations	Mean	Min	Max
Cardio	17,673	11.067	0.000	58.000	10,714	54.968	18.000	97.000	10,674	5.377	1.000	14.000
Diarrhea	17,673	5.737	0.000	58.000	4,734	51.338	19.000	87.000	4,707	5.532	1.000	15.000
Ear	17,673	5.913	0.000	58.000	5,984	56.618	20.000	92.000	5,937	6.146	1.000	15.000
Endocrine	17,673	0.441	0.000	48.000	496	57.966	29.000	80.000	485	7.080	1.000	15.000
Eyes	17,673	7.015	0.000	58.000	6,873	55.533	18.000	92.000	6,823	5.620	1.000	15.000
Gastro	17,673	8.614	0.000	57.000	7,563	53.020	17.000	92.000	7,515	5.802	1.000	14.000
G. Appearance	17,673	16.550	0.000	58.000	13,930	51.238	17.000	92.000	13,882	4.254	1.000	15.000
Genito Urinary	17,673	7.785	0.000	55.000	8,486	58.024	20.000	97.000	8,452	5.839	1.000	15.000
Infectious	17,673	1.419	0.000	54.000	1,258	52.707	20.000	89.000	1,232	6.485	1.000	15.000
Hernia	17,673	3.311	0.000	58.000	3,133	55.006	16.000	92.000	3,122	4.640	1.000	15.000
Injury	17,673	10.474	0.000	59.000	7,860	47.506	17.000	90.000	7,814	3.478	1.000	15.000
Liver	17,673	6.643	0.000	58.000	6,100	54.126	19.000	88.000	6,052	6.614	1.000	15.000
Spleen	17,673	3.387	0.000	50.000	3,142	54.549	30.000	85.000	3,108	7.634	1.000	15.000
Gallbladder	17,673	0.136	0.000	34.000	160	61.269	42.000	86.000	157	8.363	3.000	14.000
Neoplasm	17,673	0.743	0.000	45.000	794	57.105	27.000	88.000	783	5.494	1.000	14.000
Nervous	17,673	7.921	0.000	58.000	7,416	53.904	19.000	92.000	7,351	5.327	1.000	15.000
Rectum/Hemorr.	17,673	7.472	0.000	58.000	6,517	53.021	21.000	89.000	6,490	5.584	1.000	15.000
Respiratory	17,673	9.873	0.000	58.000	9,048	53.167	18.000	95.000	8,993	5.103	1.000	14.000
Rheumatism	17,673	13.129	0.000	58.000	11,487	52.956	17.000	88.000	11,448	4.498	1.000	15.000
Varicose Veins	17,673	2.000	0.000	53.000	1,921	54.506	22.000	85.000	1,907	4.835	1.000	15.000



**Table 4**

This table shows the basic statistics for the following variables: 1) the number of years a veteran was sick from any of the 20 conditions classified in the Surgeon's Certificates after the first exam, 2) the total number of disabilities a veteran had from time of first exam until time of death, 3) the total number of disabilities a veteran was diagnosed with at the first exam, 4) the age of a veteran at the first exam, and 5) the age of a veteran at time of death.

Variable Name	Description	Num of Observations	Mean	Min	Max
<i>ysickex</i>	Years sick after first exam	15,332	24.345	0.000	75.000
<i>#dis</i>	Total number of disabilities in life	16,574	7.055	0.000	19.000
<i>#disex</i>	Total number of disabilities at first exam	16,515	2.978	1.000	14.000
<i>ageex</i>	Age at first exam	16,515	47.416	16.000	85.000
<i>agedeath</i>	Age at death	15,385	71.679	19.000	108.000

**Table 5**

This table shows the distribution of known causes of death that can be classified in each of the 20 disease groups in the Surgeons' Certificates. The number of observations for each disease group is given in absolute numbers as well as in terms of the percentage of the total number of observations for which we have a cause of death. The total number of observations for which a cause of death could be classified according to the Surgeons' Certificates appears in the last row.

Cause of death	No. Observations	Percent of total sample
Cardio	3016	18.57%
Diarrhea	446	2.75%
Ear	50	0.31%
Endocrine	381	2.35%
Eyes	5	0.03%
Gastro	480	2.95%
G. Appearance	854	5.26%
Genito Urinary	1352	8.32%
Infectious	512	3.15%
Hernia	259	1.59%
Injury	547	3.37%
Liver	209	1.29%
Spleen	316	1.95%
Gallbladder	35	0.22%
Neoplasm	503	3.10%
Nervous	928	5.71%
Rectum/Hemorr.	86	0.53%
Respiratory	1837	11.31%
Rheumatism	840	5.17%
Varicose Veins	18	0.11%
Total	12,674	78.04%

**Table 6**

The second column of this table shows the number of recruits that were sick from each condition. The third column shows the number of recruits that died from the diagnosed condition. The fourth column shows the percentage of recruits that died from the diagnosed disease.

Diseases	Number of Veterans		Percentage veterans dying from same condition they were sick
	sick of	died from	
Cardio	10,674	2,079	19.48%
Diarrhea	4,707	304	6.46%
Ear	5,937	11	0.19%
Endocrine	485	30	6.19%
Eyes	6,823	4	0.06%
Gastro	7,515	270	3.59%
G. Appearance	13,882	712	5.13%
Genito Urinary	8,452	816	9.65%
Infectious	1,232	62	5.03%
Hernia	3,122	68	2.18%
Injury	7,814	324	4.15%
Liver	6,052	102	1.69%
Spleen	3,108	65	2.09%
Gallbladder	157	2	1.27%
Neoplasm	783	54	6.90%
Nervous	7,351	493	6.71%
Rectum/Hemorr.	6,490	75	1.16%
Respiratory	8,993	1,092	12.14%
Rheumatism	11,448	626	5.47%
Varicose Veins	1,907	14	0.73%
Total	116,932	7,203	6.16%

**Table 7**

The columns of this table present the results of logistic regressions using as dependent variable the cause of death for each veteran. There is one column for each of the 20 groups of diseases classified in the Surgeons' Certificates for which reliable results were obtained. Each row contains the coefficients obtained for each explanatory variable which are: 1) a dummy for each condition, where a 1 means that the veteran had the disease which caused his death at any point in the sampled period, and 2) the average specific disability rate for that condition. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of observations included in each regression and the Akaike and Schwartz Information criteria are reported in the last three rows.

