A recent investigation of hormone replacement therapy and breast cancer risk used a method called “floating absolute risks” (FARs) to compute confidence intervals for relative hazards. This method has been used in other medical studies and has received controversy. This controversy stems from the correct implementation of this method. However, there has been no direct comparison of the FAR method, as it is sometimes incorrectly applied and reported, with the conventional approach for computing confidence intervals from proportional hazards regression. In this paper, the author reports simulation results comparing these two methods and demonstrates that the FAR method, when applied incorrectly, can produce confidence intervals that are substantially too narrow.

Recently, collaborators from the Million Women Study published results from an investigation of hormone replacement therapy and breast cancer risk (1). Relative hazards of incident and fatal breast cancer by hormone replacement therapy use were estimated from proportional hazards regression models. When hormone replacement therapy groups were expressed using more than two groups (e.g., never, past, current users), confidence intervals were computed using a method called “floating absolute risks” (FARs) (2). According to the article, the FAR method does not alter the estimated relative hazards, but it produces standard errors and confidence intervals that enable valid comparisons to be made between any two groups, even if neither is the baseline group.

Easton et al. (2) originally proposed the FAR method and applied it to two studies in which the exposure consisted of more than two categories. The first study (3) was an investigation of prognostic factors in 333 patients with small-cell carcinoma. Easton et al. examined the relation between performance status, defined by the World Health Organization five-point scale, and time to death from lung cancer using proportional hazards regression. The second study (4) was a matched case-control study consisting of 755 matched pairs, in which the association between duration of oral contraceptive use (never, 1–48, 49–96, ≥97 months) and breast cancer risk was investigated. In both studies, Easton and colleagues reported standard errors and confidence intervals using both the FAR method and the conventional approach, and the widths of the FAR confidence intervals were consistently narrower than the corresponding confidence intervals computed using conventional methods. Another interesting feature of the FAR method was that it incorporated sampling variation into the referent group, and a confidence interval for the referent group was reported.

In a 1997 letter, Peto (5) stated that the FAR method should be used in place of the conventional approach for computing confidence intervals. However, in a 1999 commentary, Greenland et al. (6) stated that the FAR method does not yield valid confidence intervals for relative hazards, claiming that it can be proven mathematically that the 95 percent confidence intervals constructed using the FAR method will not cover any parameter with 95 percent
frequency. In addition, they stated that it did not make sense for a fixed, reference relative hazard of 1.0 to have a confidence interval. Furthermore, they stated that conventional confidence intervals are asymptotically efficient and that any intervals which are consistently narrower are invalid. Easton and Peto addressed these criticisms, stating that the claim that the FAR method does not directly produce confidence intervals for relative hazards is “... obviously true—they are not meant to be” (7, p. 393). However, this is how the FAR method has been applied. Easton and Peto also stated that the point of the FAR method is to allow computation of confidence intervals for relative hazards for any pair of categories. It appears that the problem is that the FAR method is not being applied correctly. The standard errors derived from the FAR method, which correspond to “absolute risks,” are applied as if they are the standard errors of the relative hazards. If the FAR method is applied as intended, then one gets the same standard errors as one gets using the conventional approach.

Clearly, there is controversy regarding the FAR method as it continues to be used. Although Greenland et al. analytically addressed the FAR method, there has been no direct comparison of the FAR method, as it is incorrectly applied, with the conventional approach for relative hazards to assess the impact of this. Hence, I conducted a series of simulation studies comparing confidence intervals computed using these two approaches. In the next section, the FAR method is briefly described, followed by simulation study results. After that, the findings are discussed.

FLOATING ABSOLUTE RISK

Easton et al. (2) derived the FAR method for proportional hazards regression and conditional logistic regression, using an easy-to-implement heuristic formulation and a formal derivation of an augmented likelihood. Both derivations yield virtually identical results.

