Overview of the State of Biomarkers of Nutrient Intake

Committee on the Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes

Marian L Neuhouser, PhD, RD
Fred Hutchinson Cancer Research Center
Seattle, WA
National Academy of Sciences, Medicine and Engineering Public Workshop
January 9, 2017
Overarching Question

• Can biomarkers inform studies of nutrition and chronic disease risk?
• We would like to have the best possible measure of nutritional exposure to test associations of diet with chronic disease risk
  • Self-report measures of diet are subject to both random and systematic error that distort associations. The difference between the “observed” and the “truth” may be large.
• Therefore, the goal of using [nutritional] biomarkers is that the biomarkers may provide a less biased estimate of exposure and may be closer to the “true exposure” or intake
• Biomarkers may be subject to random error, but less subject to systematic error

More accurate diet-disease modeling
Robust nutrition recommendations for chronic disease prevention
Biomarkers - Definitions

• **Biomarkers**: compounds/molecules from human blood, urine, feces, saliva, hair, skin and tissue (i.e., adipose tissue, organ biopsy material)

• **Biomarkers of nutritional exposure**: (1) compounds/molecules that reflect exposure to nutrients (vitamins, minerals, macronutrients), (2) bioactive compounds (carotenoids, isoflavones), (3) food contaminants (aflatoxin, PAHs, nitrosamines, acrylamide, pesticides)
  - Note: (1) and (2) are collectively called “food substances” in the Options Report. Food contaminants (3) may not be relevant for future DRIs.

• **Biomarkers of surrogate endpoints**: not direct measures of food or nutrient intake, but measures of metabolic or other processes that are influenced by food substrates and predictive of disease risk (i.e., LDL-cholesterol, triglycerides, C-reactive protein and other inflammation markers, adipokines)
  - Can be qualified surrogate disease markers or non-qualified disease markers
Classes of Nutritional Biomarkers

• Recovery biomarkers
  • Measures intake and output; typically “recovered” in urine
  • Quantitative; can be used to assess absolute intake
  • Doubly labeled water (energy), urinary nitrogen, urinary potassium, urinary sodium
    Note: potential issues with the number of measures needed for sodium as well as new data on sequestration of sodium in tissues
  • All are short term (few days to a few weeks)

• Concentration biomarkers
  • Measures concentrations or relative percentages/abundance of food substances in blood, urine and other tissues
  • Most nutritional biomarkers are concentration-type
  • Considerable inter-individual variation due to physiology, BMI and other factors
  • Metabolism and endogenous synthesis present issues for fatty acids, metabolomics, proteomics
    • Short term (days to months depending on the half life and tissue)
    • Cannot be used to assess absolute intake, but useful for ranking

• Future research may identify new classes of biomarkers
Nutritional Biomarkers vs. Self-Report for Chronic Disease Risk Estimation – Important Issues

• **Exposure assessment**: Biomarkers are objective whereas self-report is subject to random (less serious) and systematic (more serious) bias

• **Time course of exposure**: Biomarkers only measure short-term unless multiple repeat measures over a course of years. Self-report can ask about intake in distant past, but subject to memory and recall bias

  *For both biomarkers and self-report it is critical to understand and establish the exposure window and the time course between exposure and disease onset; may be many years*

• **Logistics and study cost**: Blood, urine etc can be expensive to collect, store and assay for the biomarkers. Self-report less expensive, but not without cost or need for trained staff

• **Technical issues**: Many biomarker platforms are well validated with excellent reproducibility. Others have high CVs and between lab/kit variation. Many food and nutrient databases are good, but not excellent; limited foods; imputed nutrients; lack of data on between cultivar variations in nutrient content (or influence of growing conditions on nutrient content); issues with defaults; missing/incomplete data for many bioactives
Nutritional Biomarkers vs. Self-Report for Chronic Disease Risk Estimation – Important Issues

• Data analysis:
  • Recovery biomarkers can be used for absolute risk
  • Concentration biomarkers and self-report typically rely on relative rankings. Quartile or quintile cutpoints vary by study since based on the distribution in the non-cases in the study sample. Some cutpoints may be based on clinical cutpoints [e.g., 25(OH)D] if appropriate
  • Dose-response is possible to compute if Investigators properly compute a linear trend across the quantiles, but need to be careful not to over-interpret linear trend p-values if point estimates are null or have wide 95% CIs. Example:

<table>
<thead>
<tr>
<th>Quartiles of 25(OH)D, nmol/L</th>
<th>P-trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;64.5</td>
<td></td>
</tr>
<tr>
<td>43.6-&lt;64.5</td>
<td></td>
</tr>
<tr>
<td>32.7-43.6</td>
<td></td>
</tr>
<tr>
<td>&lt;32.7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Serum 25(OH)D, nmol/L</th>
<th>OR (95% CI)</th>
<th>P-trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>0.003</td>
</tr>
<tr>
<td>2.76 (1.30,5.89)</td>
<td>1.51 (0.72,3.14)</td>
<td>4.45 (1.96,10.10)</td>
</tr>
</tbody>
</table>

• Nested case-control study; models were multivariate adjusted. OR (95% CI) presented.