	Cause of Death													
	Cardio	Diarrhea	Eyes	Gastro	Gene. App.	Genito U.	Hernia	Injury	Liver	Nervous	Rec. Hem	Respiratory	Rheumatism	Varicose V.
Constant	-2.583 (1.045)**	-14.149 (449.8)	-16.610 (2215.0)	-15.738 (672.5)	-14.027 (522.4)	-13.633 (506.1)	-15.230 (835.5)	-2.176 (1.071)**	-15.477 (1269.2)	-14.597 (546.6)	-15.676 (600.6)	-14.395 (398.5)	-12.981 (267.8)	-15.800 (1359.8)
Dummy Cardio	1.056 (1.045)													
Dummy Diarrhea		11.083 (449.8)												
Dummy Eyes			9.908 (2215.0)											
Dummy Gastro.				11.902 (672.5)										
Dummy Gen. App.					11.303 (522.4)									
Dummy Genito U.						11.098 (506.1)								
Dummy Hernia							10.709 (835.5)							
Dummy Injury								-1.127 (1.072)						
Dummy Liver									11.499 (1269.2)					
Dummy Nervous										11.968 (546.6)				
Dummy Rec. Hem.											10.590 (600.6)			
Dummy Respiratory												12.092 (398.5)		
Dummy Rheumatism													9.975 (267.8)	
Dummy Varicose V.														10.744 (1359.8)

	Cause of Death													
	Cardio	Diarrhea	Eyes	Gastro	Gene. App.	Genito U.	Hernia	Injury	Liver	Nervous	Rec. Hem	Respiratory	Rheumatism	Varicose V.
Rate Cardio	2.109 (0.477)*													
Rate Diarrhea		5.223 (1.013)*												
Rate Eyes			1.393 (2.420)											
Rate Gastro.				8.058 (1.622)*										
Rate Gen. App.					0.885 (0.487)***									
Rate Genito U.						6.930 (1.104)*								
Rate Hernia							7.828 (2.179)*							
Rate Injury								2.685 (0.508)*						
Rate Liver									8.528 (2.900)*					
Rate Nervous										2.445 (0.369)*				
Rate Rec. Hem.											10.308 (1.877)*			
Rate Respiratory												6.758 (0.644)*		
Rate Rheumatism													3.258 (0.618)*	
Rate Varicose V.														5.790 (3.151)***
Number of Observations	6,646	3,282	2,686	2,002	2,915	1,636	2,313	4,912	973	1,867	3,960	3,981	7,865	1,117
AIC	6,888.68	1,717.20	65.81	687.58	1,455.12	1,203.28	551.82	1,893.92	307.06	1,155.65	621.14	3,347.39	3,693.18	136.27
SC	6,909.09	1,735.49	83.49	704.39	1,473.05	1,219.48	569.05	1,913.42	321.70	1,172.25	639.99	3,366.26	3,714.09	151.32

\* Significant at 1%  
\*\* Significant at 5%  
\*\*\* Significant at 10%

**Table 8**

From the logistic coefficients shown in Table 7, it is possible to calculate the probability each recruit had of dying of each specific disease. These probabilities are reported here for the same groups of diseases we reported results for in Table 7. The second column shows the number of observations used to calculate the average probability of dying from each condition, which is reported in the third column. The standard deviation, the minimum and the maximum for these probabilities are listed in the last columns of the table.

<u>Probability Dying of</u>	<u>Number of Obs.</u>	<u>Mean</u>	<u>Std Dev</u>	<u>Minimum</u>	<u>Maximum</u>
Cardio	6,997	0.214	0.023	0.074	0.642
Diarrhea	3,425	0.074	0.029	0.000	0.896
Eyes	2,858	0.001	0.000	0.000	0.005
Gastro	2,089	0.042	0.027	0.000	0.655
Gene. App.	3,113	0.068	0.009	0.000	0.137
Genito U.	1,714	0.126	0.060	0.000	0.988
Hernia	2,456	0.026	0.015	0.000	0.484
Injury	5,302	0.049	0.018	0.035	0.350
Liver	1,013	0.038	0.035	0.000	0.990
Nervous	1,960	0.097	0.053	0.000	0.454
Rec. Hem	4,158	0.016	0.014	0.000	0.446
Respiratory	4,208	0.157	0.069	0.000	0.989
Rheumatism	8,306	0.063	0.018	0.000	0.563
Varicose Veins	1,185	0.011	0.007	0.000	0.168

**Table 9**

The columns of this table present the results logistic regressions using as dependent variable the cause of death for each veteran. There is a column for each of the 20 groups of diseases classified in the Surgeons' Certificates for which reliable results were obtained. Each row contains the coefficients obtained for each explanatory variable, which are dummies for each condition veterans were diagnosed during the sampled period (a 1 means that a veteran had the disease at any point in the sampled period). To select the independent variables for each regression the forward procedure was used. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of observations included in each regression and the Akaike and Schwartz Information criteria are reported in the last three rows.

	Cause of Death																			
	Cardio	Diarrhea	Ear	Endocrine	Gastro	Gene. App.	Genito U.	Hernia	Infectious	Injury	Liver	Spleen	Gallbladder	Neoplasm	Nervous	Rec. Hem	Respiratory	Rheumatism	Varicose V.	
Constant	-1.816 (0.042)*	-3.561 (0.121)*	-4.796 (0.258)*	-3.792 (0.056)*	-3.522 (0.090)*	-3.116 (0.058)*	-2.533 (0.077)*	-4.045 (0.092)*	-3.010 (0.078)*	-2.837 (0.110)*	-4.036 (0.175)*	-3.710 (0.135)*	-6.588 (0.294)*	-3.496 (0.049)*	-3.011 (0.056)*	-6.171 (0.429)*	-1.641 (0.061)*	-3.262 (0.087)*	-8.104 (0.604)*	
Dummy Cardio	0.413 (0.048)*	-0.436 (0.113)*	-0.742 (0.315)**		-0.309 (0.107)*				-0.404 (0.103)*	-0.441 (0.105)*							-0.183 (0.060)*	0.186 (0.088)**		
Dummy Diarrhea		2.050 (0.114)*																		
Dummy Ear						0.287 (0.075)*												-0.284 (0.060)*		
Dummy Endocrine				1.116 (0.197)*												1.013 (0.412)**				
Dummy Eyes		-0.279 (0.110)**				0.181 (0.074)**														-0.127 (0.056)**
Dummy Gastro.					0.555 (0.112)*															
Dummy Gen. App.			-0.771 (0.329)**				-0.246 (0.085)*			-0.416 (0.111)*	-0.399 (0.202)**	-0.324 (0.146)**								
Dummy Genito U.				-0.409 (0.105)*		0.577 (0.064)*	-0.502 (0.133)*	-0.398 (0.103)*	-0.307 (0.104)*									-0.192 (0.059)*	-0.442 (0.079)*	
Dummy Hernia						0.177 (0.087)**		0.577 (0.146)*				0.323 (0.135)**								-0.216 (0.071)*
Dummy Infectious									0.648 (0.150)*											0.197 (0.097)**
Dummy Injury		-0.221 (0.107)**								0.634 (0.092)*			0.704 (0.363)***			-0.980 (0.282)*				-0.178 (0.053)*

