Bisphosphonates and the Risk of Fractures
- Survival Analysis Using Patient Data from Swedish National Registers

Jonathan Bergman
Undergraduate Thesis
Department of Statistics
Umeå University
Spring 2016
ABSTRACT

Background Osteoporosis increases the risk of fractures. Bisphosphonates is a group of medications used to treat osteoporosis. Purpose The purpose of this study was to answer the questions: Do bisphosphonates decrease the risk of fractures? Are the effects of bisphosphonates different depending on a patient’s age or sex? Methods Data were collected for patients in the Swedish National Patient Register who got a fracture between 2006 and 2012 and who were at least fifty years of age. Kaplan-Meier curves and Cox regression were used; both were extended for time-varying covariates and multiple events. Results A patient who had received a bisphosphonate by any given time had an estimated 45.1 % (hazard ratio, 0.549; 95 % confidence interval, 0.524-0.574) lower rate of fractures than did a patient who would later receive treatment, given that they were the same age when they entered the study. The effect of bisphosphonate treatment did not vary significantly depending on a patient’s age at study entry (p = .866). A patient’s sex had no significant effect on the rate of fractures after adjusting for his/her age at study entry (p = .142). Discussion It is likely that the patients who received a bisphosphonate suffered osteoporosis to a greater extent than did others. If this had not been taken into account the results would have shown that bisphosphonates were associated with an increased rate of fractures.

POPULAR SCIENCE ABSTRACT

Osteoporosis is a medical condition that increases the risk of fractures. Bisphosphonates is a group of medications used to treat osteoporosis. The purpose of this study was to answer the questions: Do bisphosphonates decrease the risk of fractures? Are the effects of bisphosphonates different depending on a patient’s age or sex? Using a statistical method known as survival analysis, this study analyzed data of patients in the Swedish National Patient Register who got a fracture between 2006 and 2012 and who were at least fifty years of age. The results showed that a patient who had been treated with a bisphosphonate had an estimated 45.1 % lower risk of fractures than did a patient who would later receive treatment, given that they were the same age. The effect of the medication was not affected by a patient’s age or sex. It is likely that the patients who received a bisphosphonate suffered osteoporosis to a greater extent than did others. If this had not been taken into account the results would have shown that bisphosphonates were associated with an increased risk of fractures. It should be noted that the methods used do not allow us to say that bisphosphonates actually caused the decrease in the risk of fractures.
SAMMANFATTNING PÅ SVENSKA

Titel Bisfosfonater och risken för frakturer – överlevnadsanalys på patientdata från svenska nationella register

Bakgrund Benskörhet ökar risken för frakturer. Bisfosfonater är en grupp läkemedel som används vid benskörhet. Syfte Syftet med den här studien var att besvara frågorna: Minskar bisfosfonater risken för frakturer? Är effekten av bisfosfonater annorlunda beroende på vilken ålder eller vilket kön en patient har? Metod Data samlades in för patienter i Socialstyrelsens patientregister som fick en fraktur mellan 2006 och 2012 och som var minst femtio år gamla. Kaplan-Meier kurvor och Coxregression användes. Båda modellerna tog hänsyn till tidsberoende förklaringsvariabler och upprepade händelser. Resultat En patient som hade behandlats med bisfosfonat vid en given tidpunkt hade en 45.1 % (hasardkvot, 0,549; 95 % konfidensintervall, 0,524-0,574) lägre skattad hasard för frakturer än en patient som senare skulle få läkemedlet, givet att de var vid samma ålder vid studiestart. Effekten av bisfosfonat varierade inte signifikant beroende på en patients ålder vid studiestart (p = 0,866). En patients kön hade ingen signifikant effekt på hasarden för frakturer efter att ha justerad för patientens ålder vid studiestart (p = 0,142). Diskussion De patienter som fick bisfosfonat var förmodligen mer bensköra än övriga. Om hänsyn inte hade tagits till detta skulle resultaten ha visat att bisfosfonater var associerade med en ökad hasard för frakturer.
1. INTRODUCTION

Osteoporosis is defined by the World Health Organization as having a bone mineral density 2.5 standard deviations less than the mean of young healthy adults (World Health Organization 1994, 5). It is a common condition that is most frequent among older people due to the fact that bone mineral density decreases at a higher rate after the age of fifty, a process that starts in one’s twenties (Lidia et al. 2010). Both men and women are affected, but women who have gone through menopause are at particularly high risk. Osteoporosis does not have any direct symptoms; instead, it increases the risk of fractures (Favus 2010).

The majority of fractures that older adults in the United States get are linked to osteoporosis (U.S. Department of Health and Human Services 2004, 5). In Europe, one report (Hernlund et al. 2013, 5) estimated that 6 % of men and 21 % women over the age of fifty who live in the European Union will have a fracture caused by osteoporosis at some point in their remaining lifetime. Due to aging populations, they are expected to become more common in Asia (Mital & Lau 2009, 3), Africa, and Latin America as well (World Health Organization 2003, 2).

One of the most important objectives of any osteoporosis treatment is to reduce the risk of fractures (Hosking, Geusens & Rizzoli 2005; Liberman et al. 2006). Bisphosphonates is a group of medications which does this by strengthening weakened bones, which in turn is caused by a slowdown in the activity of osteoclasts (McClung 2000). Osteoclasts are cells in the body that break down bone, while there are other cells, osteoblasts, that create bone (U.S Department of Health and Human Services 2004, 21).

Favus (2010) summarizes three of the most important clinical studies that have been conducted to study the effectiveness of bisphosphonates in reducing the risk of fractures. Each study was experimental and participants received either a bisphosphonate or a placebo. There were 2027, 2458, and 7765 participants in each study, respectively. In two of these, the participants were postmenopausal women with low bone mineral density and who had had at least one fracture of the spine. The third study also examined postmenopausal women with low bone mineral density but only required them to have had a fracture of the spine if their bone mineral density did not meet the definition of osteoporosis. After having followed the patients for three years, each study showed that the patients who had taken a bisphosphonate had a lower risk for various kinds of fractures than did the patients who had taken a placebo.

The present study examined the relationship between bisphosphonates and fractures using data of patients in Swedish national registers. This observational approach made it possible to examine a larger and more diverse group of patients than was done in each of the experimental studies summarized by Favus (2010). Additionally, it made it possible to follow patients for a longer period of time. The research questions were:

- Do bisphosphonates decrease the risk of fractures?
- Are the effects of bisphosphonates different depending on a patient’s age or sex?

The first research question is similar to the ones asked in the three experimental studies, but the approach used to answer it was different. The second question is a relevant addition because, as was highlighted above, both age and sex are important contributors to the risk of developing osteoporosis. The question was feasible as well because the national registers are not limited to patients of a particular age or sex.

The data were analyzed using survival analysis. There were two reasons for this. First, patients were followed for varying lengths of time in the registers because of the way eligibility to the study was defined and because not all patients were alive at the end of the
study. Second, the variables in the national registers are measured continuously and are often dates of events. Bisphosphonate treatment and fractures are among such variables.

This paper proceeds as follows. Section 2 introduces the data and study design. Section 3 is a review of survival analysis. Section 4 outlines the statistical methods that were used. Section 5 presents results. This includes a descriptive analysis of the data as well as an account the process of fitting a Cox regression model. Finally, Section 6 is a discussion of methods and results. It concludes with the main findings of the study.

