ITT for Observational Data

Worst of Both Worlds?

Meir J. Stampfer

Abstract: Hernán et al reanalyzed Nurses' Health Study Data on hormone therapy and heart disease, to explore further the apparent discrepancy for those results compared with findings from the Women's Health Initiative Trial. Hernán et al concludes that differences in time since menopause remains the most plausible explanation for the different findings. Part of the analysis employs application of the "intention-to treat" principle to analyze the observational data. This commentary points out some of the weaknesses inherent in that approach, which combines a major limitation of observational studies—lack of randomization—with a common limitation of trials, imperfect adherence to the assigned treatment.

(*Epidemiology* 2008;19: 783–784)

In this issue, Hernán et al provide a reanalysis of Nurses' Health Study (NHS) data to explore reasons for the apparent discrepancy in results for use of estrogen plus progestin postmenopausal hormones and heart disease, which were associated with lower risk, in contrast to the positive association in the Women's Health Initiative (WHI) randomized trial. Their main conclusion was that most of the difference could be attributed to the difference in age distribution at the time of initiation of hormone therapy in these 2 studies. Consistent with prevailing clinical practice, most hormone users in NHS began around menopause, whereas in the WHI, two thirds of the participants were aged 60 years or older at the start. The hypothesis that the time of initiation of hormone therapy affects risk of coronary disease—the timing hypothesis—was explored by Grodstein et al[1] soon after publication of the initial WHI findings. Since then, considerable additional evidence, based on animal and human studies, has accumulated to support this hypothesis, as reviewed by Manson and Bassuk[2] and Mendelsohn and Karas.[3]

The NHS and WHI results have been remarkably concordant for all other clinical outcomes examined, including stroke, pulmonary embolism, breast cancer, and colorectal cancer. In particular, the virtually identical results for stroke, as confirmed by the most recent NHS analysis[4] argue against substantial confounding by lifestyle factors or other variables in NHS analyses. Hernán et al came to a similar conclusion, which is that residual confounding is unlikely to explain the apparent divergent findings.

Consistent evidence suggests a transient increase in risk among women who start hormone therapy years after menopause, in the presence of preexisting atherosclerosis, but not in women who start earlier. Further evidence suggests subsequent protection, reflected by the finding that in the WHI the cumulative incidence curves converged by 8 years. Using the retrospective data in the NHS, Hernán et al show this transient increase in risk among women who started hormone therapy more than 10 years after menopause but not in those who started earlier. This was also examined in a sensitivity analysis in our earlier publication; in agreement with Hernán et al, this transient increase in risk did not have a substantial impact on the overall findings in NHS.

In a novel approach to considering this issue, Hernán et al apply the intention-to-treat (ITT) principle in analyzing the NHS data, to "mimic the design of the randomized trial as closely as possible in the NHS." The main advantage of a randomized trial is, of course, that potential confounding factors, both known and unknown, will tend to be evenly distributed across groups. This advantage is lost if only adherent individuals are considered in the analysis, so the ITT principle is appropriately typically applied to analyze all data regardless of adherence to randomized treatment assignment. However, as a consequence, poor adherence in a randomized trial will tend to yield inaccurate estimates of the efficacy of the treatment under study due to misclassification of exposure. In observational data, there is concern about the potential for confounding, but with repeated assessments, the actual exposure of the individuals can be identified, reducing misclassification. In applying the ITT principle to observational data, those who initiate the exposure (in the present example, hormone use) are treated as though they were in that exposure group, regardless of their actual later behavior—that is, if they stopped such use. Thus, application of the intention-to-treat principle to observational data essentially combines the most important limitations of each study design.

From the [a]Department of Epidemiology, Harvard School of Public Health and [b]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA.
Supported by NIH grants HL080644 and CA87969.
Correspondence: Meir J. Stampfer, Channing Laboratory, 181 Longwood Avenue, Boston, MA 02115. E-mail: mstampfer@hsph.harvard.edu.

ISSN: 1044-3983/08/1906-0783
DOI: 10.1097/EDE.0b013e318188442e

With the resulting misclassification of exposure, it is no surprise that essentially null results emerge from the present ITT analysis. This does not "explain" the apparent discrepancy, it just tells us that substantially misclassified exposure data will tend to yield null results. It is a kind of magical thinking that by using the terminology of randomized trials and calling the 2-year intervals of observation "trials," the advantages of randomization are achieved. To the contrary, the worst features of both designs emerge.

Hernán et al attempt to deal with this misclassification by adjusting the results for adherence by using inverse probability weighting. This makes an analysis that is already complex and difficult to follow even more complicated by requiring additional assumptions and models. The magnitude of this problem is apparent by the finding that the relative risk for 8+ years was 0.87 in Hernán's unadjusted analysis and only changed to 0.85 in the adherence-adjusted analysis, despite the fact that over half of the women were misclassified by that time in the ITT analysis.

One potential use of an ITT analysis might be to evaluate whether the women who discontinue their hormone use are at elevated risk, which could lead to an apparent lower risk for those who continue. However, this can also be evaluated using conventional methods by examining risk in those who discontinue their exposure. This is an interesting attempt to use a novel method for analysis, but one that adds no new insights on the relation of hormone therapy to chronic heart disease. Because of its far greater complexity and "black box" nature, which make it difficult even to track numbers of subjects, it should not be recommended for routine use.

ABOUT THE AUTHOR

MEIR STAMPFER is Professor of Epidemiology and Nutrition at Harvard School of Public Health, where he served as chair of the Department of Epidemiology, from 2000 to 2007. He is also Professor of Medicine, and serves as Associate Director of Channing Laboratory at Brigham and Women's Hospital. In addition to his research in chronic disease epidemiology, he also directs 2 training grants.

REFERENCES

1. Grodstein F, Clarkson TB, Manson JE. Understanding the divergent data on postmenopausal hormone therapy. *N Engl J Med*. 2003;348:645–650.
2. Manson JE, Bassuk SS. Invited commentary: hormone therapy and risk of coronary heart disease—why renew the focus on the early years of menopause? *Am J Epidemiol*. 2007;166:511–517.
3. Mendelsohn ME, Karas RH. HRT and the young at heart. *N Engl J Med*. 2007;356:2639–2641.
4. Grodstein F, Manson JE, Stampfer MJ, Rexrode K. Postmenopausal hormone therapy and stroke: role of time since menopause and age at initiation of hormone therapy. *Arch Intern Med*. 2008;168:861–866.