	Cause of Death																				
	Cardio	Diarrhea	Ear	Endocrine	Gastro	Gene.App.	Genito.U.	Hernia	Infectious	Injury	Liver	Spleen	Gallbladder	Neoplasm	Nervous	Rec.Hem	Respiratory	Rheumatism	Varicose V.		
Dummy Liver										0.672 (0.157)*					-0.170 (0.074)**						
Dummy Spleen									0.324 (0.124)*												
Dummy Gallbladder					0.948 (0.310)*					1.265 (0.378)*		1.962 (0.736)*				1.071 (0.537)**					
Dummy Neoplasm													0.949 (0.150)*							1.829 (0.586)*	
Dummy Nervous	0.094 (0.043)**														0.519 (0.071)*						
Dummy Rec. Hem.					0.290 (0.107)*											3.016 (0.429)*	-0.195 (0.056)*				
Dummy Respiratory		-0.246 (0.106)**							-0.245 (0.098)**							-0.570 (0.229)**	0.636 (0.058)*				-1.303 (0.540)**
Dummy Rheumatism		-0.697 (0.109)*							-0.207 (0.104)**	-0.514 (0.153)*						-1.131 (0.228)*	-0.409 (0.059)*	0.589 (0.097)*			
Dummy Varicose V.										0.414 (0.196)**											3.578 (0.638)*
Num Observations	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385
AIC	14,606	3,433	628	3,387	4,068	6,353	8,698	2,453	4,221	4,413	2,112	2,976	469	4,279	6,651	868	10,532	6,171	469	216	216
SC	14,629	3,487	651	3,402	4,114	6,383	8,721	2,476	4,267	4,459	2,158	2,999	492	4,295	6,674	921	10,616	6,202	469	247	247

\* Significant at 1%  
\*\* Significant at 5%  
\*\*\* Significant at 10%



**Table 10**

The columns of this table present the results of logistic regressions using as dependent variable a dummy indicating whether the veteran had each condition. There is a column for each of the 20 groups of diseases classified in the Surgeons' Certificates for which reliable results were obtained. Each row contains the coefficients obtained for each explanatory variable, which are dummies for each cause of death (a 1 means that a veteran died from that particular cause), excluding causes of death that belong to the same group as the dependent variable. To select the independent variables for each regression the forward procedure was used. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of observations included in each regression and the Akaike and Schwartz Information criteria are reported in the last three rows.

	Dummy																			
	Cardio	Diarrh.	Ear	Endoc.	Eyes	Gastro	Gen App	Genito U	Hernia	Infect.	Injury	Liver	Spleen	Gallbl.	Neopl.	Nerv.	Rec.Hem	Resp.	Rheum.	V. Veins
Constant	0.730 (0.021)*	-0.872 (0.019)*	-0.536 (0.020)*	-3.461 (0.048)*	-0.318 (0.020)*	-0.147 (0.019)*	1.758 (0.028)*	0.152 (0.020)*	-1.410 (0.023)*	-2.552 (0.035)*	-0.075 (0.018)*	-0.522 (0.020)*	-1.477 (0.025)*	-4.715 (0.087)*	-2.986 (0.038)*	-0.218 (0.019)*	-0.420 (0.020)*	0.245 (0.017)*	0.931 (0.021)*	-1.999 (0.026)*
Cause Death Cardio		0.088 (0.043)**		0.096 (0.042)**	0.140 (0.042)*	0.121 (0.060)**	0.191 (0.042)*		0.180 (0.075)**		0.137 (0.043)*	0.187 (0.051)*			0.176 (0.042)*	0.088 (0.043)**		0.249 (0.048)*		
Cause Death Diarrhea	-0.437 (0.100)*		-0.889 (0.404)**	-0.369 (0.105)*	0.916 (0.105)*			-0.273 (0.100)*			-0.448 (0.106)*	0.735 (0.102)*	0.678 (0.109)*				0.962 (0.103)*	-0.245 (0.099)**	-0.673 (0.102)*	
Cause Death Ear							-0.810 (0.324)**													-0.807 (0.294)*
Cause Death Endocrine								0.424 (0.124)*				0.252 (0.128)**					0.244 (0.108)**			
Cause Death Eyes	-2.116 (1.118)***					-11.093 (121.5)					-11.134 (121.5)									
Cause Death Gastro.	-0.199 (0.098)**	0.409 (0.098)*						-0.271 (0.096)*				0.202 (0.097)**	0.868 (0.316)*				0.341 (0.096)*			
Cause Death Gen. App.			0.334 (0.073)*		0.249 (0.073)*			0.272 (0.074)*	0.199 (0.087)**								0.154 (0.072)**	0.182 (0.073)**		
Cause Death Genito U.	0.218 (0.065)*									0.226 (0.102)**										
Cause Death Hernia	-0.686 (0.160)*							-0.468 (0.159)*												
Cause Death Infectious	-0.405 (0.095)*							-0.283 (0.121)**	-0.443 (0.095)*				0.227 (0.112)**						-0.352 (0.093)*	
Cause Death Injury	-0.755 (0.090)*	-0.500 (0.111)*	-0.366 (0.098)*		-0.303 (0.094)*	-0.495 (0.094)*	-0.761 (0.101)*	-0.614 (0.092)*	-0.270 (0.123)**			-0.450 (0.099)*	-0.400 (0.130)*					-0.438 (0.097)*	-0.465 (0.090)*	-0.592 (0.091)*
Cause Death Liver													0.379 (0.161)**	1.191 (0.389)*						-0.403 (0.148)*

	Dummy																				
	Cardio	Diarrh.	Ear	Endoc.	Eyes	Gastro	Gen App	Genito U	Hernia	Infect.	Injury	Liver	Spleen	Gallbl.	Neopl.	Nerv.	Rec. Hem	Resp.	Rheum.	V. Veins	
Cause Death Spleen	-0.594 (0.155)*						-0.761 (0.217)*		0.615 (0.167)*												
Cause Death Gallbladder																					
Cause Death Neoplasm																					
Cause Death Nervous							0.216 (0.105)**														
Cause Death Rec. Hem.		1.569 (0.237)*		1.462 (0.420)*						-1.094 (0.284)*	0.646 (0.237)*	0.476 (0.243)**	1.373 (0.537)**							-0.548 (0.233)**	
Cause Death Respiratory	-0.314 (0.053)*	-0.231 (0.059)*	-0.297 (0.055)*		-0.244 (0.053)*	-0.263 (0.052)*	-0.197 (0.068)*	-0.377 (0.052)*	-0.281 (0.070)*	-0.203 (0.052)*	-0.249 (0.055)*	-0.214 (0.069)*				-0.208 (0.052)*	-0.282 (0.054)*			-0.528 (0.053)*	-0.222 (0.085)*
Cause Death Rheumatism	0.559 (0.113)*							-0.419 (0.085)*	-0.391 (0.120)*	-0.146 (0.074)**											
Cause Death Varicose V.										1.056 (0.543)***					1.807 (0.573)*						
Number of Observations	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	
AIC	19,584	18,470	20,166	4,171	20,857	21,110	13,082	21,112	14,982	8,282	21,211	20,254	15,021	1,673	5,950	21,144	20,582	21,118	18,525	11,094	
SC	19,668	18,508	20,204	4,194	20,903	21,156	13,143	21,189	15,028	8,305	21,264	20,308	15,090	1,704	5,965	21,175	20,643	21,148	18,586	11,109	