2. DATA & STUDY DESIGN

Data were collected from three sources. First, the subjects of the study were patients in the Swedish National Patient Register who got a major fracture\footnote{Major fractures were defined as follows using codes from the International Classification of Diseases 10 (ICD 10): neck (S12), ribs/sternum/thoracic spine (S22), lumbar spine/pelvis (S32), shoulder/upper arm (S42), forearm (S52), femur (S72), lower leg/ankle (S82).} as their main diagnosis between January 1, 2006, and December 31, 2012. Additionally, they had to be at least fifty years of age at the time. The variables from this source were date of birth, sex, and dates of doctor’s visits for fractures. Second, information about which subjects filled a prescription for a bisphosphonate and on what date was obtained from the Prescription Register. Third, dates of death were, if applicable, collected from the Cause of Death Register. All of the above data are held by the Swedish National Board of Health and Welfare. A list of the variables can be found in the Appendix.

Figure 2.1 provides a sketch of how the study design can be visualized. Each patient in the data set entered the study when he/she got a fracture sometime between 2006 and 2012. A patient exited the study because of death or because the study ended. During the period of time that a patient was followed (follow-up) he/she may have suffered additional fractures. These additional fractures were the events of interest. It needs to be emphasized that the fracture that caused a patient to enter the study was not regarded as an event of interest. When fractures are counted henceforth, the first fracture will refer to the first after study entry. Time until fractures/study exit was measured in days.

Because the National Patient Register contains information about doctor’s visits and diagnoses and not actual fractures, it can be difficult to distinguish a new fracture from a check-up for a previous one. Therefore, a new fracture was defined as either (i) a doctor’s visit for a different type of fracture than the previous one or (ii) a doctor’s visit at least 180 days after a visit for the same type of fracture as the previous one.

Subjects who had already taken a bisphosphonate before entering the study were excluded. Two subjects were dropped because they were reported to have had a doctor’s visit after their date of death. In sum, this left a data set containing 294,136 patients.
3. THEORY OF SURVIVAL ANALYSIS

Suppose we conduct a study over a period of time. Between the beginning and the end of the study, subjects both enter and exit. A subject enters the study because he/she becomes available or eligible (we defined eligibility before the study started). A subject exits the study because he/she withdraws for some reason, dies, or because the study ends. During the period of time that we follow a subject (follow-up) we record whether or not he/she experiences a particular event that is of interest to us. Next, we wish to describe and explain the extent to which subjects experienced the event as well as the length of time (hours, days, weeks, etc.) it took for subjects to experience the event. How do we go about this?

We could analyze the data using a binary variable indicating whether or not each subject experienced the event of interest (1 = yes, 0 = no). A bar graph might be used to describe the proportion of subjects that experienced the event visually. A regression model such as logistic regression might then be used to explain why some subjects experienced the event and why others did not. There is, however, a problem with this approach: some subjects were followed for a longer period of time than others. This alone could explain why some experienced the event and why others did not; ignoring this may lead to misleading conclusions.

Figure 2.1. Sketch of how the study design can be visualized. The horizontal lines represent the follow-ups of four patients. A patient first entered the study, then possibly got one or more fractures, and finally exited the study due to death or study end.
Alternatively, we could analyze the data using a continuous variable for the time (hours, days, weeks, etc.) from when a subject entered the study until he/she experienced the event. A histogram might be used to describe this time visually. A regression model such as linear regression might then be used to explain time until the event. There is, however, a problem with this approach, and it is similar to the previous one: not all subjects experienced the event and the reason for this may be that they were followed for a shorter period of time. One of two things could be done about subjects that did not experience the event: (i) use the time until they exited the study; or (ii) exclude them completely. The former would make little sense if analyzed as time until the event. Even so, there is value in that information since we do know that the time until event was at least the length of follow-up. For this reason, the latter option would be a waste of data.

Notice that the above scenario is similar to how the present study was designed. A group of methods known as survival analysis offers solutions to the problems of analyzing such data. These methods analyze time until an event, while taking into account the length of time that subjects are followed and the possibility that subjects did not experience the event during this time. Another advantage of survival analysis is that the values of a subject’s variables are allowed to vary over time (Guo 2010, 3-10).

The remainder of this theory section reviews two models for analyzing survival data: (i) the Kaplan-Meier curve, which can be used to describe time until an event visually; and (ii) the Cox regression model, which can be used to explain time until an event. Extensions of these models that allow for time-varying covariates and multiple events are presented as well. The former is necessary for the present study because a patient was untreated until such time as he/she received bisphosphonate treatment. The latter is necessary because a patient may have had more than one fracture over the course of follow-up.

3.1. The Kaplan-Meier Curve: A Descriptive Statistic

Let \( T \) be a random variable denoting time (hours, days, weeks, etc.) from a starting point (such as study entry) until an event occurs. The survivor function is defined as \( S(t) = P(T > t) \), the probability that time \( T \) until the event will exceed a particular time \( t \). Survivor functions that are estimated using the Kaplan-Meier method (1958) and displayed in a graph are known as Kaplan-Meier curves. These curves take the form of step functions that start at a probability of one and decrease over time (Kleinbaum & Klein 2005, 46-50).

Some key concepts need to be laid out before introducing the standard Kaplan-Meier estimator. Every subject in a data set has a value for the time-to-event variable \( T = t \). For a subject that did not experience the event, this will be the time until he/she exited the study. Let the number of distinct times-to-event be \( J \). The set of distinct times-to-event in the data set will probably be fewer than the number of subjects because subjects can share the same time \( t \). Next, order these times from smallest to largest, \( t_{(1)} < t_{(2)} < \ldots < t_{(j)} < \ldots < t_{(J)} \). Let \( n_j \) be the number of subjects that have not experienced the event or exited the study by \( t_{(j)} \) and \( d_j \) be the number of subjects that experienced the event at time \( t_{(j)} \). The Kaplan-Meier estimator is

\[
\hat{S}(t) = \hat{P}(T > t) = \prod_{t_{(j)} \leq t} \left( \frac{n_j-d_j}{n_j} \right) \tag{1}
\]


In words the Kaplan-Meier estimator says: (i) Count the number of subjects that did not experience the event at a given point in time. (ii) Divide by the number of subjects that are
still in the study and who remain free of the event until at least that point in time. This gives
the proportion of subjects that did not experience the event of those who are still at risk. (iii)
Multiply such proportions for all times-to-event in the data set leading up to the desired time \( t \).

Kaplan-Meier curves can be estimated separately and compared for subjects that are
grouped by a categorical covariate. When the value of such a covariate can change over time,
these groups will not be constant. Snapinn, Jiang, and Iglewicz (2005) propose an adjustment
to the standard Kaplan-Meier estimator for the \( k \)th group of subjects that allows for a time-varying grouping variable. It takes the form:

\[
\tilde{S}_k(t) = \tilde{p}_k(T > t) = \prod_{t(j) \leq t} \left( \frac{n'_jk - d'_jk}{n'_jk} \right). \tag{2}
\]

\( n'_jk \) is the number of subjects in the \( k \)th group at time \( t_{(j)} \) that have not exited the study or
experienced the event by that time. \( d'_jk \) is the number of subjects in the \( k \)th group at time \( t_{(j)} \)
that experienced the event at that time. To put it simply: group a subject at each time of event
by the value that the subject’s covariate happens to take at that time; then, do as in (1).

When dealing with multiple events it is necessary for the Kaplan-Meier curve be estimated
one event at a time. Thus, there will be a curve for the probability that time until the first
event is greater than a certain time, a second curve for the probability that time until the
second event is greater than a certain time, and so on (Kleinbaum & Klein 2005, 364-365).

