This issue features two abstracts. The first is a mortality abstract giving 25-year follow-up (FU) on 2700 subjects with surgical repair of congenital cardiovascular defect of one of eight different types. Estimated combined exposure was 31,000 person-years with 270 observed deaths. Mean duration of FU was 11.5 years, and overall FU was 94% complete. The second is a combined morbidity/mortality abstract giving 8-year follow-up on 176 subjects with abdominal aortic aneurysm (AAA). Estimated exposure was 825 person-years with nine ruptures and 98 deaths. Mean duration of FU was 4.8 years, with FU 100% complete.

The data upon which each abstract was based thus satisfied the minimum data requirements of Checklist A ("Finding Suitable Articles") and had Grade A or B informational value ("Guidelines for Evaluation of FU Articles"). The first abstract lacked an age/sex distribution, but the authors of the original article had employed age- and sex-adjusted Oregon population data in order to fashion their standardized mortality ratios (SMR's). Hence, Singer could use the reported d and SMR values to derive d' values, analogous to the derivation strategies given in Table I of last month's commentary.

The authors also did not provide interval exposure data. But Singer estimated exposure using a more approximate form of his "Life Table Reconstruction" method. Then armed with his estimated exposure, E, he was able to derive q and q', and arrive at overall EDR's for each type of surgically-corrected defect. When these estimated q values (based on reported d and estimated E), are compared to q values corresponding to the independently provided P values, there is reasonably close agreement. (Technically speaking, this is comparing aggregate [exposure-weighted] q values [i.e., \( \bar{q} = d/E \)] with geometric [compound rate] q values [i.e., \( q = 1 - P^{1/1} \)]. This approach presupposes a fairly regular distribution of losses (d+w) without unusual skewing, and such assumptions may not be justified in some studies.)

Singer's abstract records 159 early deaths (less than 30 days due to perioperative mortality, and 111 late deaths (up to 25 years). Of the 159 persons "lost to follow-up" 57 such withdrawals occur early (in the first 30 days after surgery), with half of them occurring in the PDA group. This observation is unexplained and is somewhat unusual, given the extensive tracing that was pursued.

Singer notes that the mortality ratios found in the study are somewhat lower than they would be if the reference group were insured lives rather than the general population used by the authors. That is to say, the q' in the denominator of the MR would be smaller for a select or ultimate group than for the U.S. or Oregon population. The impairment-specific death rates (EDRs) derived for each defect can, however, be transposed to an insured lives reference group to yield an approximate insured-life MR, in a manner described in Brackenridge.

For the AAA study, age and sex compositional data were not provided directly but were available from a prior citation which detailed the demographic characteristics of the study group. Exposure data were not furnished, but from cumulative incidence (event-rate) curves and information supplied in the article, interval event rates could be determined. As is becoming more commonly the case, the cumulative R curve also gave some \( \bar{w} \) information (\( \theta \) and \( \dot{\theta} \)) permitting some inferences about w. In Singer's mortality abstract, w represented those lost to follow-up (withdrawn alive during study) plus those alive at end of FU (withdrawn alive at end of study). But, as is typical of morbidity abstracts, the target event was one of several potential morbid events (in this case \( n = \) rupture). Hence, w consisted of deaths, plus those lost to follow-up, plus those withdrawn alive at end of study, plus those terminated from follow-up for risk-of-rupture because of some other censoring event (e.g., undergoing graft for elective repair). At the fifth year of FU, of 176 original entrants, 76 were still at risk. Of the 100 no longer in follow-up, 9 had experienced rupture. The "mortality cohort" (a closely similar though apparently not identical group) had experienced 42 deaths by the fifth year of follow-up. So the remainder — perhaps half — of those lost to follow-up would have been untraced (alive at last contact, but fate unknown) or alive-at-end-of-study. While the distribution of w over the five-year...
span is not reported, and thus, exact exposure cannot be calculated, one can still approximate the range of possible exposures by taking the two extreme assumptions — that all withdrawals occurred in the first interval, or that all withdrawals occurred in the last interval. (One could also assume any pattern in between — regular losses, evenly distributed over intervals; a pattern of loss previously seen in studies of similar impairments; and so forth — and arrive at alternative estimates.) The extreme assumptions permitted an estimate of events per 100-person years of exposure.

