Methodology for Treatment Evaluation in Patients With Cancer Metastatic to Bone

Richard J. Cook, Pierre Major

Background: Patients with cancer metastatic to bone experience several adverse and clinically important skeletal-related events, including pathologic fractures, vertebral compressions with fracture, the need for surgery to treat or prevent fractures, and the need for radiation therapy for the treatment of bone pain. We present appropriate methods for describing and modeling the clinical course of skeletal-related events and comparing treatments for such events. Methods: On the basis of data from a recently completed randomized, placebo-controlled trial involving 380 breast cancer patients with bone metastases, we tested the validity of the “events-per-person-years” method, one of the most commonly used techniques, for the analysis of skeletal-related events. We then used more robust methods of analysis that are based on fewer assumptions, including a random-effects Poisson model, and contrasted the inferences about skeletal-related event rates and treatment effects for the different analytic methods. All statistical tests were two-sided. Results: The events-per-person-years analysis underestimated substantially the variation in the data and is not appropriate to summarize the incidence rate of skeletal-related events. A random-effects Poisson model did provide a valid basis for analyzing such data. Conclusions: The underestimation of variability in data associated with the use of the events-per-person-years analysis leads to unduly narrow confidence intervals for complication rates and inflated false-positive error rates in treatment comparisons. A random-effects Poisson model provides a valid, robust basis for describing the clinical course of bone complications and evaluating treatment effects. [J Natl Cancer Inst 2001;93:534–8]
skeletal complication rate and hence better reflects the true clinical course of disease in this heterogeneous population. It has been pointed out that this generalization makes the model much more plausible for characterizing the distribution of recurrent events in medical research (3). Finally, we examined the dependence between the rate of skeletal complications and the length of follow-up by regression modelling ([12]; see the “Statistical Appendix” section). From the results of these analyses, we identified suitable models on which to base estimates of complication rates and treatment effects for studies of therapeutic interventions for patients with bone metastases. In all of the analyses, tests were two-sided and evidence was considered to be statistically significant at the 5% level.

RESULTS

Descriptive Analyses

Three hundred eighty-two patients in the study by Hortobagyi et al. (6) were randomly assigned either to treatment or to placebo. Two patients were excluded from our analyses because they did not have documentation of bone metastases. Three hundred twenty-six patients were alive at study end. Graphic analysis of the skeletal-related event and survival data suggest considerable variation in the rate of skeletal complications between patients. Moreover, patients with longer follow-up, and hence survival times, tend to have lower rates of events than patients with short periods of follow-up. We investigate each of these two features in the analyses that follow.

Test of the Events-per-Person-Years Analysis

The events-per-person-years analysis gave an estimated annual incidence rate of 3.15 skeletal complications for patients in the control arm (95% confidence interval [CI] = 2.90 to 3.41). The validity of this analysis is predicated on the assumption that all patients within the sample experience events at the same rate.

Fig. 1 contains a histogram of the distribution of the individual patient’s event rates. There is considerable variation reflected in the histogram, with approximately 40% of the patients having estimated annual event rates of less than one per year and several with event rates in excess of 10. A formal test of the hypothesis that all patients experience events at the same rate is rejected with very strong evidence (P<.001). The events-per-person-years analysis is justified only if this assumption is true, and so the naive application of this method of analysis will lead to erroneous inferences about treatment effects (see the “Discussion” section).

We next fit a random-effects model (3,11) that accommodates different event rates for different patients and makes patients, rather than the events themselves, the basis of the analysis. The random-effects model gave an estimate of the expected annual rate of bone complications of 3.88 (95% CI = 3.19 to 4.73). Note that this estimate is completely outside the CI arising from the events-per-person-years analysis, and the associated CI is just over three times as wide. The increased uncertainty is appropriate because this analysis addresses the between-patient variation in the event rates.

To illustrate the fit of the random-effects model, we have superimposed the estimated distribution (smooth curve) of the patients’ event rates over the histogram in Fig. 1. The general agreement between the observed and model-based distribution is supportive of the use of the random-effects model over the events-per-person-years analysis that assumes the common annual event rate of 3.15 (Fig. 1).