The conditional gap-time method can be used to code a data set to account for multiple
events. This means that time is reset to zero each time a subject experiences the event
(Kleinbaum & Klein 2005, 364-367). To take an example in the context of this study, suppose
a patient gets three fractures: 90, 365, and 730 days after entering the study. The observed
times-to-fracture would be 90, 255, and 365 days. Patients that never suffered a first fracture
are only included in the Kaplan-Meier curve for the first fracture. This applies to subsequent
fractures as well. In other words, the estimated probabilities are conditional on the fact a
patient has had all previous events (fractures).

Combining the two extensions of the Kaplan-Meier estimator to allow for both time-varying covariates and multiple events gives

\[
\tilde{S}_{mk}(t) = \tilde{p}_{mk}(T > t) = \prod_{t(j) \leq t} \left( \frac{n'_{mj_k} - d'_{mj_k}}{n'_{mj_k}} \right). 
\]

This is similar to (2) but estimated for the \( m \)th fracture.

The Kaplan-Meier estimator relies on the assumption that subjects exit the study
independently of their times-to-event \( T \). In other words, subjects who are in the study at any
given time must be representative of subjects who have left it in order for estimates to be
unbiased. The same applies to the Cox regression model presented below (Collett 2003, 4-5).
3.2. The Cox Regression Model

The Cox regression model does not use time \( T \) until the event of interest as the response variable directly. Instead, the hazard function is used:

\[
\lambda(t) = \lim_{h \to 0} \frac{P(t < T < t + h \mid T \geq t)}{h},
\]

where \( h \) and \( \lambda(t) \) are positive numbers (Cox 1972). The hazard function gives the rate of the event of interest at a particular time \( t \), given that the event has not occurred by time \( t \) (Kleinbaum & Klein 2005, 11).

The Cox regression model is

\[
\lambda_i(t; x_{i1}, \ldots, x_{ip}) = \lambda_0(t) \exp\{\beta_1 x_{i1} + \ldots + \beta_p x_{ip}\},
\]

where \( x_{i1}, \ldots, x_{ip} \) are covariates for the \( i \)th subject, \( \beta_1, \ldots, \beta_p \) are coefficients, and \( \lambda_0(t) \) is the hazard when all the explanatory variables are zero (Cox 1972). The factor \( \lambda_0(t) \) is called the baseline hazard and specifies how the hazard is a function of time, while the exponential factor specifies how the hazard is a function of covariates (Hosmer, Lemeshow, & May 2008, 69, 215).

Just as the Kaplan-Meier estimator, the standard Cox model does not handle time-varying covariates or multiple events per subject. A solution to this is to extend it using the counting process method (Andersen & Gill 1982; Therneau & Grambsch 2000, 69-70, 185-186). It is easiest to think of the counting process method as a data set layout; it is explained below with a hypothetical example.

Table 3.1 shows a hypothetical patient in the data set that was used in this study. The patient entered the study and was followed for 1200 days. The patient got a fracture (event) after 90 days and a second one after 730 days. He/she received bisphosphonate treatment 150 days after entering the study.

<table>
<thead>
<tr>
<th>Serial Number</th>
<th>Days of follow-up</th>
<th>Days to 1st Fracture</th>
<th>Days to 2nd Fracture</th>
<th>Days to Bisphosphonate Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1200</td>
<td>90</td>
<td>730</td>
<td>150</td>
</tr>
</tbody>
</table>

The upper panel of Table 3.2 shows how the counting process method changes the layout of the data to allow for multiple events. The subject’s follow-up has now been split into three intervals: study entry to first fracture, first fracture to second fracture, and second fracture to study exit. Each interval gets one observation. There is also a binary variable indicating whether or not each time interval ended in a fracture.

The next step is to incorporate the fact that bisphosphonate treatment is a binary time-varying covariate (1 = Yes, 0 = No). The idea is to split the patient’s observations into time intervals such that treatment takes exactly one value in each interval. This has been done in the lower panel of Table 3.2. Two things have changed. First, the time interval that started at
90 days and ended at 730 days has been split at 150 days. Second, bisphosphonate treatment takes the value zero until 150 days have passed, after which it is one.

Table 3.2. A hypothetical patient in the data set. Days are counted from study entry. In the upper panel the counting process method has been used to allow for multiple events (fractures). In the lower panel the counting process method has been used to allow for multiple events and a time-varying covariate (bisphosphonate treatment).

<table>
<thead>
<tr>
<th>Data Set Layout</th>
<th>Serial Number</th>
<th>Start Time</th>
<th>Stop Time</th>
<th>Fracture (Yes = 1, No = 0)</th>
<th>Days to Bisphosphonate Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Events</td>
<td>1</td>
<td>0</td>
<td>90</td>
<td>1</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>90</td>
<td>730</td>
<td>1</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>730</td>
<td>1200</td>
<td>0</td>
<td>150</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Set Layout</th>
<th>Serial Number</th>
<th>Start Time</th>
<th>Stop Time</th>
<th>Fracture (Yes=1, No=0)</th>
<th>Bisphosphonate Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Events</td>
<td>1</td>
<td>0</td>
<td>90</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>90</td>
<td>150</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&amp; Time-varying Covariate</td>
<td>1</td>
<td>150</td>
<td>730</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>730</td>
<td>1200</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The data set is now set up according to the counting process method. The next question is: how do we change the Cox model in (3)? In the standard Cox model a subject is considered to exit the study when he/she experiences the event. In the extended model, the subject continues to be in the study as long as he/she has additional follow-up. This is incorporated into the model by introducing the indicator \( Y_i(t) \) which is one as long as the \( i^{th} \) subject has remaining follow-up and zero thereafter. For time-varying covariates, we simply state that the explanatory variables in (3) are functions of time \( x(t) \) (Therneau 2005, 39, 185-186). For covariates that are time-fixed we simply specify that \( x(t) = x \) (Hosmer, Lemeshow, & May 2008, 215). The model is now:

\[
\lambda_i(t; x_{i1}(t),..., x_{ip}(t)) = Y_i(t) \lambda_0(t) \exp[\beta_1 x_{i1}(t) + ... + \beta_p x_{ip}(t)].
\]  

(4)

3.3. The Cox Model: Estimation, Interpretation, & Inference

The beta-coefficients in the Cox model are estimated using the partial maximum likelihood method. This is a variant of the maximum likelihood method that has been adjusted because not all subjects experience the event(s). The factor \( \lambda_0(t) \) is not estimated at all. The advantage of this is that it avoids having to assume a functional form for \( \lambda_0(t) \). The disadvantage is that the hazard of a subject cannot be estimated. However, the association between event and covariates can still be found by interpreting beta-coefficients (Kleinbaum & Klein 2005, 94-99, 340).
There are often subjects in a data set that share the same times-to-event. This is, however, not allowed when fitting the Cox model because the partial maximum likelihood method assumes that time is continuous. The Breslow (1974) approximation is a standard method for handling this due to its computational simplicity. The Efron (1977) approximation is another one and is considered to be more reliable (Therneau & Grambsch 2000, 48).