Truncation

If all studies could follow all their entrants to the study’s end-point (death or defined morbid event), no truncation (abbreviation, curtailment) would occur. But few studies are untruncated. Cases where complete follow-up and complete outcome data might be available would include some single-decrement studies. In such studies there are only target events (d or n) and no withdrawals. If they follow either a rapidly-mortal or rapidly-morbid course, or if they involve the study of a stable population (no migratory losses) by researchers with sufficient patience and longevity (able to follow many years until the final end-points), truncation might also be avoidable. Beginning with an inception cohort (a group with a common time-zero), that is followed over time for outcome-events, would also be a help. Such an inception cohort would be uni-serial (representing one series) instead of multi-serial (e.g., formed over time such as during a rolling entry period, or formed cross-sectionally but without uniform disease severity at time of inception of cohort). An inception cohort also has a good chance of being uni-serial. A uni-serial cohort is one which is studied during a narrow calendar period, reducing the chances that secular changes in therapy, diagnosis or other characteristics might influence the study’s process or outcome. Study group stability and homogeneity can result from attention to these various factors. The Framingham Study is unserial and now in its thirty-third year of study.

Singer’s abstract is a double-decrement study with a rolling entry period that is 30 years wide (1958-1989). The subjects are a nearly "captive" population, in part because they are unizonal (geographically fixed within Oregon) and in part because the record-keeping and tracking system is extensive, diligent, and tightly-linked. Nonetheless, 6% were lost to follow-up (untraced). The rolling entry period resulted in a multi-serial cohort where some subjects shared a common time-zero, some shared a common termination-time (withdrawn alive at end of study) and others entered and left at various times during the study (contributing their individual follow-up durations to overall exposure). Hence, contributing to the first year of follow-up are some who entered (had surgery and first year of survival) in 1958 and some who did not enter until 1989. The cohort is not uni-secular, either. Changes in surgical procedure occurred over the 30 years, and the average age of surgical repair steadily declined in all groups. All of these factors must be borne in mind when evaluating the aggregate data of such truncated cohorts.

The AAA study is a smaller double-decrement study using a multi-serial cohort (aneurysms of different size in persons enrolled at different times) from a captive population. The rolling entry-period is 1951-1984. Part of the problem of this study’s multi-serial character is handled by stratification of the aneurysms into three groups by size on initial exam. However, one form of withdrawal was withdrawing to undergo surgical repair (grafting) and it is not clear that criteria for this decision remained uniform over the secular period. Because the study population was "captive" (most of them from Olmsted County, but with assiduous tracking of out-migrants), FU could be 100% complete. But because of the wide rolling-entry period, many of the subjects were late entrants who were withdrawn alive at the end of study and who contributed most heavily to early follow-up intervals. Thus, while FU was 100% complete, 100% did not complete the FU duration. In fact, although the total time period of the study was 37 years, the average duration of follow-up was only about five years and very few subjects were alive and being followed beyond nine years after their entry.

When The Numbers Don’t Add Up

The articles on which the two abstracts were based originated in the New England Journal of Medicine and the Journal of the American Medical Association — both highly-respected peer-reviewed journals. Yet, due to either typographical errors or editorial lapse, the articles did not reach publication without some of the numbers being in error. For instance, Morris’s article (used by Singer) describes 2701 subjects in the text, yet the tables add up to 2700 subjects. Although the abstract to Nevitt’s article refers to 16 ruptures, the text and tables list only 11. Such discrepancies are typically minor, and it is usually easy to determine which number is the mistaken one. Kita’s abstract was additionally confronted with morbidity data (Nevitt) and mortality data (Ballard) purporting to be on the same AAA cohort, yet differing in $d_0$ and $d_5$ numbers, and perhaps in other ways. And, while Nevitt referred to Melton’s article for compositional information on the morbidity
cohort, Melton's sample is of 296 AAA's identified between 1951 and 1980 while Nevitt's cohort is drawn from an extended sample of 370 AAA's identified between 1951 and 1984. For the purposes of the abstract Melton's demographics were generalized to Nevitt's cohort.