Consider a categorical exposure \( z \) consisting of \( s \) categories. In data analyses, \( z \) typically is expressed as \( z = (z_1, \ldots, z_{s-1}) \) denoting \( (s-1) \) indicator variables; the category not having a designated indicator is the referent group. Under the proportional hazards model, the hazard function at time \( t \) is

\[
\lambda(t; z; \beta) = \lambda_0(t) \exp(\beta' z),
\]

where \( \lambda_0(t) \) is the baseline hazard, and \( \beta = (\beta_1, \ldots, \beta_{s-1}) \) denotes the logarithm of the hazards of each category relative to the referent group. Let \( \hat{\beta}_j \) denote the estimated \( \beta_j \)s \( (j = 1, \ldots, s - 1) \) from the proportional hazards regression. The \( \hat{\beta}_j \)s are, of course, correlated. Easton and colleagues defined independent parameters \( \alpha_0, \ldots, \alpha_{s-1} \) such that \( \hat{\beta}_j = \alpha_j - \alpha_0 \), and they interpreted the \( \alpha \)s as the logarithms of the “absolute risks” associated with each category. Hence, subtraction of \( \alpha_0 \) produces the logarithm of the relative hazards \( \beta_j \)s. However, when the FAR method has been implemented, the variances of the estimated \( \hat{\beta}_j \)s \( (j = 1, \ldots, s - 1) \) are used to compute confidence intervals for relative hazards instead of the variances of \( \hat{\beta}_j - \alpha_0 \). Easton and colleagues developed two methods to estimate the variances of the \( \hat{\beta}_j \)s, a heuristic formulation and an augmented likelihood.

Under the heuristic formulation, the covariance matrix for the \( \hat{\beta}_j \)s is calculated from the covariance matrix for the \( \hat{\beta}_j \)s based on the relation \( \hat{\beta}_j = \alpha_j - \alpha_0 \), that is, \( \text{cov} (\beta_j, \hat{\beta}_j) = \text{cov} (\alpha_j, \alpha_j) - \text{cov} (\alpha_j, \alpha_0) - \text{cov} (\alpha_0, \alpha_j) + \text{cov} (\alpha_0, \alpha_0) (i, j = 1, \ldots, s - 1) \). For a detailed description of this formulation, refer to the article by Easton et al. (2).

The augmented likelihood is a formal derivation of the \( \hat{\beta}_j \)s by adding a term to the log-partial likelihood of the proportional hazards model, which is used to estimate the log relative hazards \( \beta_j \)s. Easton and colleagues stated that the \( \hat{\beta}_j \)s can be used to estimate the \( \hat{\beta}_j \)s and that the variances and covariances are precisely the same as those obtained from the original partial likelihood. For a detailed description of the augmented likelihood, refer to Easton et al. (2).

SIMULATION STUDIES

Extensive simulation studies were conducted to compare the confidence intervals based on the FAR method with those based on the conventional approach. The FAR method was implemented using the heuristic formulation. The exposure \( z \) was a simulated categorical variable, and a number of different configurations of categories were considered. Survival was generated using the hazard function \( \lambda(t; z; \beta) = \exp(\beta' z) \), where the baseline hazard \( \lambda_0(t) = 1 \) and various \( \beta \)s were considered to assess the effect of different relative hazards. Survival was censored such that approximately 25 percent of each sample was censored to reflect real data settings. For each set of simulation parameters, 1,000 samples were generated. For each simulated sample, the data were fit using proportional hazards regression, and \( \hat{\beta} \) as well as the covariance matrices for \( \hat{\beta} \) and \( \hat{\beta}_j \), was recorded. Then, 95 percent confidence intervals were computed using the conventional approach and the FAR method, in which the FAR method was applied as described in the previous section. In addition, the mean standard errors of \( \hat{\beta} \) and \( \hat{\beta}_j \), as well as their standard error estimates (SEEs), were computed from the 1,000 simulated samples. The SEE is the standard deviation of the 95 percent confidence intervals that includes the true relative hazard, were also computed for each method. If a method has proper coverage, then for a 95 percent confidence interval, the coverage should be very close to 0.95.