Beta-coefficients are interpreted as relative hazards, known as hazard ratios. Let \( x^* \) and \( x \) be two different values of the same covariate. After dropping the subscripts \( i \) for simplicity, the hazard ratio using (4) is

\[
HR = \frac{Y(t)\lambda_0(t)\exp[\beta x^*(t)]}{Y(t)\lambda_0(t)\exp[\beta x(t)]} = \exp[\beta(x^*(t) - x(t))],
\]

where \( x^*(t) - x(t) \) is the desired change in \( x \) at time \( t \). The interpretation of the hazard ratio is by what factor the hazard in the denominator must be multiplied to get the hazard in the numerator at a particular time. Hence, they are similar to odds ratios in logistic regression. If the covariate is time-fixed there will be a single estimate for all points in time (Kleinbaum & Klein 2005, 32-33, 221, 340).

The standard inferential methods for the Cox model are the Wald (\( Z \)), score and likelihood ratio tests. The likelihood ratio test is considered to be the most reliable, but the three tend to give similar results. The Wald statistic can be used to construct confidence intervals for beta-coefficients; the upper and lower bounds can then be used as exponents to the base \( e \) to get confidence intervals for hazard ratios. When using a quantitative variable where the desired change \( x^*-x \) is not one, confidence intervals can be constructed by:

\[
100(1-\alpha)\%\ CI = \exp[(x^*(t) - x(t))\beta \pm z_{\alpha/2} \left| \frac{(x^*(t) - x(t))}{SE(\beta)} \right|],
\]

where the vertical lines denote the absolute value (Hosmer, Lemeshow, & May 2008, 77-79, 99, 106, 216).

Standard errors need to be estimated by so called robust estimation (Lin & Wei 1989) when the counting process method is used for multiple events. The reasoning behind this is that a subject’s events will be dependent events (Hosmer, Lemeshow, & May 2008, 291). Therneau and Grambsch (2000, 172) write that p-values will be too low unless the Wald and score statistics are estimated using robust standard errors. The likelihood ratio test will give p-values that are too low as well, but the authors offer no explanation as to how it can be adjusted. Nevertheless, it is still used without adjustment by both Hosmer, Lemeshow, and May (2008, 292) and by Kleinbaum and Klein (2005, 346).

### 3.4. The Cox Model: Proportional Hazards Assumption, Time-Dependent Coefficients, & Functional Form of Covariates

It was previously stated that covariates can be time-fixed in a Cox model that allows for time-varying covariates by specifying that \( x(t) = x \). The proportional hazards assumption states that the ratio between two hazards (the hazard ratio) must be constant over time if the covariate is time-fixed. Put differently, we assume that the effect (beta-coefficient) is always the same (Therneau 2000, 127). We can see why the Cox model makes this assumption by examining the hazard ratio more closely when \( x(t) = x \). Let \( x^* \) and \( x \) be two different values of a time-
fixed covariate. After dropping the subscripts $i$ for simplicity, the hazard ratio using model (4) is

$$HR = \frac{Y(t)\lambda_0(t)\exp[\beta x^*]}{Y(t)\lambda_0(t)\exp[\beta x]} = \exp[\beta(x^*-x)];$$

the hazard in the numerator is proportional to the hazard in the denominator at all times. Referring back to (5), we see that this assumption is not made if the covariate is time-varying because time $i$ is not canceled out of the expression.

Whether or not it is reasonable to assume that a hazard ratio is constant over time must be checked. Hosmer, Lemeshow, and May (2008, 179-180, 184) recommend two methods: (i) plotting scaled Schoenfeld residuals (1982) against functions of time $g(t)$ and (ii) using a score test to test for correlation between scaled Schoenfeld residuals and functions of time $g(t)$. The functions of time can be $g(t) = t$ and $g(t) = \ln(t)$, for instance. For the purposes of this paper, it is enough to know that Schoenfeld residuals can be calculated for each covariate individually. If the proportional hazards assumption is met, there should be no correlation between these and the functions of time (Hosmer, Lemeshow, & May 2008, 179-180).

What if a time-fixed covariate has Schoenfeld residuals that are correlated with time? One way of dealing with this is by letting the coefficient $\beta$ depend on time $\beta(t)$. This means that the effect of the covariate on the hazard ratio is allowed to vary over time. Including a time dependent coefficient in a model is equivalent to including an interaction term between a function of time and the covariate (Therneau & Grambsch 2000, 147). To see why, let $xg(t)$ be an interaction term between a time-fixed covariate and a function time. Let $\beta'$ be the coefficient of this interaction term. We have

$$\beta x + \beta' xg(t) = (\beta + \beta' g(t))x,$$

and by defining the first factor on the right hand side as a function of time

$$\beta'(t) = \beta + \beta' g(t),$$

we have

$$(\beta + \beta' g(t))x = \beta'(t)x,$$

where $\beta'(t)$ is the time-dependent coefficient of $x$ (Allison 1995, cited in Guo 2010, 95-96).

Hosmer, Lemeshow, and May (2008, 208) urge caution when including this extension to the Cox model because it adds complexity and a greater risk for overfitting. Scatter plots of Schoenfeld residuals and results of hypothesis tests should agree and be reasonable considering the topic at hand. Therneau and Grambsch (2000, 142) write that often no action needs to be taken even when there is some indication that the proportional hazards assumption is violated.

Finally, the Cox model that allows for multiple events, time-varying covariates, and time-dependent coefficients becomes

$$\lambda_i(t; x_{i1}(t),...,x_{ip}(t)) = Y_i(t)\lambda_0(t)\exp[\beta_1(t)x_{i1}(t)+...+\beta_p(t)x_{ip}(t)].$$

For time-varying covariates, using a time-dependent coefficient will not be the result of finding a violation of the proportional hazards assumption, as was explained above.
According to Allison (1995, cited in Guo 2010, 96) the interaction term can be added to see whether it changes the interpretation of the covariate.

All Cox models presented so far have been linear functions of covariates; adding a covariate means simply adding $\beta x$. Hosmer, Lemeshow, and May (2008, 136) explain that all quantitative covariates must be checked to see whether other functional forms are more appropriate, such as a power function. The fractional polynomial method is one way to do this using only a small number of transformations (Royston & Altman 1994).

A description of how the fractional polynomial method can be applied is as follows. Step (i): Consider the exponents \{-2,-1,-0.5,0,0.5,1,2,3\} for a quantitative variable in the model; \{0\} refers to $\ln(x)$. Step (ii): Fit a Cox model for each transformation. Step (iii): Use a likelihood ratio test with two degrees of freedom (one for the coefficient and one for the transformation) to test whether the model with the greatest likelihood is significantly better than the model with the untransformed variable. If it is, continue to the next step. If it is not, conclude that no transformation is needed. Step (iv): Fit a model for each pair of transformations. When the two transformations are identical, such as $\beta x^2 + \beta x^3$, multiply one of these terms by $\ln(x)$. Step (v): Use a likelihood ratio test with two degrees of freedom to test whether the best two-term model is significantly better than the best single-term model. If it is, add a third term in a similar fashion. Identical transformations are in this case handled by squaring one of the $\ln(x)$, such as $\beta x^2 + \beta x^3 \ln(x) + \beta x^2 (\ln(x))^2$. If it is not significantly better, conclude that no additional transformations are required. Step (vi): Stop adding terms when the likelihood ratio test indicates no significant improvement.

4. STATISTICAL ANALYSIS

The covariates that were used in this study need to be defined thoroughly. “On/off treatment” refers to a time-varying covariate. A patient was on treatment from the day that he/she first filled a prescription for a bisphosphonate. Thus, a patient was assumed to have taken a filled prescription. Also note that a patient who was off treatment may never have received the medication. “Treatment” refers to a time-fixed variable. A patient was treated if he/she filled a prescription for a bisphosphonate at some point during follow-up; a patient was untreated otherwise. “Entry age” refers to a patient’s age at the time when he/she entered the study.

