



Research Opportunities and Needs in the Study of Dietary Modification and Cancer Risk Reduction: The Role of Biomarkers^{1,2}

Ross L. Prentice*

Fred Hutchinson Cancer Research Center and University of Washington, Seattle, WA 98109

Expanded Abstract

Overview

Dietary and physical activity modifications have great potential to reduce the risk of obesity and major cancers as well as other chronic diseases such as heart disease and diabetes. However, considerable uncertainty exists concerning the reliability and interpretation of nutritional epidemiology reports as well as concerning the national infrastructure for developing dietary and physical activity hypotheses and interventions that merit full-scale testing. Biomarkers provide important research avenues to address both of these issues; for example, dietary and physical activity biomarkers in conjunction with suitable measurement error methods can be used to calibrate standard assessments that use frequencies, records, or recalls to enhance the reliability of nutritional and physical activity association studies. Changes in proteomic and metabolomic biomarkers can mediate between a dietary or physical activity change and chronic disease risk. As the knowledge base linking these biomarker changes to chronic disease risk develops, it will become possible to develop and initially evaluate interventions in terms of changes in these biomarkers. Research activities under way or under development in the Women's Health Initiative (WHI)³ are described below to illustrate the applicability of some of these approaches.

Core findings and discussion

It has been hypothesized that changes in nutrition and physical activity patterns could reverse the obesity epidemic in the United States and elsewhere and also reduce the risk of such prominent age-related chronic diseases as cancer, cardiovascular disease,

osteoporosis, and diabetes. However, because of nutrient consumption and physical activity assessment issues, it is a challenging task to obtain research results of sufficient credibility to favorably influence a number of ameliorating factors such as individuals' nutrition and physical activity choices; advice given by primary care providers; agricultural policies; food production and processing choices; environmental design; educational choices; and food fortification and regulation activities, as may be required to achieve related public health goals. A recent review and study of nutrition and physical activity research strategies and recommendations (1) recognized these challenges and concluded that conduct of the needed research is a "demanding task that is now becoming scientifically achievable."

Several research designs, including randomized controlled intervention trials and intermediate outcome clinical trials that use high-dimensional genomic, transcriptomic, proteomic, and metabolomic outcomes to identify promising dietary and physical activity hypotheses and interventions could play an important role in an overall nutrition, physical activity, and chronic disease research strategy. However, few full-scale intervention trials can reasonably be conducted for reasons of cost and logistics, and observational study designs, particularly cohort studies, can be expected to continue to provide a mainstay approach to the identification and testing of diet, physical activity, and chronic disease hypotheses. Hence, it is important that cohort studies be carried out in a manner that yields reliable information on these important topics.

The possibility that measurement errors in dietary assessment obscure important diet and disease associations in observational studies was raised some years ago (2,3). Early work to assess the importance of this topic focused mainly on food frequency questionnaires (FFQs) and used another more detailed dietary assessment method as a so-called reference instrument. This focus was appropriate because FFQs have been ubiquitous in nutritional epidemiology in recent decades as a result of their cost and logistical advantages over such other self-report approaches as food records (diaries) or recalls. These studies tended to show positive correlations, in the vicinity of 0.2–0.5, between FFQ and reference instrument measures of absolute nutrient intake and somewhat higher correlations in the vicinity of 0.5 for energy-adjusted nutrient intake. For example, an important "validation" study by Willett et al. (4) compared their semiquantitative food frequency intake estimates to 28 d of food records for 173 female nurses collected over a 1-year period.

¹ Published in a supplement to *The Journal of Nutrition*. Presented as part of the conference "The Use and Misuse of Biomarkers as Indicators of Cancer Risk Reduction Following Dietary Manipulation" held July 12–13, 2005 in Bethesda, MD. This conference was sponsored by the Center for Food Safety and Applied Nutrition (CFSAN), Food and Drug Administration (FDA), Department of Health and Human Services (DHHS); the Office of Dietary Supplements (ODS), National Institutes of Health, DHHS; and the Division of Cancer Prevention (DCP), National Cancer Institute, National Institutes of Health, DHHS. Guest Editors for the supplement publication were Harold E. Seifried, National Cancer Institute, NIH; and Claudine Kavanaugh, CFSAN, FDA. Guest editor disclosure: H.E. Seifried, no relationships to disclose; C. Kavanaugh, no relationships to disclose.

