Randomized controlled trials of therapies for the primary and secondary prevention of cardiovascular disease increasingly are powered to study overall mortality. The interventions frequently treat asymptomatic risk factors among high-risk patients and have a diverse spectrum of potential adverse effects. Thus, medications such as clofibrate or encainide that successfully alter surrogate endpoints but increase the rate of death are of little clinical interest. Many believe that a new preventive therapy for cardiovascular disease should be used sparingly, if at all, until clinical trials establish equivalence or even superiority to existing treatments in terms of mortality from all causes.

When randomized, controlled trials, for a variety of reasons, fail to provide data on overall mortality, should epidemiologic studies step into the breach? An observational post-marketing study by Wang et al. in this issue of the *Journal* (pages 2335–2341) compares conventional and atypical antipsychotic medications and shows that the former are associated with a 37 percent increased risk of death. Such data on mortality from all causes will be of keen interest to patients and their physicians. However, there are formidable challenges involved in conducting studies of mortality that are clinically relevant and scientifically accurate.

A finding of an increased rate of death from any cause is always important. However, those who design and fund studies should consider whether negative results would be of similar scientific interest. That is, is death from any cause plausibly a sensitive indicator of the risks and benefits associated with a given medication? Mortality is always pertinent with regard to medications (such as lipid-lowering agents) used to treat asymptomatic patients, because the only benefit is prevention of future disease. For medications that have a broad spectrum of potentially serious side effects or that are used to treat patients at high risk of disease, overall mortality is an excellent integrated measure of overall benefit in comparison with risk.

However, an exclusive focus on mortality may obscure important effects of medication, particularly when study findings are negative. For example, in the treatment of gastroesophageal reflux disease with proton-pump inhibitors, it probably would make little sense to base decisions about therapy and policy on negative results with regard to overall mortality in studies comparing these inhibitors with antacids. Among typical patients, the baseline mortality risk would be too low to provide adequate power, and even in high-risk populations, the further loss of power and potential bias introduced by deaths unaffected by these drugs would conceal medication effects. Furthermore, an analysis of overall mortality may contribute little to the elucidation of a biologic basis for a given drug effect. Recognizing that overall mortality often does not tell the full story, the investigators involved in individual clinical trials almost always report other primary end points.

In a randomized, controlled trial, study treatments are initiated at or after the beginning of follow-up. This approach has two important advantages. First, it ensures the identification of adverse events that occur early in therapy, which is often a period of increased risk. For example, in comparisons of medical and surgical therapies, follow-up is begun before surgery to account for perioperative events. Second, prognostic factors can be measured before they are influenced by the treatments. Observational studies can emulate this characteristic of randomized, controlled trials with the inclusion of new user designs, which synchronize the beginning of follow-up with the start of drug use.

Like randomized, controlled trials that study mortality, observational studies usually require very large sample sizes. Thus, most post-marketing studies use data that are already available, such as those that are obtained from automated administrative databases or those from multipurpose studies. Among the minimum requirements for the use of such databases are a defined population for which entries and exits are tracked, information about the occurrence and date of death (and, ideally, the cause), data regarding potential prognostic factors (i.e., confounders) that may differ among the study groups, and ongoing quality control and validation of crucial elements of the data.

The database must include information that permits the tracking of drug use on a day-by-day basis. The capacity to define the
point at which medication use begins is essential. Many drug effects can occur only when the patient is using the drug, and thus both the database and the analysis must account for medication exposure that changes with time. For this reason, cohort studies that collect information on an annual basis (or even less frequently) often are not suitable for post-marketing studies. For case-control studies, which are often recommended for rare events, obtaining accurate, unbiased information about medication used by persons who have since died and by those in comparable control groups is challenging.

Controlling for confounding variables is particularly difficult in studies involving overall mortality. Effects of medication on overall death rates may be small, because most drugs are unlikely to affect all causes of death equally. For example, coxibs may increase the risk of death from cardiovascular causes but decrease the risk associated with colorectal cancer. Yet it is precisely in the study of small effects that the influence of confounding is most difficult to rule out.