\* Significant at 1%  
\*\* Significant at 5%  
\*\*\* Significant at 10%

**Table 11**

From the logistic coefficients showed in Table 10, it is possible to calculate the probability each recruit had of having each specific disease at any point in his life. These probabilities are reported here for the same groups of diseases for which we reported results in Table 10. The second column shows the number of observations used to calculate the average probability of having each condition, which is reported in the third column. The standard deviation, the minimum and the maximum for these probabilities are reported in the last columns of the table.

<u>Probability of Having</u>	<u>Number of Obs.</u>	<u>Mean</u>	<u>Std Dev</u>	<u>Minimum</u>	<u>Maximum</u>
Cardio	16,245	0.658	0.055	0.200	0.818
Diarrhea	16,245	0.291	0.038	0.168	0.751
Ear	16,245	0.367	0.032	0.232	0.471
Endocrine	16,245	0.030	0.006	0.013	0.119
Eyes	16,245	0.417	0.031	0.225	0.507
Gastro	16,245	0.464	0.049	0.000	0.713
Gene. App.	16,245	0.847	0.030	0.547	0.890
Genito U.	16,245	0.524	0.057	0.204	0.700
Infectious	16,245	0.076	0.007	0.072	0.105
Hernia	16,245	0.191	0.020	0.087	0.311
Injury	16,245	0.470	0.034	0.000	0.727
Liver	16,245	0.376	0.045	0.228	0.743
Spleen	16,245	0.193	0.032	0.110	0.514
Gallbladder	16,245	0.010	0.006	0.009	0.217
Neoplasm	16,245	0.048	0.006	0.048	0.235
Nervous	16,245	0.450	0.025	0.395	0.528
Rec. Hem	16,245	0.402	0.052	0.242	0.744
Respiratory	16,245	0.553	0.028	0.361	0.561
Rheumatism	16,245	0.702	0.058	0.228	0.765
Varicose Veins	16,245	0.117	0.007	0.098	0.119

**Table 12**

The columns of this table present the results of logistic regressions using as dependent variable the cause of death for each veteran. There is a column for each of the 20 groups of diseases classified in the Surgeons' Certificates for which reliable results were obtained. Each row contains the coefficients obtained for each explanatory variable, which are dummies for each of the most frequent comorbidities (a 1 means that a veteran had that particular combination of diseases at any point in the sampled period) To select the independent variables for each regression the forward procedure was used. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of observations included in each regression and the Akaike and Schwartz Information criteria are reported in the last three rows.

	Cause of Death																			
	Cardio	Diarrh.	Ear	Endoc.	Eyes	Gastro	G App	Genito U	Hernia	Infect.	Injury	Liver	Spleen	Gallbl.	Neopl.	Nerv.	Rec. Hem	Resp.	Rheum.	V. Veins
Constant	-1.498 (0.024)*	-3.607 (0.058)*	-5.858 (0.153)*	-3.736 (0.057)*	-9.774 (0.955)*	-3.505 (0.048)*	-2.910 (0.037)*	-2.393 (0.030)*	-4.202 (0.068)*	-3.429 (0.048)*	-3.603 (0.054)*	-4.367 (0.075)*	-3.882 (0.062)*	-6.292 (0.207)*	-3.428 (0.047)*	-2.881 (0.037)*	-5.505 (0.138)*	-2.149 (0.028)*	-2.989 (0.041)*	-7.346 (0.314)*
gw				-0.509 (0.248)**																
gr		-0.996 (0.414)**								-0.699 (0.322)**										0.445 (0.150)*
wn											0.377 (0.187)**	0.789 (0.291)*								
gp	-0.357 (0.140)**																			1.055 (0.115)*
wr																				-0.519 (0.172)*
gwr		-1.154 (0.584)**																		
wp											1.041 (0.203)*									0.512 (0.171)*
yw														1.767 (0.747)**						
gwn											0.782 (0.222)*				-1.417 (0.712)**	0.890 (0.189)*				
cr	0.483 (0.122)*																			0.703 (0.179)*
yg	-0.545 (0.246)**								0.864 (0.247)*											
hw															0.797 (0.316)**					
gwp																				0.869 (0.179)*
gnr					3.712 (1.283)*								1.471 (0.757)**							0.797 (0.233)*
cgp		-1.805 (1.005)**								-0.721 (0.342)**				1.617 (0.739)**						



	Cause of Death																			
	Cardio	Diarrh.	Ear	Endoc.	Eyes	Gastro	G App	Genito U	Hernia	Infect.	Injury	Liver	Spleen	Gallbl.	Neopl.	Nerv.	Rec.Hem	Resp.	Rheum.	V. Veins
gn																0.584 (0.269)**			-2.150 (1.007)**	
dg		1.211 (0.241)*						-0.755 (0.387)***												
cg	0.385 (0.191)**																			
cw	1.050 (0.179)*																			
gh																			0.806 (0.321)**	
wm								-1.109 (0.586)***					2.079 (1.030)**							
gm									0.800 (0.394)**								1.088 (0.527)**			
gv																				4.042 (0.660)*
qg																			1.973 (0.623)*	
dqg		1.218 (0.247)*				1.163 (0.275)*						1.358 (0.370)*					1.363 (0.498)*	0.566 (0.211)*		
cwr				0.726 (0.350)**						-1.154 (0.589)**					-1.821 (1.003)***					
gww																1.026 (0.274)*				2.026 (0.933)**
cgw				1.171 (0.378)*			0.637 (0.258)**					0.935 (0.423)**								
gmr															1.046 (0.334)*				0.673 (0.304)**	
dgm		1.243 (0.301)*															2.047 (0.443)*			2.360 (1.094)**
guw														1.849 (1.057)***						
wnr																				
ygn					2.874 (1.589)***															
egp																			0.611 (0.257)**	
ewp										0.839 (0.369)**			2.318 (1.036)**							-11.952 (200.9)
dqgm		0.992 (0.247)*	1.353 (0.727)***														1.398 (0.418)*	-0.604 (0.303)**		
guwr	0.477 (0.211)**							0.583 (0.264)**												