Test of Independence of Events and Follow-up Duration

We next examine the relationship between the empiric rates of events and the survival by fitting a random-effects Poisson regression model (11) to the skeletal-related events with the log of the time from randomization to death as an explanatory variable (12). Direct implementation of this method was not possible because 54 patients were alive at last contact and hence their survival times were not observed (i.e., the survival times were

![Fig. 1. Distribution of individual control patient’s skeletal-related event rates. This figure contains a histogram of the empiric distribution of the individual patient’s skeletal complication rates in the control group computed as the number of events that each patient experienced divided by the duration of time each was followed. The curve is the random-effects Poisson model-based estimate of the distribution of individual patient’s skeletal complication rates. SER = skeletal-related event.](image-url)

Journal of the National Cancer Institute, Vol. 93, No. 7, April 4, 2001
censored). To address this challenge, a joint model was defined with a random-effects Poisson regression model specified conditionally on the (possibly censored) survival time and a survival model specified for the time to death. The “Statistical Appendix” provides the technical details on how the analysis was carried out. This analysis revealed a highly significant ($P<.001$) inverse relationship between the duration of time on study and the event rate. This result implies that patients who experience skeletal complications at a relatively low rate tend to live longer and, therefore, statements about the rate of events must be made in conjunction with statements about survival.

A test of the hypothesis that this random-effects Poisson regression model provides a good fit to the skeletal-related event data suggests that it is reasonable to base inferences on this model. Diagnostic plots based on Pearson residuals (not shown) also suggest a reasonable fit to the data. We, therefore, consider the inferences drawn from this model in Table 1, where we conclude that the rate of events over the first year of follow-up for patients dying 1 year after randomization is 3.96 (95% CI = 3.23 to 4.66). In contrast, among patients who survived exactly 2 years from randomization, this annual event rate is only 2.61 (95% CI = 2.02 to 3.21).

**Treatment Comparisons**

Having identified a suitable model for the data analysis, we turn our attention to the evaluation of therapeutic interventions. Table 2 reports the estimates of treatment effects under various methods of analysis. The five columns contain the regression coefficient for treatment ($\beta$), the standard error for the $\beta$ coefficient, the relative rate, the 95% CI for the relative rate, and a $P$ value for the test of the null hypothesis of no effect of pamidronate on the incidence of skeletal-related events. Comparison of the standard errors shows the substantial increase in variability that results from use of the random-effects model instead of the events-per-person-years method. Specifically, the standard error for the regression coefficient in the events-per-person-years analysis is only 40% as large as it should be for valid inferences. Again, the CI for the regression coefficient is appropriately more than 2.5 times larger in the random-effects analysis. The $P$ value for testing the treatment effect is statistically significant at .008. Controlling for the survival time gives qualitatively similar inferences about the statistically significant effect of the treatment.

**DISCUSSION**

Methods for describing the pattern of skeletal complications resulting from bone metastases must be based on plausible assumptions to yield clinically relevant information and a rigorous basis for statistical analyses. The importance of appropriate statistical models for drawing reliable conclusions in trials of chronic diseases with recurrent events has been actively debated (7,9,13,14). In our analysis of the data reported by Hortobagyi et al. (5,6), we found that the rate of occurrence of bone complications is highly variable between patients ($P<.001$). The finding is in accord with clinical experience (4) but is in conflict with the assumption underlying the events-per-person-years analysis that all patients within each arm of the study experience skeletal complications at the same rate. This between-patient variation in complication rates means that tests for treatment effects on the basis of the events-per-person-years analysis will feature false-positive error rates potentially much higher than the assumed level (typically 5%). That is, ineffective treatments may generate spurious evidence of a treatment benefit simply because the events-per-person-years method fails to reflect adequately the variation in the complication rates between patients. On the basis of the standard errors of the $\beta$ coefficient in Table 2, even when ignoring the dependence on survival, only approximately 40% of the total variation is accounted for in the events-per-person-years analysis. This result implies that, had an events-per-person-years analysis been performed in a similar population of patients to evaluate an ineffective treatment, there would be an approximately 20% probability of falsely concluding that the treatment was beneficial, a number considerably greater than the typically assumed 5% level. Therefore, a general random-effects model, which accommodates variation in complication rates between different patients, provides close agreement with the experimental data (Fig. 1) and provides a more robust basis for inference about treatment effects.