Kaplan-Meier curves were used to examine the relationship between bisphosphonates and fractures visually. The method proposed by Snapinn, Jiang, and Iglewicz (2005) was used to extend the estimator for time-varying covariates; the conditional gap-time method (Kleinbaum & Klein 2005, 364-367) was used to extend it for multiple events.

The Cox (1972) regression model was used to estimate hazard ratios (HRs). The counting process method was used to account for time-varying covariates (Therneau & Grambsch 2000, 69-70, 185-186) and multiple events (Andersen & Gill 1982). The Breslow (1974) method was used to handle ties. The Wald statistic with robust standard errors (Lin & Wei 1989) was used to test the significance of covariates and to give 95% confidence intervals (CIs) for hazard ratios.

Product terms were used to examine interactions between the effect of on/off treatment and other variables. Fractional polynomials (Royston & Altman 1994) were used to test whether different functional forms were necessary for quantitative covariates.

The proportional hazards assumption for time-fixed covariates was checked in two ways: (i) scatter plots of scaled Schoenfeld residuals (1982) against time and $\ln(time)$, and (ii) robust score tests for correlation between scaled Schoenfeld residuals and time and $\ln(time)$. 
Violations of the proportional hazards assumption were dealt with by adding interaction terms between covariates and time or ln(time). These interactions were also considered for on/off treatment, but for theoretical reasons: the effect of a medication may vary over time. In sum, (6) is the Cox regression model that was used in this study.

An analysis of the sensitivity of the final results was conducted to assess whether or not patients exited the study due to death independently of their times-to-fractures. (Recall that this is an assumption of the Kaplan-Meier estimator and the Cox model.) This was done by matching each patient that died with a patient that was alive at the end of the study and who was: (i) the same age (± 5 years), (ii) of the same sex, and (iii) also treated/untreated. Next, the patient that died in each matched pair was dropped and the final Cox model was refitted on the remaining patients.

The reasoning behind the matching procedure was that the patients who died were given counterparts that were alive at the end of the study but similar in other respects. Therefore, the counterparts were expected to have times-to-fractures that were representative of what the deceased patients would have had if they not died. If fitting a Cox model on the counterparts gave a different result compared to when the full data set was used, this was interpreted as evidence of a bias caused by the fact that some patients were not followed until the end of the study.

The full data set was used to estimate Kaplan-Meier curves and other descriptive statistics. Such statistics included means with standard deviations, medians with interquartile ranges, minimum and maximum values, and percentages. For computational reasons, Cox models were fit using a random sample of 29,414 (10 %) patients. The exception to this was the final Cox model which was fitted on the full data set.

All tests of hypotheses were performed at the 5 % level. Stata IC version 14 (StataCorp. 2015) was used for statistical analyses.

5. RESULTS

This section is divided into three subsections: (i) a description of the patients in the study; (ii) a comparison of Kaplan-Meier curves; and (iii) an account of the process of fitting a Cox model.

5.1. Study Patients

Table 5.1 lists characteristics of the subjects. There were 22,161 patients (7 %) that received bisphosphonate treatment. All patients were between 50 and 111 years of age when they entered the study. The majority of patients were women, in particular among the treated. The mean age was greater for women than it was for men. Treated patients were younger than untreated patients on average.

Figure 5.1 shows the percentage distribution of length of follow-up. The average patient was followed for a period of 3.0 years (standard deviation, 2.1 years). As can be seen in Table 5.1, 28.8 % of patients exited the study because of death. The proportion of untreated patients that died during follow-up was about twice that of treated patients.

It took a median of 208 days from study entry for treated patients to receive treatment, as shown in Table 5.1. The timing of bisphosphonate treatment in relation to having fractures was as follows: 82.1 % of treated patients received their treatment before getting any
fractures; 13.8 % after a first but before a second; 3.1 % after a second but before a third; and 1 % at some point after a third fracture.

Table 5.2 shows that most patients did not have any fractures during follow-up, regardless of whether or not they received treatment. A greater percentage of treated patients than untreated patients sustained one to six fractures. The percentage of treated patients that sustained seven or eight fractures was less than that of untreated patients, although this only concerned 15 people.

Table 5.1. Characteristics of treated and untreated patients. NA=Not Applicable.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treated</th>
<th>Untreated</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22,161</td>
<td>271,975</td>
<td>294,136</td>
</tr>
<tr>
<td></td>
<td>(7%)</td>
<td>(93%)</td>
<td>patients</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>89.4</td>
<td>69.2</td>
<td>70.7</td>
</tr>
<tr>
<td>Men</td>
<td>10.6</td>
<td>30.8</td>
<td>29.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age at study entry</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Min-Max</td>
<td>50-101</td>
<td>50-111</td>
<td>50-111</td>
</tr>
<tr>
<td>Mean (Standard Deviation)</td>
<td>72.1 (9.5)</td>
<td>75.2 (11.9)</td>
<td>75.0 (11.8)</td>
</tr>
<tr>
<td>Men</td>
<td>72.7 (9.7)</td>
<td>73.9 (11.8)</td>
<td>73.9 (11.7)</td>
</tr>
<tr>
<td>Women</td>
<td>72.1 (9.5)</td>
<td>75.8 (12.0)</td>
<td>75.5 (11.8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Died during follow-up (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>16.4</td>
<td>29.8</td>
<td>28.8</td>
</tr>
<tr>
<td>No</td>
<td>83.6</td>
<td>70.2</td>
<td>71.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Days from study entry to bisphosphonate treatment</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Min-Max</td>
<td>0-2556</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Median (Interquartile Range)</td>
<td>208 (439)</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Figure 5.1 Percentage distribution of the length of time that subjects were followed (follow-up).

Table 5.2. The count and percentage distributions of fracture for treated patients and untreated patients, respectively.

<table>
<thead>
<tr>
<th>Number of Fractures</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treated</td>
<td>Untreated</td>
</tr>
<tr>
<td>0</td>
<td>15236</td>
<td>224676</td>
</tr>
<tr>
<td>1</td>
<td>4827</td>
<td>36524</td>
</tr>
<tr>
<td>2</td>
<td>1436</td>
<td>8121</td>
</tr>
<tr>
<td>3</td>
<td>455</td>
<td>1928</td>
</tr>
<tr>
<td>4</td>
<td>141</td>
<td>543</td>
</tr>
<tr>
<td>5</td>
<td>48</td>
<td>132</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>22,161</td>
<td>271,975</td>
</tr>
</tbody>
</table>
5.2. Bisphosphonates and Fractures: Kaplan-Meier curves

This subsection compares Kaplan-Meier curves obtained by using different grouping variables. In the order that they are presented, the grouping variables are: (i) treated and untreated; (ii) on/off treatment; and (iii) on/off treatment for treated patients.

Figure 5.2 displays Kaplan-Meier curves for treated and untreated patients, respectively. The vertical axes are estimated probabilities (values of the survivor functions) and the horizontal axes are days (time). The four panels correspond to first, second, third, and fourth fractures. Plots of additional fractures are not provided because few subjects suffered more than four.