	Cause of Death																				
	Cardio	Diarrh.	Ear	Endoc.	Eyes	Gastro	G. App	Genito U	Hernia	Infect.	Injury	Liver	Spleen	Gallbl.	Neopl.	Nerv.	Rec. Hem	Resp.	Rheum.	V. Veins	
gwnr														2.400 (0.772)*							
gwpr																			-0.929 (0.428)**		
gupr																0.733 (0.322)**					
cygr					4.078 (1.256)*		0.834 (0.338)**														
cgmr				0.968 (0.466)**																0.820 (0.323)**	
yn					2.237 (1.666)						0.984 (0.468)**										
egw									1.232 (0.517)**												
cnr																				-1.924 (1.014)***	
Num Observations	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385	15,385
AIC	14,607.2	3,713.3	637.9	3,391.7	77.4	4,118.9	6,355.6	8,748.4	2,463.0	4,280.3	4,424.9	2,129.8	2,959.3	462.2	4,292.0	6,638.1	923.4	10,570.1	6,140.4	240.0	
SC	14,683.6	3,812.6	660.8	3,452.9	130.9	4,134.2	6,393.8	8,809.6	2,485.9	4,310.9	4,493.7	2,160.4	3,012.8	531.0	4,330.2	6,706.9	984.5	10,661.8	6,247.4	285.9	

\* Significant at 1%  
\*\* Significant at 5%  
\*\*\* Significant at 10%

**Table 13**

This table shows the results of an OLS regression using the age of death for each veteran as dependent variable. Each row shows the coefficients obtained for each explanatory variable, which are dummies for each of the diseases a veteran could have (a 1 means that a veteran had that particular disease at any point in the sampled period). We controlled for veteran's age at first exam. To select the independent variables for the regression the forward procedure was used. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of observations included and the R squared are reported in the last rows.

	Age of Death
Constant	49.656 (0.502)*
Dummy Cardio	1.662 (0.213)*
Dummy Diarrhea	-0.517 (0.254)**
Dummy Ear	1.577 (0.197)*
Dummy Endocrine	-0.819 (0.512)
Dummy Eyes	1.213 (0.188)*
Dummy Gastro.	0.461 (0.233)**
Dummy Gen. App.	1.432 (0.269)*
Dummy Genito U.	3.556 (0.202)*
Dummy Hernia	2.043 (0.221)*
Dummy Infectious	-0.277 (0.337)
Dummy Injury	1.382 (0.183)*
Dummy Liver	0.193 (0.233)
Dummy Spleen	
Dummy Gallbladder	2.031 (0.889)**



	Age of Death
Dummy Neoplasm	0.530 (0.406)
Dummy Nervous	0.461 (0.188)**
Dummy Rec. Hem.	1.093 (0.221)*
Dummy Respiratory	-0.184 (0191)
Dummy Rheumatism	2.196 (0.213)*
Dummy Varicose V.	
Age at first exam	0.285 (0.009)*
Number of Observations	15,336
R Squared	0.175

\* Significant at 1%

\*\* Significant at 5%

\*\*\* Significant at 10%

**Table 14**

This table shows the results of an OLS regression using the age of death for each veteran as dependent variable. Each row contains the coefficients obtained for each explanatory variable, which are dummies for each of the most frequent comorbidities (a 1 means that a veteran had that particular combination of diseases at any point in the sampled period). To select the independent variables for the regression the forward procedure was used. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of observations included and the R squared are reported in the last rows.

	Age of Death
Constant	71.488 (0.129)*
gp	-4.007 (0.605)*
wr	2.510 (0.526)*
gwr	1.475 (0.599)**
cgr	1.26133 (0.591)**
yw	1.598 (0.779)**
gwn	-1.543 (0.750)**
gpr	1.248 (0.782)
yg	1.757 (0.946)***
hw	2.958 (0.932)*
gwp	-1.794 (0.836)**
cgp	0.593 (0.859)
cgur	1.761 (0.853)**
cgwr	1.886 (0.788)**
cgpr	0.575 (0.798)
pr	2.126 (0.870)**
cp	0.982 (0.955)
ur	1.421 (0.964)

	Age of Death
ep	2.812 (1.009)*
yr	1.380 (1.077)
mr	1.799 (1.111)
cur	0.733 (1.084)
hr	2.350 (1.201)***
dq	0.937 (1.109)
w	-3.337 (0.287)*
y	1.303 (0.595)**
r	0.605 (0.473)
h	-0.650 (0.625)
g	-0.780 (0.590)
p	-0.642 (0.645)
v	-1.405 (1.168)
c	-1.170 (0.968)
e	1.523 (1.168)
u	-1.819 (1.348)
dg	-0.969 (0.916)
cg	0.729 (1.001)
uw	2.655 (1.108)**
wm	3.599 (1.167)*
gm	2.538 (1.152)**
ew	1.295 (1.181)

	Age of Death
gv	-1.851 (1.430)
qg	1.041 (1.400)
gur	1.209 (0.916)
cwr	0.695 (0.875)
gww	1.820 (1.124)
ygr	1.525 (1.097)
cgw	3.002 (1.080)*
dgm	1.137 (1.162)
uwr	2.691 (1.295)**
guw	4.633 (1.263)*
ghr	3.077 (1.344)**
cgn	2.643 (1.445)***
wnr	-1.345 (1.251)
ghw	2.042 (1.454)
ewp	2.365 (1.360)***
dqgm	1.673 (0.882)***
cgnr	2.747 (0.891)*
guwr	1.745 (1.093)
gwnr	1.316 (1.086)
gwpr	2.132 (1.131)***
gupr	4.085 (1.168)*
cygr	1.787 (1.255)
cmr	2.619 (1.359)***
cyr	1.305 (1.352)
egw	2.323 (1.321)***
cuwr	5.234 (1.376)*
yn	1.247 (1.426)
Number of Observations	15,385
R Squared	0.030

\* Significant at 1%

\*\* Significant at 5%

\*\*\* Significant at 10%

**Table 15**

This table shows the results of OLS regressions using the age at death for each veteran as dependent variable. Each column presents the results of a different regression, grouped by veterans who had each disease reported (of the 20 diseases in the Surgeons' Certificates the coefficients reported here are only those for the conditions with reliable results). Each row contains the results for the different independent variables, which are: a dummy taking the value of 1 when a veteran was ill of the disease; a dummy taking the value of 1 when a veteran also died of that disease; the number of years a veteran was sick of that disease; the age of a veteran when first diagnosed with the disease; and the disease's specific rate at that diagnosis. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of observations included in each regression and the R squared are reported in the last rows.

