Invited Commentary

Invited Commentary: Dietary Fiber, Estradiol, and Cholesterol

Emily B. Levitan*

* Correspondence to Dr. Emily B. Levitan, Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Ryals Public Health Building 203K, 1530 3rd Avenue South, Birmingham, AL 35294-0022 (e-mail: elevitan@uab.edu).

Initially submitted August 12, 2002; accepted for publication August 26, 2002.

The limitations of examining mediating factors by adjusting for them in standard regression models have been well-documented in the literature. Although alternative analytic models have been suggested, they are rarely used. In the accompanying article, Mumford et al. (Am J Epidemiol. 2010;173(2):145–156) use marginal structural linear mixed models to determine the association between dietary fiber intake and cholesterol through pathways that do not involve estradiol. Their findings suggest that overall high fiber intake decreases levels of total and low density lipoprotein (LDL) cholesterol and that there are multiple pathways through which fiber can act. The estradiol-mediated pathway seems to lead to increases in total and LDL cholesterol which are more than counterbalanced by pathways leading to decreases in total and LDL cholesterol. In addition to answering a scientifically interesting question, this work provides a concrete example of the use of marginal structural models for examination of direct effects and may serve as a guide to future researchers.

cholesterol; dietary fiber; estradiol; lipoproteins; menstrual cycle

Abbreviations: LDL, low density lipoprotein; MSM, marginal structural model.

As epidemiologists, we often want to know not only whether exposure X causes outcome Y but also whether the exposure causes the outcome independently of known pathways. One frequently used approach is to construct models with and without the potential mediator. The model without the mediator is used to estimate the total effect of exposure, and the model adjusting for the mediator is interpreted as showing the effect of exposure independent of the pathway that includes the mediator. However, this approach is only appropriate under strong assumptions that can be easily ignored in practice (1–3). These assumptions include no interaction between the exposure and the mediator on the scale being assessed and no unmeasured confounding of both the exposure-outcome relation and the mediator-outcome relation (1–3). In addition, the mediator-outcome relation cannot be confused by a factor that is affected by exposure (1). The assumptions necessary for the usual analysis to be valid have been documented in the literature for a number of years, and alternative analytic strategies that require less stringent assumptions have been proposed but rarely used (1, 3, 4).

Mumford et al. (5) set out to determine the total effect of dietary fiber intake on blood lipid levels, as well as the effect of fiber intake not mediated through estradiol. To parse out the effects of fiber independent of estradiol, they took the approach of calculating “controlled direct effects”—terminology which is somewhat confusing but in line with previous work. The term direct does not imply that there are no intermediate steps between fiber intake and blood lipid levels. In fact, the authors suggested that fiber may affect blood lipid levels through bile acid metabolism (5). The effects are direct in the sense that they do not act through other factors explicitly included in the causal model. Because there were no measures of bile acid metabolism in Mumford et al.’s study, this pathway was not analyzed. Controlled in this setting means that the effect of fiber intake independent of estradiol is estimated at specified levels of estradiol (6). The controlled direct effects allow for exploration of heterogeneity in the direct effect of fiber intake by estradiol level. Instead of the usual multivariable-adjusted regression models, the authors used more elaborate modeling that sought to avoid several of the pitfalls encountered.
when trying to analyze the effect of exposure through different pathways. This work is informative on several different levels, including the interesting scientific question, the demonstration of the marginal structural model (MSM) for handling mediators, and the clear discussion of the assumptions on which the analysis rests.