The Kaplan-Meier curve for treated patients in the upper left panel of Figure 5.2 lies consistently below that for untreated patients; the estimated probability that time until a first fracture exceeded a particular number of days (or years) from study entry was greater for untreated patients. For example, the estimated probability that it took more than a year from study entry for a patient to get a first fracture was 83.7 % if he/she was treated and 88.9 % if not. A treated patient had an estimated 63.8 % probability of having time until a first fracture that was more than five years. This probability was 73.8 % for an untreated patient.

The Kaplan-Meier curves for second and third fractures in Figure 5.2 are similar to those for a first. The curves for a fourth fracture appear to diverge and then converge. Also note that the curves get lower for each fracture. This suggests that the time between fractures tended to decrease as a patient had more of them.

**Figure 5.2.** Kaplan-Meier curves for patients who were treated with a bisphosphonate at some point during follow-up versus those who were untreated. The four panels are estimates for first, second, third, and fourth fractures.
Figure 5.3 displays Kaplan-Meier curves for patients on and off treatment, respectively. The Kaplan-Meier curves for a first fracture in the top left panel are inseparable; the estimated probability that it took more than a certain number of days (or years) from study entry to get a first fracture was approximately the same regardless of whether a patient was on or off treatment at any particular time. For instance, a patient who had been treated in the first year of follow-up had an estimated 88.8% probability of it taking more than a year to get a first fracture. This estimate was 88.5% for a patient who did not receive treatment by that time or who never did. The estimated probability that it took more than five years for a patient to get a first fracture was 73%, given that he/she was treated by that time, and 72.8% if he/she was not.

Beyond the first fracture, the curves for patients off treatment in Figure 5.3 are higher than those for patients on treatment, suggesting that patients off treatment did better. This difference is more pronounced for third and fourth fractures. For a fourth fracture, the two curves diverge and then converge again.

Plotting Kaplan-Meier curves for treated patients and comparing those who were on treatment to those who were off treatment at any given time gave the plots in Figure 5.4. Note that this means that all patients off treatment eventually received treatment.

The upper left panel in Figure 5.4 presents estimates for time until first fracture. The curve for patients on treatment is higher than that for patients off treatment. For instance, the estimated probability that it took more than a year for a patient to get a first fracture was 88.8%, given that he/she had received a bisphosphonate by then and 80.7% if he/she had not. A patient who was on treatment five years after entering the study had an estimated 73.0% probability that time until a first fracture exceeded five years, as opposed to 34.7% for a patient who was still off treatment at that time.

Figure 5.3. Kaplan-Meier curves for patients who had received a bisphosphonate by a particular time (on treatment) and those who had not yet or never received them (off treatment). The four panels are estimates for first, second, third, and fourth fractures.
The Kaplan-Meier curves for a second fracture in Figure 5.4 are similar to those for a first fracture but the curves appear to be closer together. A patient who had had one fracture during follow-up had an estimated 65.9 % probability that the time until he/she sustained a second fracture exceeded a year, given that he/she was on treatment by then. This estimate was 33.1 % for a patient who had not yet been treated one year after his/her first fracture. A patient who was treated four years after his/her first fracture or earlier had an estimated 52.7 % probability that time until second fracture exceeded four years. The corresponding probability for a patient who had not received treatment by then was 33.1 %.

The curves for third and fourth fractures in Figure 5.4 are different from those for first and second fractures; the curves for patients on treatment are higher than or equal to those for patients off treatment until two to three years have passed from the previous one. After this time, the curves for treated patients lie somewhat higher.

![Kaplan-Meier curves for patients](image)

**Figure 5.4.** Kaplan-Meier curves for patients who had received a bisphosphonate by a particular time (on treatment) and those who had not (off treatment). All of these patients eventually received treatment. The four panels are estimates for first, second, third, and fourth fractures.

### 5.3. Bisphosphonates and Fractures: Fitting and Interpreting a Cox Model

Table 5.3 shows the results of fitting five Cox models, starting with on/off treatment only. On/off treatment (1 = on treatment, 0 = off treatment) had a statistically significant coefficient of .148 (p = .004). Adding a variable for treatment (1 = treated, 0 = untreated) led to a decrease in that coefficient to -.523. The variable treatment had a coefficient of .694 and was significant (p < .001).

Model C in Table 5.3 shows that the effect of on/off treatment did not change by adding sex, given that treatment was also included. Sex was significant (p = .001). Doing the same for entry age (Model D) led to a decrease in the effect of on/off treatment from -.523 to -.553. Age was also significant (p < .001). Including both age and sex led to a coefficient for on/off
treatment that was practically unchanged in comparison to when only age was included. Notice, however, that the effect of sex was halved and became insignificant (p = .142). Hence, sex was dropped from further examination.

Table 5.3. Four fitted Cox models with covariates: on/off treatment (1 = on treatment, 0 = off treatment), treatment (1 = treated, 0 = untreated), sex (1 = man, 2 = woman), and age at study entry.

<table>
<thead>
<tr>
<th>Model</th>
<th>Variable(s)</th>
<th>Coefficient</th>
<th>Robust Standard Error</th>
<th>P-value (Robust Wald Test)</th>
<th>95 % CI (Robust Wald Statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>On/off treatment</td>
<td>.148</td>
<td>.051</td>
<td>.004</td>
<td>.048 - .247</td>
</tr>
<tr>
<td>B</td>
<td>On/off treatment</td>
<td>-.523</td>
<td>.067</td>
<td>&lt;.001</td>
<td>-.654 - -.392</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td>.694</td>
<td>.047</td>
<td>&lt;.001</td>
<td>.602 - .787</td>
</tr>
<tr>
<td>C</td>
<td>On/off treatment</td>
<td>-.523</td>
<td>.067</td>
<td>&lt;.001</td>
<td>-.654 - -.392</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td>.677</td>
<td>.047</td>
<td>&lt;.001</td>
<td>.584 - .77</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>.095</td>
<td>.028</td>
<td>&lt;.001</td>
<td>.039 - .151</td>
</tr>
<tr>
<td>D</td>
<td>On/off treatment</td>
<td>-.553</td>
<td>.067</td>
<td>&lt;.001</td>
<td>-.683 - -.421</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td>.747</td>
<td>.047</td>
<td>&lt;.001</td>
<td>.654 - .840</td>
</tr>
<tr>
<td></td>
<td>Entry age</td>
<td>.019</td>
<td>.001</td>
<td>&lt;.001</td>
<td>.017 - .021</td>
</tr>
<tr>
<td>E</td>
<td>On/off treatment</td>
<td>-.552</td>
<td>.67</td>
<td>&lt;.001</td>
<td>-.684 - -.421</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td>.739</td>
<td>.048</td>
<td>&lt;.001</td>
<td>.645 - .832</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>.042</td>
<td>.028</td>
<td>&lt;.142</td>
<td>-.014 - .097</td>
</tr>
<tr>
<td></td>
<td>Entry age</td>
<td>.019</td>
<td>.001</td>
<td>&lt;.001</td>
<td>-.016 - .021</td>
</tr>
</tbody>
</table>

Continuing with Model D in Table 5.3, entry age is the only continuous covariate. All transformations using a single fractional polynomial term for age gave Cox models with partial likelihoods that were lower than when the variable was untransformed. Hence, there was no evidence that a different functional form was necessary.

An interaction term between on/off treatment and entry age was not significant (p = .866). An interaction term between age and treatment was dropped by the software due to collinearity. It was not possible to include an interaction between on/off treatment and treatment because such a variable would be identical to on/off treatment.