Another conclusion from our analyses is that the survival time is related to the number of complications. This finding

---

**Table 1. Estimated event rate for the control arm of Hortobagyi et al. (6)**

<table>
<thead>
<tr>
<th>Events/person-years</th>
<th>Yearly event rate</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random-effects model</td>
<td>3.15</td>
<td>2.90 to 3.41</td>
</tr>
<tr>
<td>Adjusted random-effects model</td>
<td>3.88</td>
<td>3.19 to 4.73</td>
</tr>
<tr>
<td>12-mo survival/follow-up</td>
<td>3.96</td>
<td>3.23 to 4.66</td>
</tr>
<tr>
<td>18-mo survival/follow-up</td>
<td>3.10</td>
<td>2.50 to 3.70</td>
</tr>
<tr>
<td>24-mo survival/follow-up</td>
<td>2.61</td>
<td>2.02 to 3.21</td>
</tr>
</tbody>
</table>

*The results reported under the adjusted random-effects model provide estimates of the yearly skeletal complications rates controlling for the survival time of patients, which provides the most appropriate basis for inferences about complication rates. Yearly complication rates for three representative survival times are reported for illustration.

**Table 2. Estimates of treatment effects on the rate of skeletal complications in the trial by Hortobagyi et al. (6)**

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>SE</th>
<th>RR</th>
<th>95% CI</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random-effects model</td>
<td>-0.54</td>
<td>0.07</td>
<td>0.58</td>
<td>0.51 to 0.66</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Adjusted random-effects model</td>
<td>-0.38</td>
<td>0.16</td>
<td>0.68</td>
<td>0.49 to 0.94</td>
<td>.022</td>
</tr>
</tbody>
</table>

*$The values under $\beta$ are the log-relative rates arising from regression models comparing the complication rates in patients randomly assigned to receive pamidronate versus those assigned to receive placebo. SE is the standard error of the log-relative rate. RR is the corresponding relative rate, and the corresponding 95% confidence intervals (CIs) are reported along with associated $P$ values. The results reported under the adjusted random-effects model provide estimates of the treatment effect on the rate controlling for the survival time of patients, which provides the most appropriate basis for assessing treatment effects.

†All $P$ values were two-sided.
is not entirely surprising, since patients who have advanced disease with extensive bone involvement are at higher risk for both skeletal-related events and death. The link between survival and the number of complications and, in particular, the importance of considering these responses jointly have not been discussed, to our knowledge, in the literature. For the purpose of estimating event rates, the implications of this relationship are important as can be seen from Table 1. The analyses we report, which adjust for follow-up duration, ensure that treatment effects are assessed on event rates among patients with comparable times in the study. For example, as shown in Table 2, among patients with comparable survival times, the rate of skeletal complications is reduced by 32% with pamidronate compared with placebo. Analyses of event rates that are dependent on survival times should be accompanied by regular survival analyses to provide a complete understanding of how the treatment effects are manifested.

Both features of the clinical course of bone complications (the highly variable event rates and their relation to survival) must be addressed in the statistical analysis to provide meaningful statements about the rate of bone complications and the effects of treatments on the entire course of disease.

These findings suggest that caution is warranted in the interpretation of trials that have used the events-per-person-years method to test treatment effects in bisphosphonate trials (1, 2). Valid analyses are reported in these studies based on the distribution of the time to the first skeletal-related event, but these analyses failed to show a statistically significant treatment effect (15). Hortobagyi et al. (5, 6) did not report any analyses based on the events-per-person-years method. The straightforward and conservative analysis of the time to the first event in this trial showed that pamidronate could delay the first event. Our analysis also shows that a statistically significant benefit is obtained in terms of reducing the overall rate of skeletal complications over the 24-month period of observation.

**STATISTICAL APPENDIX**

Suppose there are m subjects in a study. Let $T_i$ be the random variable representing the time from randomization to death for patient $i$, and let $t_i$ denote its realized value. Let $N(t_i)$ represent the number of skeletal-related events experienced by patient $i$ over the interval $(0, t_i]$ for the skeletal-related events, where $t_i \leq T_i$.

Let $g(t)$ denote a monotonic function of $t$. For a one-sample problem, we may specify a mixed-time homogeneous Poisson regression model with conditional rate of the form

$$
\lambda_i(s(t_i), u_i) = u_i \lambda_0 \exp[g(t_i)] \quad 0 < t_i \leq T_i
$$

where $u_i$ is a subject-specific random effect that we take to follow a gamma distribution with mean one and variance $\phi$, and $a \otimes b = \min(a,b)$. For given $t_i$, this defines a negative binomial regression model for $N(t_i)$, with $g(t_i)$ as a covariate, where the probability of $n_i$ events over $(0, t_i \leq t_j)$ is given by

$$
p(n_i | (t_i \leq t_j)) = \frac{\Gamma(n_i + \rho)}{\Gamma(n_i) \Gamma(\rho)} (1 - \pi_i)^{-\rho} \pi_i^{n_i},
$$

where $\pi_i = \mu_i \phi/(1 + \mu_i \phi)$ and $\mu_i = \lambda_0 \exp[g(t_i)]$, $i = 1, 2, \ldots, m$, and $\theta = (\lambda_0, \delta, \phi \beta^2) (11)$.