In one series of studies, I considered three categories and defined \( z = (z_1, z_2) \) in which \( z_1 \) and \( z_2 \) denoted indicator variables. I simulated samples of 50, 100, and 1,000 patients. Different frequency distributions of the exposure were considered (table 1). In particular, I considered categorical exposures in which the referent group was common as well as rare and in which one of the exposure categories was common as well as rare. The log relative hazards, \( \beta = (\beta_1, \beta_2) \), were specified such that the corresponding relative hazards were 1.5 and 2.0. Simulation results are summarized in table 1. The mean standard errors from the conventional approach approximate the SEEs well, and the 95 percent confidence intervals have proper coverage. The mean standard errors from the FAR method are consistently
and substantially smaller than the SEE, and the corresponding confidence intervals have poor coverage. For the FAR method, coverage is particularly poor when the referent group is rare. When two groups other than the referent group are compared, that is, $z_1$ against $z_2$, the FAR and conventional methods are identical (results not reported).

Another series of simulations mimicked the lung cancer study Easton et al. (2) reported in their paper. For the simulation study, I sampled 300 persons, where performance status was simulated to mimic the frequency distribution reported by Easton et al. (table 2). Time to death was generated, with the hazard function mimicking the relative hazards reported by Easton et al. (table 2). Simulation results are summarized in table 2. The mean standard errors from the conventional approach approximate the SEE well, and the 95 percent confidence intervals have proper coverage. The mean standard errors from the FAR method are consistently and substantially smaller than the SEE, and the corresponding confidence intervals have poor coverage.

Another series of simulation studies mimicked the result reported from the hormone replacement therapy and breast cancer analysis of the Million Women Study (1). The time to

<p>| TABLE 2. Simulation results mimicking small cell lung cancer study (Eur J Cancer Clin Oncol 1987; 23:1598–9) (3) |</p>
<table>
<thead>
<tr>
<th>Performance status</th>
<th>Frequency (%)</th>
<th>Relative hazard</th>
<th>SEE*</th>
<th>Conventional</th>
<th>FAR*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean SE*</td>
<td>CP*</td>
</tr>
<tr>
<td>0</td>
<td>13</td>
<td>1.0</td>
<td></td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>51</td>
<td>1.93</td>
<td>0.21</td>
<td>0.20</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>4.15</td>
<td>0.26</td>
<td>0.25</td>
<td>0.95</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>5.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.94</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>5.06</td>
<td>0.60</td>
<td>0.55</td>
<td>0.93</td>
</tr>
</tbody>
</table>

* SEE, standard error estimate; FAR, floating absolute risk; SE, standard error; CP, coverage probability.
invasive and fatal breast cancer and distribution of hormone replacement therapy use were simulated to mimic the results reported in the article, specifically figures 1 and 6. Although the hormone replacement therapy and breast cancer analysis included over 800,000 patients, because of computational considerations, samples of 1,000 patients were generated for the simulation study. Simulation results are summarized in table 3. As observed in the previous simulation studies, the conventional approach produces appropriate standard errors and confidence intervals. For the FAR method applied to invasive breast cancer, the current users have poor coverage. For the other hormone replacement therapy groups, the mean standard errors and coverage probabilities are reasonable. This may be due to a couple of factors, such as a prevalent referent group or relative hazards of former hormone replacement therapy users being close to one. For fatal breast cancer, the conventional analysis produces appropriate results, whereas the FAR method underestimates the standard errors and has poor coverage.

**DISCUSSION**

In this report, simulation studies have demonstrated that confidence intervals for relative hazards based on an incorrect application of the FAR method, which sometimes occurs in practice, can be too narrow and potentially misleading. Conventional analyses provide correct confidence intervals for relative hazards, even when comparing two categories in which neither is the referent group. However, to do so requires the investigators either to provide all of the potential associations of interest in their report or, which is rarely if ever done, to provide the covariance matrix so that the reader can make his/her own calculations. When the FAR method is implemented as designed, the resulting confidence intervals are identical to those obtained from the conventional approach.

The purpose of the FAR method is to allow the reader to compute confidence intervals of any exposure level relative to another, even if neither is the referent group. However, investigators who choose to report confidence intervals based on the FAR method need to exercise great care as to its correct implementation. Failure to do so can lead to potentially misleading results, specifically associations that appear to be statistically significant but in fact are not, which can have a substantial public health impact.

**ACKNOWLEDGMENTS**

The author wishes to thank Dr. Frank Harrell, Dr. Wayne Ray, and Dr. Marie Griffin for their valuable comments that greatly enhanced the quality of this manuscript. Conflict of interest: none declared.

**REFERENCES**