² Author disclosure: no relationships to disclose.

³ Abbreviations used: 24HR, 24-h recall; DLW, doubly labeled water; EPIC, European Prospective Investigation of Cancer; FFQ, food frequency questionnaire; NBS, Nutrition Biomarker Study; OPEN, Observing Protein and Energy Nutrition Study; WHI, Women's Health Initiative.

* To whom correspondence should be addressed. E-mail: rprentic@whi.org.

However, to serve as a suitable reference instrument, consumption estimates should adhere to a classical measurement model in which the assessed value can be represented as the actual intake plus random measurement error that is independent of the actual intake and of other study subject characteristics and, importantly, is independent of the measurement error in the FFQ assessment. Without this independent measurement error criterion, it would be possible that the positive correlations between the FFQ and reference instrument result from correlated measurement errors rather than correlations with actual intake (5).

Good biomarkers are available for total energy and for protein energy consumption, and these have allowed this issue of correlated measurement errors to be evaluated empirically in some important settings. The National Cancer Institute's Observing Protein and Energy Nutrition (OPEN) Study represents the most substantial effort of this type to date. This study (6) involved doubly labeled water (DLW) assessments (7) of energy consumption and urinary nitrogen assessments of protein consumption (8) along with FFQs and 2 24-h recalls (24HR) for 261 men and 223 women in Maryland. Under a plausible measurement model (9), the (log) FFQ assessments were only weakly correlated with actual consumption. Moreover, corresponding estimated correlations would have been considerably higher had the 24HR assessments, rather than the biomarker, been used (inappropriately) as reference instrument, reflecting positively correlated measurement errors between the FFQ and 24HR assessments. Analyses of these data indicate that regression coefficients to relate a dietary factor to disease risk among women would be attenuated by estimated factors of 0.039 for total energy, 0.137 for protein, and 0.316 for protein density by the measurement characteristics of the FFQ.

Positive measurement error correlations arise when a person systematically under- or overreports consumption on differing self-report instruments or on repeat applications of the same instrument. The biomarker studies just described allowed for such systematic bias through the inclusion of a person-specific bias term as an independent random effect in a linear measurement model (9). However, there is now evidence from DLW and nitrogen biomarker studies that both energy and protein underreporting bias may depend on such individual characteristics as body mass (6,10) and social desirability factors (11) and may plausibly depend also on age and ethnicity. Systematic biases of this type no longer simply attenuate the regression coefficient for a dietary factor in a disease risk model but can otherwise distort and even reverse the direction of a relative risk trend. Measurement models are available (12) to accommodate this type of systematic bias by allowing the mean of a person-specific bias term to have a linear regression dependence on factors such as obesity and ethnicity that may relate to the magnitude of an individual's under- or overreporting practice. The potential importance of the measurement properties of available assessment instruments is illustrated by a study (13) of dietary fat and breast cancer in the Norfolk component of EPIC in which total and saturated fat showed a noteworthy association with breast cancer incidence when consumption was assessed using a 7-d food diary, but the association was modest and not statistically significant when consumption was assessed using a FFQ. A related larger study using 4-day food records and FFQs in the comparison group of the WHI dietary modification trial (14) is nearing completion.

The importance of these measurement issues also motivated the WHI research group to expend substantial resources on a nearly completed Nutrition Biomarker Study (NBS) among 543 women (50% intervention, 50% control) in the Dietary Modification component of the WHI Clinical Trial. Calibrated

energy and protein consumption estimates based on DLW and urinary nitrogen, and calibrated estimates for various other nutrients using blood concentration biomarkers, will provide an opportunity for the reliable identification of dietary patterns associated with weight change and disease risk over an average 8.5-y follow-up on the 161,808 women in the WHI cohorts.

Research needs

Research needs in the context of observational studies with calibrated nutrient consumption data include the development of so-called recovery biomarkers (15) for additional nutrients. These biomarkers, typically recovered in urine as a nutrient is expended, plausibly adhere to a classical assessment model with an error that is independent of that for nutrient consumption estimates via self-report. Concentration biomarkers are available for many other nutrients from nutrient concentrations in blood or another body compartment. These biomarkers typically cannot be expected to adhere to a classical measurement model, although their measurement errors are plausibly independent of those for self-report assessments. Additional research is needed on the modeling and use of concentration biomarkers. Human feeding study data may be needed to supplement biomarker and self-report data for the development and application of such models.