Continuing with Model D in Table 5.3, Figure 5.5 displays plots of scaled Schoenfeld residuals versus time and ln(time) for the variables treatment and entry age, respectively. There appears to be curvature in the residuals for treatment in the two left panels, suggesting that the proportional hazards assumption may be violated. The same is not true for age in the two right panels.

There was significant evidence of correlation between scaled Schoenfeld residuals and the two functions of time for variables entry age and treatment (p < .001 in all cases). Moreover, the plot in Figure 5.5 does not suggest that these test results can be attributed to a few extreme data points. I concluded that the proportional hazards assumption was not violated for age because of the lack of pattern in the plots. The variable treatment was further examined with a time-dependent coefficient.

On/off treatment did not have a significant interaction with time (p = .932) or ln(time) (p = .701), given that treatment and entry age were included in the model. The variable treatment had significant interaction terms regardless of which function of time was used (p < .001 in both cases).
Table 5.4 shows the results of two fitted models, each including an interaction term between the variable treatment and a function of time. The interpretation of the variable treatment differed depending on which function of time was used, as can be seen after taking $e$ to the power of the coefficients. For instance, a year after entering the study, the hazard ratio for treatment was estimated to be 2.17 using $time$ and 2.56 using $\ln(time)$. After four years, the hazard ratios were 3.02 and 3.18, respectively.

Both interaction terms with time were dropped for the variable treatment because the choice of function was arbitrary and affected results. I judged that this arbitrariness was a greater concern than the curvature of the residuals.

![Figure 5.5](image)

**Figure 5.5.** Schoenfeld residuals for treatment (left) and entry age (right) against $time$ (top) and $\ln(time)$ (bottom). 2500 refers to days after study entry (2556 was the longest follow-up in the data set).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>Robust Standard Error</th>
<th>P-value (Robust Wald Test)</th>
<th>95 % CI (Robust Wald Statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>On/off treatment</td>
<td>-.7293</td>
<td>.0873</td>
<td>&lt;.001</td>
<td>-9004</td>
</tr>
<tr>
<td>Treatment</td>
<td>.6665</td>
<td>.0508</td>
<td>&lt;.001</td>
<td>.5670</td>
</tr>
<tr>
<td>Treatment* (t)</td>
<td>.0003</td>
<td>.0001</td>
<td>&lt;.001</td>
<td>.0001</td>
</tr>
<tr>
<td>Entry age</td>
<td>.0188</td>
<td>.0011</td>
<td>&lt;.001</td>
<td>.0167</td>
</tr>
<tr>
<td>On/off treatment</td>
<td>-.8481</td>
<td>.0792</td>
<td>&lt;.001</td>
<td>-1.0032</td>
</tr>
<tr>
<td>Treatment</td>
<td>.0273</td>
<td>.1253</td>
<td>&lt;.001</td>
<td>-.2182</td>
</tr>
<tr>
<td>Treatment* ln(t)</td>
<td>.1550</td>
<td>.0237</td>
<td>.828</td>
<td>.1085</td>
</tr>
<tr>
<td>Entry age</td>
<td>.0188</td>
<td>.0011</td>
<td>&lt;.001</td>
<td>.0166</td>
</tr>
</tbody>
</table>
Model D in Table 5.3 was used as a final model after refitting it on the full data set. The results are shown in Table 5.5. After taking $e$ to the power of each coefficient, the hazard ratios can be interpreted as follows. A patient who had received a bisphosphonate by any given time had an estimated 45.1% (hazard ratio, .549; 95% confidence interval, .524-.574) lower rate of fractures than did a patient who would later receive treatment, given that they were the same age when they entered the study.  

A patient who received a bisphosphonate at some point during follow-up but who had not yet been treated was estimated to have fractures at a rate 2.1 (HR, 2.073; 95% CI, 2.006-2.143) times greater than that of a patient who never received treatment, given that they were the same age when they entered the study.  

A patient that was older than another patient at study entry had a higher estimated rate of fractures, given that the two were identical with respect to being treated or untreated and being on or off treatment. For instance, a ten year difference in entry age was associated with having fractures at a 19.7% (HR, 1.197; 95% CI, 1.185-1.209) higher rate. A twenty-year difference in entry age was associated with having fractures at a 43.3% (HR, 1.433; 95% CI, 1.405-1.462) higher rate. Finally, a thirty year difference in entry age was associated with having fractures at a 71.6% (HR, 1.716; 95% CI, 1.665-1.768) higher rate.

Table 5.5. The final Cox model fitted on the full data set with variables: on/off treatment (1 = on treatment, 0 = off treatment), treatment (1 = treated, 0 = untreated), and age at study entry.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>Robust Standard Error</th>
<th>P-value (Robust Wald Test)</th>
<th>95% CI (Robust Wald Statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>On/off treatment</td>
<td>-.600</td>
<td>.023</td>
<td>&lt;.001</td>
<td>-.646 -.555</td>
</tr>
<tr>
<td>Treatment</td>
<td>.729</td>
<td>.017</td>
<td>&lt;.001</td>
<td>.696 .762</td>
</tr>
<tr>
<td>Entry age</td>
<td>.018</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>.018 .019</td>
</tr>
</tbody>
</table>

In order to check the assumption that patients exit the study independently of their times-to-fractures, the final Cox model was fitted on 68,769 patients that were alive at the end of the study. This led to a drop in the coefficient for on/off treatment by 11.7% to -.670 (HR, .512; 95% CI, .458-.572). The coefficient for treatment dropped less than 1% to .723 (HR, 2.06; 95% CI, 1.896-2.239). The coefficient for entry age was unchanged.

---

2 The hazard ratio compares a patient who was on treatment (on/off treatment=1) to one who was off treatment (on/off treatment=0) and who had identical values for all other variables included in the model. If the hazard ratio is $HR = \frac{\exp[-.600*1+.729*\text{Treatment}]}{\exp[-.600*0+.729*\text{Treatment}]}$, this reasonably means that both patients were treated at some point (treatment=1) because the patient in the numerator could not have been on treatment if treatment=0. 

3 The hazard ratio compares a patient who was treated (treatment=1) to one who was untreated (treatment=0) and who had identical values for all other variables included in the model. If the hazard ratio is $HR = \frac{\exp[+.600*\text{OnOffTreatment}+.729*1]}{\exp[+.600*\text{OnOffTreatment}+.729*0]}$, this reasonably means that the treated patient was off treatment (on/off treatment=0) because the untreated patient could not been untreated if on/off treatment=1.
6. DISCUSSION

All methods that were used provided evidence of a relationship between bisphosphonates and fractures; what was not obvious was whether or not bisphosphonates were associated with a decreased risk of fractures. A simple percentage distribution showed that patients who were treated at some point during follow-up had more fractures than did those who remained untreated. Kaplan-Meier curves for first, second, third, and fourth fractures showed the same pattern because the curves for treated patients lay consistently below those for untreated patients. However, it needs to be emphasized that these results were not expected to provide an assessment of the effect the medication itself. There was both a theoretical and an empirical reason for this: on/off treatment was a time-varying variable and most patients that received the medication did so several months after entering the study.

Kaplan-Meier curves and an unadjusted Cox regression model (Model A in Table 5.3) showed that bisphosphonate treatment was associated with a greater risk of fractures even after accounting for the fact that the variable was time-varying (on/off treatment). It should, however, be added that the Kaplan-Meier curves were practically identical for the two groups with regard to a first fracture. This is interesting and something that the Cox model does not capture, but the positive association holds overall after examining curves for subsequent fractures.