	Age of Death												
	Cardio	Diarrhea	Ear	Eyes	Gastro	Genito U.	Hernia	Infectious	Injury	Nervous	Rec. Hem	Rheumatism	Varicose V.
Constant	-4.772 (1.636)*	-3.007 (3.204)	-1.867 (1.832)	-4.774 (3.058)	-3.009 (1.970)	0.779 (3.130)	-1.370 (3.106)	-3.470 (1.803)***	-8.607 (1.569)*	-8.014 (2.308)*	-1.433 (1.948)	-4.239 (3.197)	0.856 (2.077)
Dummy for having condition	1.795 (1.591)	0.536 (3.150)	0.310 (1.720)	2.282 (3.009)	0.279 (1.771)	-3.413 (3.040)	-2.086 (3.038)	1.076 (1.547)	4.210 (1.527)*	4.186 (2.190)***	-0.372 (1.871)	2.330 (3.179)	-3.439 (1.943)***
Cause Death (that condition)	0.721 (0.096)*	-0.324 (0.217)	-1.093 (1.218)	-1.085 (1.506)	-0.466 (0.395)	-0.052 (0.230)	-0.304 (0.398)	0.144 (0.471)	0.249 (0.205)	-0.260 (0.247)	-0.438 (0.418)	-0.488 (0.148)*	-0.531 (0.800)
Years sick (ysickdiag)	1.141 (0.005)*	1.124 (0.007)*	1.126 (0.008)*	1.128 (0.007)*	1.150 (0.010)*	1.123 (0.010)*	1.113 (0.008)*	1.122 (0.013)*	1.099 (0.004)*	1.126 (0.008)*	1.148 (0.006)*	1.133 (0.005)*	1.092 (0.010)*
Age first diagnosis (agediag)	1.011 (0.006)*	1.004 (0.008)*	0.993 (0.009)*	1.005 (0.008)*	1.006 (0.012)*	1.008 (0.011)*	1.029 (0.008)*	1.010 (0.015)*	1.046 (0.005)*	1.031 (0.010)*	0.985 (0.008)*	0.992 (0.005)*	1.014 (0.012)*
Rate first diagnosis	-1.448 (0.651)**	-1.094 (1.064)	-2.085 (0.633)*	-0.711 (0.336)**	-1.055 (1.639)	1.250 (1.233)	-1.331 (1.305)	-3.816 (1.921)**	-0.881 (0.482)***	-0.498 (0.495)	-1.054 (1.182)	-1.112 (0.648)***	-0.745 (1.761)
Number of Observations	6,638	3,279	2,032	2,684	1,998	1,635	2,311	791	4,908	1,864	3,954	7,861	1,116
R Squared	0.894	0.909	0.907	0.910	0.872	0.902	0.904	0.918	0.931	0.911	0.894	0.896	0.922

\* Significant at 1%

\*\* Significant at 5%

\*\*\* Significant at 10%

**Table 16**

This table shows the results obtained when a lifetest procedure is done using the years each veteran was sick (years from first diagnosis until death) from any of the 20 conditions according to the Surgeons' Certificates. The censoring is determined by whether the veteran died during the sampled period. In parentheses are the standard errors. The total number of observations is 15,332.

Years sick with any disease in Surgeons (vsickex)	Number of deaths from diseases in Surgeons	Number of deaths other causes	Probability dying from diseases Surgeons	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	2420	0	0.158 (0.003)	1.000 (0.000)	24.838 (0.167)	0.017 (0.000)
10 to 20	3442	37	0.267 (0.004)	0.842 (0.003)	18.092 (0.153)	0.031 (0.001)
20 to 30	3634	367	0.393 (0.005)	0.617 (0.004)	14.925 (0.240)	0.049 (0.001)
30 to 40	1652	1642	0.358 (0.007)	0.375 (0.004)	17.573 (0.393)	0.044 (0.001)
40 to 50	471	1046	0.292 (0.011)	0.241 (0.004)	19.511 (0.568)	0.034 (0.001)
50 to 60	135	369	0.309 (0.022)	0.170 (0.004)		0.037 (0.003)
60 to 70	0	113	0.000 (0.000)	0.118 (0.005)		0.000 ( )
70 to 80	0	4	0.000 (0.000)	0.118 (0.005)		0.000 ( )

(1) Probability of not dying from any disease in the Surgeons' Certificates until the beginning of the interval or more.

(2) Remaining time until death from any disease in the Surgeons' Certificates for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

**Table 17**

Each of these tables show the results obtained when a lifetest procedure is done using the years each veteran was sick (years from first diagnosis until death) with one of the following conditions: cardio, diarrhea, gastro, genito urinary, injury, nervous, rectum hemorrhoids, respiratory and rheumatism. Each table corresponds to one of these conditions and in each case the censoring is determined by whether the veteran died during the sampled period from that particular condition. In parentheses are the standard errors. The total number of observations is 16,221.

Years sick with cardio (vsickdiag)	Number of deaths from cardio	Number of deaths other causes	Probability dying from cardio	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	1232	6713	0.096 (0.003)	1.000 (0.000)	36.802 (0.201)	0.010 (0.001)
10 to 20	706	2748	0.102 (0.004)	0.904 (0.003)	28.989 (0.249)	0.011 (0.001)
20 to 30	706	2604	0.201 (0.007)	0.812 (0.004)	21.734 (0.492)	0.022 (0.001)
30 to 40	331	1062	0.337 (0.015)	0.649 (0.006)	17.589 (0.745)	0.041 (0.002)
40 to 50	27	71	0.323 (0.051)	0.430 (0.011)		0.039 (0.007)
50 to 60	6	15	0.444 (0.135)	0.291 (0.023)		0.057 (0.022)

(1) Probability of not dying from cardio until the beginning of the interval or more.

(2) Remaining time until death from cardio for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

Years sick with diarrhea (vsickdiag)	Number of deaths from diarrhea	Number of deaths other causes	Probability dying from diarrhea	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	245	12026	0.024 (0.002)	1.000 (0.000)		0.002 (0.001)
10 to 20	106	1196	0.032 (0.003)	0.976 (0.002)		0.003 (0.001)
20 to 30	66	1343	0.033 (0.004)	0.945 (0.003)		0.003 (0.001)
30 to 40	27	1036	0.037 (0.007)	0.914 (0.005)		0.004 (0.001)
40 to 50	2	147	0.020 (0.014)	0.879 (0.008)		0.002 (0.001)
50 to 60	0	27	0.000 (0.000)	0.862 (0.014)		0.000 ( )

(1) Probability of not dying from diarrhea until the beginning of the interval or more.