There are also important research needs in the area of dietary intervention hypothesis development and initial testing. Observational studies, even when carefully calibrated using nutrient exposure biomarkers, may lack the specificity to distinguish the disease prevention potential of consumption changes for highly correlated nutrients or for complex dietary patterns. The new types of genomic, proteomic, and metabolomic data becoming available have great potential to strengthen the nutritional prevention of cancer enterprise. For example, potential interventions could be studied under controlled human feeding study settings to examine changes in a high-dimensional set of proteomic and metabolomic biomarkers. Then, as the knowledge base develops on the relation of such changes to the risk of cancer and a range of other clinical outcomes, it will become possible to project the health benefits and risks of a dietary pattern change in a much more comprehensive manner than has been possible to date. For example, studies of proteomic changes in relation to postmenopausal hormone therapy, and studies of high-dimensional single nucleotide polymorphisms in relation to the risk of each of breast cancer, coronary heart disease, and stroke, are under way in the WHI. A greatly enhanced research program of this type is needed. Such research could be facilitated by the formation of a multidisciplinary body with representation across health outcomes and research foci (basic, clinical, population) to help stimulate needed research and to advise on preventive interventions that may be ready for full-scale testing in randomized controlled trials (1).

In conclusion, a vigorous nutrition and chronic disease research agenda could usefully include observational studies in well-selected populations with nutrient exposure data calibrated using biomarkers; a substantial intervention development and initial testing research enterprise including small-scale human feeding studies with disease risk biomarkers as outcomes; and an interdisciplinary forum to stimulate needed research and to assess readiness for full-scale dietary intervention trials.

Literature Cited

1. Prentice RL, Willett WC, Greenwald P, Alberts D, Bernstein L, Boyd NF, Byers T, Clinton SK, Fraser G, et al. Nutrition, physical activity and

- chronic disease prevention: research strategies and recommendations. *J Natl Cancer Inst.* 2004;96:1276–87.
2. Prentice RL, Pepe M, Self SG. Dietary fat and breast cancer: a quantitative assessment of the epidemiologic literature and a discussion of methodologic issues. *Cancer Res.* 1989;49:3147–56.
 3. Willett W. *Nutritional Epidemiology*. New York: Oxford University Press; 1990.
 4. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. Reproducibility and validity of a semi-quantitative food frequency questionnaire. *Am J Epidemiol.* 1985;122:51–65.
 5. Schatzkin A, Kipnis V. Could exposure assessment problems give us wrong answers to nutrition and cancer questions? *J Natl Cancer Inst.* 2004;96:1564–5.
 6. Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, Sharbaugh CO, Trabulsi J, Runswick S, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol.* 2003;158:1–13.
 7. Schoeller DA. Measurement of energy expenditure in free-living humans using doubly-labeled water. *J Nutr.* 1988;118:1278–89.
 8. Bingham SA. The use of 24-h urine samples and energy expenditure to validate dietary assessments. *Am J Clin Nutr.* 1994;59:227S–31S.
 9. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, Bingham S, Schoeller DA, Schatzkin A, Carroll RJ. Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol.* 2003;158:14–21.
 10. Heitmann BL, Lissner L. Dietary underreporting by obese individuals: is it specific or non-specific? *BMJ.* 1995;311:986–9.
 11. Hebert JR, Ebeling CB, Matthews CE. Systematic errors in middle-aged women's estimates of energy intake: comparing three self-report measures to total energy expenditure from doubly-labeled water. *Ann Epidemiol.* 2002;12:577–86.
 12. Prentice RL, Sugar E, Wang CY, Neuhausser M, Patterson R. Research strategies and the use of nutrient biomarkers in studies of diet and chronic disease. *Public Health Nutr.* 2002;5:977–84.
 13. Bingham SA, Luben R, Welch A, Wareham N, Khaw KT, Day N. Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet.* 2003;362:212–4.
 14. Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials.* 1998;19:61–109.
 15. Kaaks R, Ferrari P, Ciampi A, Plummer M, Riboli E. Uses and limitations of statistical accounting for random error correlations, in the validation of dietary questionnaire assessments. *Public Health Nutr.* 2002;5:969–76.