Nutrient-Wide Association Studies: Another Road to the Same Destination

Running title: Forman et al.; Nutrient-wide association studies

John P. Forman, MD, MSc1; Walter C. Willett, MD, DrPH2,3

1Renal Division; 2Channing Division of Network Medicine; 3Depts of Nutrition and Epidemiology, Harvard School of Public Health & Dept of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA

Address for Correspondence:
John P. Forman, MD, MSc
Kidney Clinical Research Institute
Renal Division, Department of Medicine
Brigham and Women’s Hospital
41 Avenue Louis Pasteur, Suite 120
Boston, MA 02115
Tel: 617-732-8239
Fax: 617-525-2008
E-mail: jforman@partners.org

Journal Subject Codes: [109] Clinical genetics; [14] Other hypertension

Key words: blood pressure; diet; Editorials; nutrigenomics genetics
Apart from the specific findings of this analysis, the authors encourage the use of what they call a nutrient-wide association study (NWAS), analogous to a genome-wide association study (GWAS), to investigate dietary determinants of health outcomes\(^1\). Similar to the agnostic approach used in a GWAS, the features of their approach include: the inclusion of all available nutrient exposures; non-incorporation of prior knowledge; adjustment of statistical tests for multiple comparisons; division of the primary dataset into training and test subsets; and the use of a confirmatory dataset. The primary objectives are to analyze all possible nutrient-disease hypotheses and to avoid false positive conclusions.

This general approach has proven highly valuable for genetic analyses; a large number of associations with various diseases have been firmly established. However, the analogy with GWAS is only partial. In the genetic studies, prior knowledge has been quite limited, and the number of hypotheses tested has been hundreds of thousands and, more recently, millions. In contrast, in the present analysis of nutrients and blood pressure, much is already known and the number of variables that could be tested was less than 100. In addition, low-quality genetic information is usually discarded using quality control measures prior to testing associations; however, errors in dietary reporting during a 24-hour recall are less identifiable and therefore more likely to inflate the type II error rate. Finally, genotypes are immutable while individuals may change their diet in relation to their health status, which may inflate the type I error rate.

Despite the value of the statistical paradigm commonly used for GWAS, the price of a high level of confidence in statistically significant findings is a large number of false negative results. In genetic studies, the risk of false negative results is increasingly being addressed by the incorporation of prior biological knowledge, such as with pathway analyses using machine kernel logistic regression,\(^2\) or extensive polymorphism analyses in one or a small number of
strong candidate genes. This price of false negatives is reflected in the low-to-moderate power described by Tzoulaki et al. for their analysis, and by the failure of some nutrients with well documented relationships with blood pressure to be confirmed in the training and testing dataset. In contrast to the OMNIHEART trial, for example, inverse associations of vegetable protein and monounsaturated fat with blood pressure were not identified in the present analysis, nor was a positive association with carbohydrate consumption. In addition, an inverse relation between fiber intake and blood pressure was not validated in the current study, although randomized trials appear to show that fiber lowers blood pressure.

On the other hand, the issue of multiple comparisons in exploratory studies such as this one does need to be addressed. This can be done in multiple ways such as the Bonferroni corrections (usually too extreme), the False Discovery Ratio (used in the present study), or by simply reporting the nominal P-values while interpreting the findings appropriately as exploratory and requiring confirmation. Given the differences between GWAS and nutrient-wide analyses described above, the traditional latter approach will usually be appropriate. In addition, the value of splitting the available data into training and testing sets is now a less common practice in genetic studies. For GWAS as well as non-genetic studies, the most powerful use of the data appears to be to analyze the data as a single combined dataset with appropriate adjustment of statistical significance. Confirmation of findings in independent datasets is desirable, and sometimes these are available to authors, but often this is not feasible and is done by others. Fortunately, false positive reports are ultimately corrected by the scientific process, although more careful accounting for multiple comparisons and caution in the interpretation of findings can make this more efficient.

The authors argue that their statistical approach is superior to a single nutrient analysis,
but we believe that there is a place for analyses of single (or a few) nutrients, which may be hypothesis-driven based upon animal studies, mechanistic studies, previous observational studies, and exploratory analyses such as this. It is possible that the agnostic approach used here may be more appropriate when studying a disease for which dietary associations are unknown. However, there is already extensive knowledge of the effects of diet on blood pressure, knowledge derived from studies superior in design than the one currently in question. Consequently, although the methodology employed by the authors may be novel, the findings were not. The authors correctly point out that most of the relationships they identified have already been reported. Although they suggest that the inverse association of folate with blood pressure is poorly studied, it has been shown in three large prospective cohort studies,7,8 as well as in several randomized trials.9-11

The fact that the authors found known risk factors may, at first glance, appear to validate the nutrient-wide association approach as a paradigm for future studies of diet and disease. However, the current study represents a simple application of the approach with limited generalizability: a relatively easy and inexpensive exposure (24-hour recall) was analyzed with a relatively easy to measure and inexpensive continuous outcome (blood pressure) in a large pre-existing cross-sectional study. Even then, the authors had low-to-moderate power and, as noted above, known associations of diet and blood pressure were missed. This was particularly notable in their multivariate analyses as only alcohol intake, urinary calcium, and urinary sodium-to-potassium ratio were independently associated with blood pressure. Attempting to use the nutrient-wide approach to detect dietary associations in randomized trials or prospective studies with dichotomous outcomes (eg, myocardial infarction, cancer) would have even lower statistical power and therefore a higher rate of false negatives.
The more challenging issues in the investigation of diet and health outcomes include improving precision in measuring exposures, accounting for confounding, and the investigation of realistic latency periods, which may be many decades. Although there is more to be learned about the relation of diet to blood pressure, we already know much (http://www.cdc.gov/features/Hypertension/). The greatest challenge is to translate this knowledge into clinical and public health action because our failure to do so contributes enormously to the burden of cardiovascular disease.

**Conflict of Interest Disclosures:** None.

**References:**


Nutrient-Wide Association Studies: Another Road to the Same Destination
John P. Forman and Walter C. Willett

Circulation. published online October 23, 2012;
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2012 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/early/2012/10/23/CIRCULATIONAHA.112.149989

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/