(2) Remaining time until death from diarrhea for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

Years sick with gastro (vsickdiag)	Number of deaths from gastro	Number of deaths other causes	Probability dying from gastro	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	279	9758	0.025 (0.001)	1.000 (0.000)		0.002 (0.000)
10 to 20	85	2180	0.017 (0.002)	0.975 (0.001)		0.002 (0.001)
20 to 30	74	2236	0.026 (0.003)	0.959 (0.002)		0.003 (0.001)
30 to 40	37	1422	0.041 (0.007)	0.934 (0.004)		0.004 (0.001)
40 to 50	5	131	0.059 (0.026)	0.895 (0.007)		0.006 (0.003)
50 to 60	0	14	0.000 (0.000)	0.842 (0.024)		0.000 ( )

(1) Probability of not dying from gastro until the beginning of the interval or more.

(2) Remaining time until death from gastro for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.



Years sick with genito u. (vsickdiag)	Number of deaths from genito u.	Number of deaths other causes	Probability dying from genito u.	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	709	9270	0.061 (0.002)	1.000 (0.000)		0.006 (0.001)
10 to 20	338	3085	0.072 (0.004)	0.939 (0.002)		0.007 (0.001)
20 to 30	224	1870	0.119 (0.007)	0.871 (0.004)		0.013 (0.001)
30 to 40	75	599	0.176 (0.019)	0.768 (0.007)		0.019 (0.002)
40 to 50	5	43	0.170 (0.069)	0.632 (0.015)		0.019 (0.008)
50 to 60	0	3	0.000 (0.000)	0.525 (0.046)		0.000 ( )

(1) Probability of not dying from genito u. until the beginning of the interval or more.

(2) Remaining time until death from genito u. for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

Years sick with injury (vsickdiag)	Number of deaths from injury	Number of deaths other causes	Probability dying from injury	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	292	9418	0.025 (0.001)	1.000 (0.000)		0.003 (0.000)
10 to 20	93	1876	0.017 (0.002)	0.975 (0.001)		0.002 (0.001)
20 to 30	73	1920	0.020 (0.002)	0.958 (0.002)		0.002 (0.001)
30 to 40	53	1521	0.030 (0.004)	0.939 (0.003)		0.003 (0.001)
40 to 50	27	613	0.040 (0.008)	0.911 (0.005)		0.004 (0.001)
50 to 60	8	327	0.047 (0.016)	0.874 (0.008)		0.005 (0.002)

(1) Probability of not dying from injury until the beginning of the interval or more.

(2) Remaining time until death from injury for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

Years sick with nervous (vsickdiag)	Number of deaths from nervous	Number of deaths other causes	Probability dying from nervous	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	565	9997	0.050 (0.002)	1.000 (0.000)		0.005 (0.001)
10 to 20	158	2108	0.034 (0.003)	0.950 (0.002)		0.003 (0.001)
20 to 30	132	1895	0.054 (0.005)	0.917 (0.003)		0.006 (0.001)
30 to 40	56	1049	0.067 (0.009)	0.868 (0.005)		0.007 (0.001)
40 to 50	12	191	0.073 (0.020)	0.810 (0.009)		0.008 (0.002)
50 to 60	2	56	0.067 (0.046)	0.751 (0.018)		0.007 (0.005)

(1) Probability of not dying from nervous until the beginning of the interval or more.

(2) Remaining time until death from nervous for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

Years sick with rec. hem. (vsickdiag)	Number of deaths from rec. hem.	Number of deaths other causes	Probability dying from rec. hem.	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	37	10801	0.003 (0.001)	1.000 (0.000)		0.000 (0.000)
10 to 20	31	1859	0.007 (0.001)	0.997 (0.001)		0.001 (0.001)
20 to 30	15	2143	0.006 (0.002)	0.990 (0.001)		0.001 (0.001)
30 to 40	3	1270	0.004 (0.002)	0.984 (0.002)		0.000 (0.000)
40 to 50	0	60	0.000 (0.000)	0.979 (0.003)		0.000 ( )
50 to 60	0	2	0.000 (0.000)	0.979 (0.003)		0.000 ( )

(1) Probability of not dying from rec. hem. until the beginning of the interval or more.

(2) Remaining time until death from rec. hem. for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

Years sick with respiratory (vsickdiag)	Number of deaths from respiratory	Number of deaths other causes	Probability dying from respiratory	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	1076	8065	0.088 (0.003)	1.000 (0.000)		0.009 (0.001)
10 to 20	339	2414	0.058 (0.003)	0.912 (0.003)		0.006 (0.001)
20 to 30	267	2392	0.085 (0.005)	0.859 (0.004)		0.009 (0.001)
30 to 40	131	1321	0.130 (0.011)	0.786 (0.005)		0.014 (0.001)
40 to 50	16	153	0.115 (0.027)	0.684 (0.010)		0.012 (0.003)
50 to 60	5	42	0.192 (0.077)	0.605 (0.020)		0.021 (0.009)

(1) Probability of not dying from respiratory until the beginning of the interval or more.

(2) Remaining time until death from respiratory for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

Years sick with rheumatism (vsickdiag)	Number of deaths from rheumatism	Number of deaths other causes	Probability dying from rheumatism	Survival probability (1)	Residual lifetime (2)	Hazard (3)
0 to 10	351	6394	0.027 (0.001)	1.000 (0.000)		0.003 (0.001)
10 to 20	234	3141	0.030 (0.002)	0.973 (0.001)		0.003 (0.001)
20 to 30	176	3672	0.041 (0.003)	0.944 (0.002)		0.004 (0.001)
30 to 40	73	2004	0.058 (0.007)	0.905 (0.004)		0.006 (0.001)
40 to 50	5	142	0.048 (0.021)	0.853 (0.007)		0.005 (0.002)
50 to 60	1	28	0.067 (0.064)	0.812 (0.019)		0.007 (0.007)

(1) Probability of not dying from rheumatism until the beginning of the interval or more.

(2) Remaining time until death from rheumatism for an individual who survived till the beginning of the interval.

(3) Evaluated at the midpoint of the interval. It is the instantaneous probability of an event at time t.

**Table 18**

This table shows the results of a proportional hazard regression where the dependent variable is the number of years each veteran was sick of any of the 20 conditions classified by the Surgeons' Certificates after his first exam. The independent variables, appearing by row, are: age of veteran at time of first exam, number of conditions diagnosed at first exam, and total rate at first exam. The censoring is determined by whether a veteran died during the sampled period. The second column contains the coefficients obtained for the regression. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of total observations included in the regression, the number of censored observations, and the Akaike Information criteria are reported in the last three rows. The third column shows the change in the probability of being sick of any of the 20 conditions, given a unit increase in the correspondent explanatory variable.