ML Neuhouser et al AJE 2012
Nutritional Biomarkers vs. Self-Report for Chronic Disease Risk Estimation

• Data analysis:
  • Dose-response can also be computed if investigators are reasonably assured of a linear relationship between serum concentration/relative abundance and outcome. In this case, compute a linear model where the nutrient is linear (continuous) rather than categorical.
  • Example:
  • In a cohort of breast cancer survivors (the HEAL Study) when total dietary fiber (g/d) was modeled as a continuous variable, each gram/day increase in total fiber was associated with a 3% decrease in serum C-reactive protein (a non-qualified disease marker)
  • Total fiber $\beta$ (95% CI) = -0.029 (-0.049, 0.008)
    • Villasenor et al *Br Cancer Res Treat* 2011
Use of Nutritional Biomarkers

• Are there some novel ways we can use nutritional biomarkers to better understand the relationship of nutrient intake with chronic disease risk?
Objective: To use biomarkers and regression calibration to better understand the relationship of diet with chronic disease risk in the Women’s Health Initiative (WHI). This approach uses the biomarkers to correct or ‘calibrate’ the measurement error in the self report.

We now have biomarker data plus self-report on over 1000 WHI participants from both the Dietary Modification Clinical Trial and the Observational Study. We recently completed a controlled feeding study in further efforts for nutritional biomarker identification.
Biomarker-calibration approach

The biomarkers are measured in a representative subset
Regression Calibration Study Procedures in the Women’s Health Initiative

- Doubly labeled water protocol (total energy expenditure)
- 24-hour urine collection (nitrogen, sodium and potassium)
- Fasting blood draw (vitamins, carotenoids)
- Body weight
- Indirect calorimetry (resting energy expenditure)
- Self-reported assessment of diet, physical activity, smoking
- Other data (age, race/ethnicity, height, education etc) – obtained from WHI master database
Regression Calibration Study Procedures in the Women’s Health Initiative

- We examine the difference between the biomarker (‘truth’) and the self-report (i.e., energy from food frequency questionnaire)

<table>
<thead>
<tr>
<th></th>
<th>Intervention</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFQ energy (mean kcal/d)</td>
<td>1379</td>
<td>1505</td>
</tr>
<tr>
<td>DLW energy (mean kcal/d)</td>
<td>2059</td>
<td>2053</td>
</tr>
</tbody>
</table>

Regression Calibration Study Procedures in the Women’s Health Initiative

• Given the substantial under-reporting in energy we need to identify the personal characteristics associated with under-reporting. This will tell us whether or not systematic bias exists.

• We regress the objective biomarker on a set of a priori personal characteristics. The following were significant in the model:

  • WHI Diet Change Arm ↓
  • BMI (higher BMI) ↓
  • Age (older) ↑
  • Black (vs White) ↓
  • Hispanic (vs White) ↓

• Systematic misreporting (note – not all is under-reporting)
Regression Calibration Study Procedures in the Women’s Health Initiative

• We next run a series of simulations using bootstrapping to predict a participant’s “true” energy intake given their age, BMI, WHI intervention arm assignment and race/ethnicity (the variables associated with misreporting).

• This predicted value using the biomarker given age etc etc is then substituted in the diet-disease models instead of the self-report. The regression model is now calibrated.
Biomarker-calibration approach

The biomarkers are measured in a representative subset
APPENDIX TABLE. Estimates of energy intake (kcal/day) obtained by self-reported food frequency questionnaire, a biomarker (total energy expenditure), and a calibrated food frequency questionnaire, according to body mass index category, Women’s Health Initiative Nutritional Biomarkers Study, 2004–2005*  