It is perhaps surprising that bisphosphonates were associated with an increased risk of fractures due to the fact that they are treatment for osteoporosis (whose symptom is fractures) and that previous research has shown the opposite. Nevertheless, it was possible to show that patients who were on treatment had a lower rate of fractures by including a variable that indicated whether or not patients received treatment at all (Model B in Table 5.3). Notice that this is exactly what was done when the Kaplan-Meier curves were plotted for treated patients who were on and off treatment. This leads to two questions: Why did treated patients have more fractures? Why was bisphosphonate treatment associated with a lower risk of fractures only after adjusting for whether or not patients were treated at all?

In answering the above two questions, the first thing one should keep in mind is that the medication was not randomly assigned; that is, there was probably a reason why certain patients received prescriptions for bisphosphonates. What could that reason be? Bisphosphonates are used to treat osteoporosis. Therefore, it is not unreasonable to suspect treated patients suffered the condition to a greater extent than did others, which in turn could explain why they had more fractures. Furthermore, if this interpretation is correct, the unadjusted relationship between on/off treatment and fractures would be misleading; it is only after taking into account the increased risk of fractures that treated patients had before receiving the medication that we get a fair depiction of the relationship. This also highlights a weakness of the study: bone mineral density is an important but unavailable variable.

Sex was not significantly associated with the rate of fractures after adjusting for age at study entry. This may have been caused by the fact that age was significantly associated with fractures while that there was an age difference between the sexes. However, this does not coincide with previous research which has shown that women are at particularly high risk for osteoporosis and fractures.

Because sex was not significantly associated with fractures in the Cox model, there was no reason to include an interaction term between that variable a function of time. Thus, there was no empirical evidence that the effect of bisphosphonates varied depending on a patient’s sex, given adjustment for his/her age.

The age of a patient was incorporated into the model using the variable entry age. Entry age was significantly associated with fractures but the interaction term with on/off treatment
was not. This implies that there was no empirical evidence that the effect of bisphosphonates varied depending on a patient’s age.

A word of caution must be voiced with the regard to the interpretation of the variable treatment because it may have violated the proportional hazards assumption. The first two pieces of evidence of this were the curvature in the plots of scaled Schoenfeld residuals and the p-values from the correlation tests. However, it was not clear whether the curvature was enough to violate the assumption even though the p-values were low because this was also the case for the variable entry age whose residuals showed no pattern. A stronger piece of evidence was perhaps the change in effect that occurred after introducing interaction terms with time. These time-dependent coefficients changed the effect of treatment from simply a greater rate of fractures for treated patients to one that increased throughout the course of follow-up relative to that of untreated patients. These interactions were problematic to use, however, because they gave hazard ratios that were different depending on which function of time was used. I considered it to be safer to drop the interaction terms because the choice of function was arbitrary and because time-dependent coefficients increase the risk of overfitting. Furthermore, this decision did no affect the overall conclusion of a positive relationship between bisphosphonate treatment and fractures.

Because 28.8% of patients exited the study due to death, it was necessary to assess whether there was a relationship between death and patients’ times-to-fractures; an additional warning was that death was twice as common among untreated patients. The matching procedure did not lead to any substantial changes in the estimated hazard ratios and, hence, no indication that the results of this study were biased in this respect. Nonetheless, this conclusion was drawn with caution because it relies on the assumption that the matching procedure led to a set of patients that were in fact representative of those who died.

The second way of exiting the study, being alive at the end of it, was not assessed for a violation of the assumption that patients exited independently of their times-to-fractures. According to Collett (2003, 4), this is not problematic as long as the final date of the study was selected before examining the data, which was the case.

With regard to the use of the counting process method to account for multiple fractures, Hosmer, Lemeshow, and May (2008, 296) write that this is an unreliable method because it depends on the unrealistic assumption of independent events. However, Therneau and Grambsch (2000, 68-69, 229) write that it gives reliable results for the overall effect of covariates. Estimating an overall effect of bisphosphonates was in line with the purpose of this study, which spoke in favor of using it.

The Breslow approximation was used to handle ties due to its computational efficiency. The more reliable Efron approximation was tested for models A to E in Table 4.3 and gave identical results, suggesting that the Breslow method was not problematic.

One may point out that adding both on/off treatment and treatment to the same model might lead to collinearity problems because the variables are similar. Hosmer, Lemeshow, and May (2008, 167) write that coefficients can become unrealistic and standard errors can become large if collinearity is extreme. Using that statement as guidance, I judged that collinearity was not a large enough problem to justify removing the variable treatment; standard errors did not stand out as being high and coefficient estimates made more sense after including that variable.

To conclude this discussion, the results showed that bisphosphonates were associated with a decreased risk of fractures after taking into account the fact that patients who received treatment had a greater risk of fractures from the beginning. There was no statistically significant evidence that the effect of bisphosphonates varied depending on a patient’s age or sex. However, an observational study of this kind does not permit a conclusion with regard to
whether the medication actually caused the decrease in the rate of fractures. This was the main drawback of the present study compared to previous ones that have been experimental.
ACKNOWLEDGEMENTS

First of all, I would like to thank Anna Nordström and Peter Nordström for unhesitatingly providing me with a topic for my thesis and for trusting me with their data. Anna Nordström deserves recognition for her kindness in checking up on how I was doing from time to time. Peter Nordström deserves recognition for the time that he took to explain the substantive issue at hand and to provide me with the data set. He has also shown great patience in answering my questions regarding how best to analyze the data, several of which I have asked on multiple occasions. Second of all, I would like to thank my advisor Anders Lundquist for his thoughtful advice with regard to the statistical methods, without which I would not have learned nearly as much as I did. Additionally, I am grateful for his genuine wish to help and for his interest in the project from the beginning. Finally, I would like the Department of Statistics at Umeå University to know that I am very grateful that they let me do this project on my own even though the norm is that undergraduate theses are written in pairs.
REFERENCES


StataCorp. 2015. *Stata Statistical software: Release 14.* College Station, TX: StataCorp LP.


APPENDIX: LIST OF VARIABLES

Variables in the data set that were used in the study. The data were obtained from the Swedish National Board of Health and Welfare. The variables come from the National Patient Register unless otherwise specified. NA= Not Applicable.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial number</td>
<td>Subject Identification number</td>
<td>Meaningless digits</td>
</tr>
<tr>
<td>Birth</td>
<td>Date of birth</td>
<td>DD/MM/YYYY</td>
</tr>
<tr>
<td>Death*</td>
<td>Date of death</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Bisphosphonates**</td>
<td>Date of first filled prescription for bisphosphonates between Jan. 1, 2006, and Dec. 31 2012.</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Sex</td>
<td>Sex</td>
<td>1=male, 2=female</td>
</tr>
<tr>
<td>Visit 1</td>
<td>Date of first doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 2</td>
<td>Date of second doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 3</td>
<td>Date of third doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 4</td>
<td>Date of fourth doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 5</td>
<td>Date of fifth doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 6</td>
<td>Date of sixth doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 7</td>
<td>Date of seventh doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 8</td>
<td>Date of eighth doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 9</td>
<td>Date of ninth doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
<tr>
<td>Visit 10</td>
<td>Date of tenth doctor’s visit between Jan. 1, 2006, and Dec. 31, 2012, where main diagnosis was a major fracture</td>
<td>DD/MM/YYYY, NA</td>
</tr>
</tbody>
</table>

* Obtained from the Cause of Death Register.
** Obtained from the Prescription Register.