	Years sick after first exam (ysickex)	Change in prob of being sick for unit increases in
Age first exam (ageex)	0.067 (0.001)*	0.069
Number of diseases first exam (#disex)	-0.038 (0.006)*	-0.037
Total rate first exam	0.669 (0.100)*	0.952
Number of Observations	14,327	
Censored Observations	3,408	
AIC	191,471.81	

- \* Significant at 1%
- \*\* Significant at 5%
- \*\*\* Significant at 10%

**Table 19**

The columns of this table show the results of a proportional hazard regression where the dependent variable is the number of years each veteran was sick of each particular disease classified by the Surgeons' Certificates after first having been diagnosed with that disease. Of the 20 diseases in the Surgeons' Certificates the coefficients reported here are only those for the conditions for which we obtained reliable results. The independent variables, appearing by row, are: age of veteran at time of first diagnosis of each disease; number of conditions a veteran was diagnosed with at that time; and the specific rate given at time of first diagnosis. The censoring is determined by whether a veteran died during the sampled period. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of total observations included in the regression, the number of censored observations and the Akaike Information criteria are reported in the last three rows.

	Years sick after first diagnosis (vsickdiag)																				
	Cardio	Diarrh.	Ear	Endoc.	Eyes	Gastro	Gen	App	Genito	U	Hernia	Infect.	Injury	Liver	Spleen	Neopl.	Nerv.	Rec. Hem	Resp.	Rheum.	V. Veins
Age first diagnosis (agediag)	0.085 (0.002)*	0.085 (0.003)*	0.084 (0.003)*	0.063 (0.020)*	0.082 (0.003)*	0.086 (0.004)*	0.088 (0.003)*	0.085 (0.004)*	0.087 (0.003)*	0.071 (0.006)*	0.080 (0.002)*	0.084 (0.005)*	0.058 (0.016)*	0.074 (0.010)*	0.078 (0.003)*	0.090 (0.003)*	0.082 (0.003)*	0.083 (0.002)*	0.085 (0.005)*		
Num diseases first diagnosis (disdiag)	-0.007 (0.007)	0.008 (0.010)	-0.004 (0.011)	0.073 (0.063)	0.014 (0.010)	0.041 (0.014)*	-0.007 (0.012)	0.019 (0.012)	-0.019 (0.010)	0.015 (0.018)	-0.010 (0.009)	0.003 (0.020)	0.089 (0.055)	0.009 (0.033)	0.010 (0.012)	-0.009 (0.009)	-0.007 (0.009)	0.011 (0.006)***	-0.027 (0.015)***		
Rate first diagnosis	1.137 (0.232)*	0.615 (0.395)	0.489 (0.243)**	2.473 (1.447)***	0.351 (0.125)*	0.790 (0.540)	-0.082 (0.170)	1.711 (0.434)*	0.812 (0.521)	0.211 (0.692)	0.361 (0.176)**	1.471 (0.815)***	-1.114 (3.468)	4.260 (0.715)*	0.630 (0.166)*	0.894 (0.421)**	2.541 (0.309)*	0.842 (0.236)*	-0.144 (0.732)		
Number of Observations	6,634	3,278	2,029	88	2,683	1,994	2,912	1,634	2,310	787	4,904	972	109	213	1,862	3,951	3,974	7,860	1,114		
Censored Observations	1,688	841	472	20	670	593	652	384	527	207	1,105	271	39	42	474	1,107	1,031	1,979	207		
AIC	78.479	35.219	20.976	500	28.182	18.959	31.694	16.159	24.145	6.800	56.840	8.507	560	1.486	18.358	42.611	43.759	95.588	10.991		

\* Significant at 1%  
 \*\* Significant at 5%  
 \*\*\* Significant at 10%

**Table 20**

Changes in the probability of being sick longer given a unit increase in the correspondent independent variable, computed from the coefficients in Table 19, are shown in this table. This table is organized as is Table 19, so that each lists the numbers describing the change in probability of living with a particular disease for a longer period. Of the 20 diseases in the Surgeons' Certificates the coefficients reported here are those for the conditions reported in Table 19.

	Change in the prob. of living with the disease for each one-unit increase in																			
	Cardio	Diarrh.	Ear	Endoc.	Eyes	Gastro	Gen	App	Genito U	Hernia	Infect.	Injury	Liver	Spleen	Neopl.	Nerv.	Rec. Hem	Resp.	Rheum.	V. Veins
Age first diagnosis (agediag)	0.088	0.089	0.088	0.065	0.086	0.090	0.092	0.089	0.091	0.074	0.084	0.088	0.060	0.077	0.081	0.094	0.085	0.087	0.088	
Num diseases first diagnosis (disdiag)	-0.007	0.008	-0.004	0.076	0.014	0.042	-0.007	0.019	-0.019	0.015	-0.010	0.003	0.093	0.010	0.010	-0.009	-0.007	0.011	-0.027	
Rate first diagnosis	2.117	0.850	0.631	10.857	0.420	1.203	-0.079	4.532	1.254	0.235	0.435	3.353	-0.672	69.818	0.877	1.446	11.692	1.321	-0.134	

**Table 21**

This table shows the results of a proportional hazard regression where the dependent variable is the number of years each veteran was sick of any of the 20 groups of diseases classified in the Surgeons' Certificates after first having been diagnosed with one of those diseases. The independent variables, appearing by row, are: age of veteran at time of first exam; number of conditions a veteran was diagnosed with from first exam until death; and the difference between the total rate at last exam and the total rate at first exam. The censoring is determined by whether a veteran died from any of the 20 conditions during the sampled period. The second column contains the coefficients obtained for the regression. In parentheses the standard errors are shown, and asterisks indicate the significance level of the correspondent independent variable. The number of total observations included in the regression, the number of censored observations and the Akaike Information criteria are reported in the last three rows. The third column shows the change in probability of being sick of any of the 20 conditions given a unit increase in the correspondent explanatory variable.

	Years sick after first exam (ysickex)	Change in prob of being sick for unit increases in
Age at first exam (ageex)	0.066 (0.001)*	0.069
Total number of disabilities in life (#dis)	-0.041 (0.004)*	-0.040
Difference in total rate	-0.176 (0.073)**	-0.162
Number of Observations	14,327	
Censored Observations	7,246	
AIC	119,379.010	

\* Significant at 1%

\*\* Significant at 5%

\*\*\* Significant at 10%

### Graph 1

The following graphs show the plot of the survival distribution function calculated from Table 17 for each of the following conditions (those of Table 17): cardio, diarrhea, gastro, genito urinary, injury, nervous, rectum hemorrhoids, respiratory and rheumatism.