The health effects of fiber have been the subject of epidemiologic and experimental investigations for many years. Mumford et al. present 2 potential pathways through which dietary fiber could affect blood lipid levels: increased bile acids and decreased estradiol (5). In premenopausal women, these mechanisms would be expected to have opposing effects on total and low density lipoprotein (LDL) cholesterol. This topic is particularly suited to the examination of controlled direct effects because of the inconsistencies in previously reported associations between fiber and blood lipid levels and because of the observation that postmenopausal women have larger decreases in total cholesterol in response to fiber than premenopausal women (7). In fact, the authors found that the total effects of high fiber intake (defined as \( \geq 22 \) g/day) on total and LDL cholesterol expressed through estradiol- and non-estradiol-mediated pathways were smaller in magnitude than the effects of fiber expressed through non-estradiol-mediated pathways alone (5). This suggests that the pathway through estradiol results in an increase in total and LDL cholesterol which is more than counterbalanced by other pathways which decrease total and LDL cholesterol. However, because mediators of nonestradiol pathways, such as fecal bile acids, were not measured in this study, the direct effects of fiber controlled for nonestradiol pathways could not be estimated. In addition, the authors found that the effect of high fiber on total cholesterol varied by estradiol level and was greatest in the setting of relatively low estradiol concentrations.

In order to estimate controlled direct effects, the authors used MSMS. These models allowed the authors to control for different sets of confounders of the fiber-lipid relation and the estradiol-lipid relation. Mumford et al.’s use and detailed description of these models grants the reader an opportunity to see the practical application of this type of analysis under conditions that many investigators are familiar with. Although the BioCycle Study, which provided data for the analysis, was well-designed and well-executed, the authors had to contend with missing data, unmeasured confounders, the potential for measurement error, and a moderate sample size (5). In many theoretical treatments of MSMS, the messiness of real studies is not addressed in order to simplify presentation. In addition, much of the previous work on MSMS has been very mathematical. This has permitted a high degree of precision in describing ideas but may be a barrier to adoption of these techniques by a broader audience. Mumford et al. describe their decisions and reasoning and provide a concrete example of the utility of MSMS which could help other investigators implement these models to answer their own questions.

A side benefit of using new methods seems to be an increased focus on the assumptions that underlie the statistics. Because the inverse probability weights are higher for people who are exposed when most people like them are not, MSMS have the potential to give odd cases a high degree of influence. Mumford et al. describe their examination of the weights to rule out this possibility, as well as examination of the exposure and covariate data (5). Additionally, they seem to have done an uncommonly thorough job of assessing the assumptions of no unmeasured confounding and correct model specification—assumptions which MSMS share with standard regression models. They present the results of a sensitivity analysis assessing how much unmeasured confounding by plasma volume would be needed to explain the observed results. They conclude that the associations would have to be implausibly strong to explain away the results. The authors examined many different ways to model fiber intake before deciding on dichotomizing intake at 22 g/day. They state that they tried alternative parameterizations for the models for the weights and the final weighted model, but the results were unchanged, although details of these alternative parameterizations are not provided.

Even an analysis as detailed and thorough as the one presented by Mumford et al. (5) is necessarily an oversimplification. In order to interpret the direct effects controlled for specific estradiol levels, the authors proposed an oral contraceptive intervention. This assumed that oral contraceptive pills would adjust estradiol to set levels across the population, which may not be plausible given individual differences in absorption and metabolism of sex hormones. Oral contraceptives also contain progestogens, which could modify the effect of estradiol on blood lipid levels. Dietary fiber is not a single chemical entity, and different types of fiber may have specific effects on blood lipids (8). The foods which are high in fiber also contain other potentially bioactive substances, the effects of which are difficult to separate from the effects of fiber using statistics. Although dichotomizing fiber in this population seems to have been a reasonable decision, in other populations with higher fiber intake, this simplifying measure may not have been appropriate. The MSM technique requires correctly specifying many models. In addition to the final weighted random-effects model, 2 models for calculating the weights associated with fiber intake and 2 models for calculating the weights associated with estradiol level must all be correctly constructed.

Mumford et al. answer a scientifically interesting question about the effects of fiber intake, demonstrate how weighting schemes can be used to construct MSMS for direct and total effects, and clearly and thoroughly document their modeling assumptions (5). This work is an important addition both to the literature on the impact of fiber on cardiovascular health and to the literature on analyzing the impact of mediators. Hopefully, other investigators will consider the techniques demonstrated in this paper when attempting to separate the influences of different causal pathways.

**ACKNOWLEDGMENTS**

Author affiliation: Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama.

Conflict of interest: none declared.
REFERENCES