<table>
<thead>
<tr>
<th>Body mass index † category</th>
<th>Self-reported FFQ ‡</th>
<th>Total energy expenditure</th>
<th>Calibrated FFQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Geometric mean</td>
<td>IQR ‡</td>
<td>Geometric mean</td>
</tr>
<tr>
<td>Normal (＜25.0)</td>
<td>1,407</td>
<td>1,157–1,759</td>
<td>1,894</td>
</tr>
<tr>
<td>Overweight (25.0–29.9)</td>
<td>1,462</td>
<td>1,196–1,837</td>
<td>2,043</td>
</tr>
<tr>
<td>Obese (≥30)</td>
<td>1,454</td>
<td>1,161–1,897</td>
<td>2,213</td>
</tr>
</tbody>
</table>

* Note that the difference between FFQ energy intake (self-report) and total energy expenditure (biomarker) increases as body mass index increases. The biomarker-calibrated estimates, for the same women, correct for the measurement error using the model shown in table 4.

† Weight (kg)/height (m)².
‡ FFQ, food frequency questionnaire; IQR, interquartile range (25th–75th percentiles).
<table>
<thead>
<tr>
<th>Disease outcome</th>
<th>Total energy consumption - uncalibrated</th>
<th>Total energy consumption- calibrated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myocardial infarction</td>
<td>1.0 (0.98,1.03)</td>
<td>1.49 (1.13,1.97)</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td>1.04 (1.01,1.08)</td>
<td>3.51 (2.12,5.82)</td>
</tr>
<tr>
<td>Total stroke</td>
<td>0.97 (0.95,1.0)</td>
<td>1.36 (1.05,1.76)</td>
</tr>
<tr>
<td>Total CVD</td>
<td>0.99 (0.97,1.0)</td>
<td>1.49 (1.18,1.88)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1.06 (1.04,1.07)</td>
<td>4.17 (2.68,6.49)</td>
</tr>
<tr>
<td>Breast cancer (invasive)</td>
<td>1.01 (0.99,1.02)</td>
<td>1.47 (1.18,1.84)</td>
</tr>
<tr>
<td>Colon</td>
<td>1.0 (0.96,1.03)</td>
<td>1.86 (1.18,2.93)</td>
</tr>
<tr>
<td>Total invasive cancer</td>
<td>1.01 (1.0,1.02)</td>
<td>1.43 (1.17,1.73)</td>
</tr>
</tbody>
</table>

Outcomes based on incidence rates in the WHI from baseline through 2010. For example, n=3798 invasive breast cancers and n=6494 diabetes

Zheng et al Am J Epidemiol 2014
Continued Nutritional Biomarker Development

- A useful nutritional biomarker should have good correlation with the nutrient (or ‘food substance’). We are using a benchmark correlation of 0.7 since this approximates the correlation for doubly labeled water (benchmark).

- From recent feeding study:
  - Vitamin B$\text{_{12}}$, alpha carotene >0.70
  - Folate, beta-carotene, lutein+zeaxanthin, alpha-tocopherol 0.6 - 0.7

- We are developing new biomarkers using metabolomics (multiple metabolites)
  - Metabolomics promising for carbohydrate, alcohol, fat (total, specific types)
    - Work in progress

- These approaches may be transportable to other cohorts
  - Freedman et al Am J Epidemiol 2014;

- Biomarkers only needed on a subset when self-report is available on the entire cohort
Research Gaps

• Only a few true “quantitative” biomarkers

• Most concentration or relative abundance biomarkers are influenced by age, sex, BMI, metabolism and physiology. Need a better understanding of how to use concentration/relative abundance biomarkers in a meaningful way that represents “intake”

• Current nutritional biomarkers are measured at one (or a few) points in time and have short half-lives. Need for multiple measures over time and/or biomarkers that reflect long term intake (cumulative risk) – likely relevant for chronic disease risk

• Need to understand the biomarker(s) in relation to the disease susceptibility window

• Efforts to understand the unexplained variance in serum nutrient concentrations

• Continued search for biomarkers for certain macronutrients (fat, carbohydrate)
Acknowledgements

Women’s Health Initiative (WHI)

Ross Prentice
Johanna Lampe
Lesley Tinker
Ying Huang
Xiaoling Song
Dan Raftery
Cheng Zheng
Dale Schoeller

WHI Clinical Centers and PIs
NHLBI Project Office, HHSN268201100046C
R01 CA119171

Health, Eating, Activity and Lifestyle (HEAL)

Rachel Ballard
Adriana Villasenor
Anne McTiernan
Anita Ambs
Kathy Baumgartner
Richard Baumgartner
Leslie Bernstein

NCI N01-CN-75036-20