Kaplan-Meier

Each of the two abstracts incorporated Kaplan-Meier survival data, provided by the original authors. This is an increasingly common clinical format for representing survival data. It differs from an actuarial life table method in several respects.

The actuarial method customarily uses annual (one-year) intervals, and groups deaths for subtraction at the end of the interval, regardless of where in the interval a death actually occurs. Having been developed by actuaries for life insurance purposes, and with most life policies being renewed or extended in annual segments on the policy anniversary, the actuarial method met an actuary's mortality accounting needs. Withdrawals were treated as if they occurred evenly, and were tallied, on the average, as if they were experienced at the midpoint of the annual interval. Hence, \( E = \xi - w/2 \), and \( \xi+1 = \xi - (w+d) \). An actuarial survival curve typically has segments of identical (annual) width.

The Kaplan-Meier method is a variable-interval method, because deaths and withdrawals are tallied as they happen. The envelope of the curve more closely approximates the actual survival curve, and accordingly, the Kaplan-Meier curve is also called a product-limit estimator. (It is considered to be the limiting form of an actuarial survival curve.) When Kaplan and Meier developed it, they were looking for a non-parametric way to estimate survival curves from incomplete information (i.e., with losses to follow-up). It makes no assumption about a particular shape for the underlying survival curve or for the corresponding hazard function, and assumes only that censoring occurs non-informatively.10 Much as the cumulative actuarial survival rate \( P \) is the product of the several interval survival rates, \( \pi_n \), the Kaplan-Meier product limit estimator is the product of serial multiplication of conditional probabilities for successive (but variable) intervals.11 In older literature, its mathematical symbol is sometimes a capitalized Greek letter \( \Pi \) (for product), just like the sum of several numbers can be given by a Greek letter \( \Sigma \) (sigma). One convention adopted in the Kaplan-Meier method is that deaths and withdrawals cannot occupy the same time boundary, and when they occur simultaneously, a death is presumed to precede a withdrawal. Exposures derived from a variable-interval or Kaplan-Meier tabulation are generally smaller than those from an actuarial tabulation, as illustrated in Table 402M1-3 of the AAA abstract. Here, the K-M total exposure over ten years was 5% less than the actuarial exposure. Because a withdrawal in the Kaplan-Meier life table does not contribute to the denominator of the survival rate for the interval in which the censoring occurs, a K-M curve can underestimate survival rates compared to actuarial \( \pi_n \)'s (see example in Feinstein,6 pp. 344 and 350). In the AAA abstract, the differences between the two were but a few tenths of a percent, and for a few intervals the K-M survival exceeded the actuarial survival. The main disadvantage of a variable-interval method is that it does not as readily permit the discernment of more clinically familiar survival indices, like median survival, or survival rates at \( 'x' \) years from time zero.

Advanced Mortality Methodology

One of the abstracts in this issue is a product of a prior Advanced Mortality Methodology Workshop. Three such workshops have been held to date, imparting methodologic insights to 20 medical directors and resulting in eight published abstracts to date, with several more pledged. The next workshop is in October 1993, in Toronto, at the conclusion of the joint Academy and CLIMOA meeting. It will run from Wednesday night, October 6th through Friday noon, October 8th. Eight participants can be accommodated, and the time to apply is now. The basic qualifications sought — and an application form — were printed on page 315 of the last issue of this Journal. Spaces are still available, so act now! If the course does not fill with medical directors, participation from qualified underwriters or actuaries will be solicited. Benefits of participation include CME credit, a chance to work closely with one of the oldest living, not to mention most knowledgeable, professors of the subject (the venerable Dr. Singer), and the opportunity to produce a publishable abstract of your own.

References