Illnesses that precede death may alter drug use and introduce confounding. For example, patients with life-threatening diseases often receive hypnotic therapy, which could lead to a noncausal positive association of such treatments with overall mortality. Conversely, the use of nonsteroidal antiinflammatory drugs may be avoided in patients with serious illnesses, leading to a negative association.3

To avoid this confounding, one might consider emulating the intention-to-treat analysis of a randomized, controlled trial. Follow-up would begin at a defined point in the patient’s clinical course, and drug exposure would be determined at this time and remain fixed throughout follow-up. However, unlike randomized, controlled trials, observational studies provide no incentives for patients to continue initial therapy. Fixed-exposure studies thus are likely to have substantial misclassification of data regarding exposure, as patients stop and start medications. This misclassification will obscure the true effects of medication and exaggerate the relative effect of residual confounding by baseline differences, such as drug indication and related factors, between drug users and nonusers.

A serious error in fixed-exposure studies is the determination of drug exposure after follow-up has begun. For example, some investigations of the effects on mortality of inhaled corticosteroids among patients with chronic obstructive pulmonary disease initiated follow-up at the time of hospital discharge but classified patients as drug users if they started therapy at any time in the subsequent 90 days. As Suissa notes,4 this approach gives drug users “immortal” person-time (i.e., from hospital discharge to start of drug) and thus substantially overestimates the beneficial effects of corticosteroids on the rate of death.

Death is affected by both the occurrence of disease and the patient’s capacity to withstand the consequences of disease. The latter is influenced by a host of factors that are potentially difficult to measure. For example, the care given to patients who receive new medications may be better (or worse) than that for patients who receive older therapy, and this may affect mortality. The potential influence of these factors is illustrated by an analysis of data from trials in patients with cardiovascular disease demonstrating that among patients assigned to placebo, death rates are lower among those who adhere to therapy. Many believe that this “healthy drug user” effect accounts for at least some of the discrepancies between randomized and nonrandomized studies involving hormone-replacement therapy and vitamin supplements.

Studies comparing drugs that have similar indications may provide the best defense against confounding. However, even in such a case, patients will be assigned to therapies for reasons that are difficult to measure (“channeling”). The meticulous measurement of all factors plausibly related to prognosis is essential. One of the criteria for causality in epidemiologic studies is specificity of effect, because a change of similar magnitude in multiple end points is consistent with channeling and other biases. When possible, the study of individual causes of death or other, more specific end points is desirable.1

The relative effectiveness and long-term safety, including effects on mortality, of many widely used medications are poorly understood. Randomized trials would provide the most reliable data; however, in the absence of material reform of the system for the approval of new drugs, there is little incentive to conduct such trials.5 Nonrandomized studies can provide valuable information, as does the thoughtful study by Wang and colleagues. However, observational studies of overall mortality are particularly susceptible to numerous biases and thus must be conducted with extreme care.
Politically Correct Human Embryonic Stem Cells?

Davor Solter, M.D., Ph.D.

Human embryonic stem cells are currently viewed as a very promising basis for regenerative medicine of the future. However, to be eligible for federal funding in the United States, researchers must work with federally approved human embryonic stem-cell lines — that is, the few lines derived before August 2001. There is a concerted effort and hope among scientists and legislators that federal funding could be extended to cover as yet nonexistent embryonic stem-cell lines if such lines could be derived without destroying a viable human embryo. The authors of two recent studies have suggested that such lines can be derived either from one cell of a cleaving embryo, leaving the remaining embryo to develop normally, or from an “embryo” that is rendered genetically incapable of normal development.

Chung et al. derived mouse embryonic stem-cell lines from single blastomeres of embryos at the eight-cell stage and transferred the remaining seven-cell embryos into surrogate mothers, in which they developed into normal mice. They argue that the same procedure (single-blastomere biopsy) could be applied to human embryos obtained by in vitro fertilization (IVF), thus allowing the derivation of an embryonic stem-cell line concomitant with the normal development of the embryo from which the cell line originated (see Figure 1).

Though theoretically possible, the procedure poses considerable problems that make its use unlikely in humans. Infertile couples resorting to IVF are unlikely to accept the additional risk imposed by both embryo biopsy (reduced probability of success) and the probable need to freeze the embryo while the embryonic stem cells are being obtained. It has been argued that the additional risk to the embryo will be balanced by the benefit of having genetically compatible embryonic stem cells in case of therapeutic need. This benefit can be realized only if the derivation of embryonic stem cells is successful; thus, all embryos that have undergone biopsy would have to be frozen until the results of embryonic stem-cell derivation are known. Institutional review boards would also be unlikely to approve this addition to standard IVF protocols, given the risk involved. For the procedure to be morally and politically justified, every embryo from which embryonic stem cells are derived must be given the chance to develop. How is this to be guaranteed, and who will act as the embryo recipient if, for example, 10...