If we observe the time of death for all patients, then the likelihood is simply proportional to the product of terms like [A.2] arising from each subject

$$
L(\theta) \propto \prod_{i=1}^{m} P(N(t_i \leq t_j) = n_i | t_i ; \theta).
$$

Let $f(t; \psi)$ and $\tilde{F}(t; \psi)$ denote the density and survival functions governing the lifetimes of the subjects. If some subjects are censored before they die, then let $\tau_i^*$ denote the corresponding censoring times for survival times where $\tau_i = \tau_i^*$. The likelihood is then constructed as

$$
L(\theta, \psi) = \prod_{i=1}^{m} P(N(t_i \leq t_j) = n_i | t_i ; \theta)
$$

$$
= \left[ f(t; \psi) \right]^{n_i} \left[ \int_0^{\infty} P(N(t_i)) \right]^{1 - n_i} \Delta_i
$$

where $\Delta_i = \int_{t_i}^{\tau_i^*} dt$ is one if patient $i$ is observed to die and $\Delta_i = 0$ otherwise. The likelihood may be maximized by a Newton–Raphson algorithm to give $\hat{\theta}$ and $\hat{\psi}$. Asymptotic variance estimates may be obtained based on the inverse of the observed information matrix.

Note that when $\delta = 0$, this model simplifies to the ordinary random effects Poisson model (the negative binomial model), and when $\phi = \delta = 0$, the model reduces further to the ordinary time-homogeneous Poisson model.

If $x_i$ is a binary treatment indicator where $x_i = 1$ if patient $i$ is on treatment and $x_i = 0$ if they are on control therapy, we may specify a mixed-time homogeneous Poisson regression model for the skeletal related events as

$$
\lambda_i(x_i, t_i ; \theta) = u_i \lambda_0 \exp[\beta_0 + (1 - x_i) \beta_1 g(t_i) + x_i \beta_2 g(t_i)]
$$

where again $u_i$ is gamma distributed with mean 1 and variance $\phi$. Here the dependence of survival time on the event process may be different for the different treatment groups, but the interpretation of treatment effects is simplified considerably if $\delta_0 = \delta_1 = \delta$. A formal test of the hypothesis $H_0: \delta_0 = \delta_1 = \delta$ did not provide any evidence against the null and so we base conclusions on the reduced model with $\delta_0 = \delta_1 = \delta$. In this case, $\beta$ is the log-relative rate of events for a treatment versus a control patient who have the same survival times. The likelihood function can again be constructed as in [A.3] if some survival times are right censored.

**REFERENCES**


NOTES

Editor’s note: The authors have performed research sponsored by Novartis Pharmaceuticals, East Hanover, NJ, including that described in this report. P. Major has served as an expert on hypercalcemia at a meeting between Novartis Pharmaceuticals and the U.S. Food and Drug Administration and has advised the company on the design of clinical trials of bisphosphonates. R. J. Cook has given workshops that were funded by Novartis Pharmaceuticals.

The opinions expressed by P. Major are his own and do not necessarily reflect those of Cancer Care Ontario.

Supported in part by grants from the Canadian Institutes for Health Research (CIHR), the Natural Sciences and Engineering Research Council of Canada, and Novartis Pharmaceuticals.

We acknowledge the collaboration of Novartis Pharmaceuticals personnel Drs. Bee Chen and Das Purkayastha for assistance in the data retrieval. We thank Ms. Lynda Clarke and Ms. Sharon D’Angelo for assistance in the preparation of the manuscript and Ms. Ker-Ai Lee and Mr. Bing-shu Chen for programming assistance. We also thank Dr. John Seaman for his continued support of this research and Professor J. D. Kalbfleisch for his comments. R. J. Cook is a CIHR Investigator.

Manuscript received April 25, 2000; revised January 17, 2001; accepted January 31, 2001.