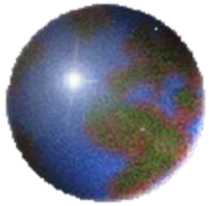

Nutrition Biomarkers in Chronic Disease Prevention Research

Laurence Freedman

Gertner Institute for Epidemiology,

Tel Hashomer, Israel

Acknowledgments

- Statisticians:

  Victor Kipnis
  Raymond Carroll
  Douglas Midthune
  Kevin Dodd
  Janet Tooze

- Epidemiologists/Nutritionists:

  Arthur Schatzkin
  Amy Subar
  Nancy Potischman
  Natasa Tasevska
  Julie Mares
  Lesley Tinker

Content

- Brief overview
- Effects of measurement error in self-report instruments
- Validation studies with recovery biomarkers and their implications
- Combining self-reports and biomarkers

Background I

- Interest in measuring dietary intake and relating this to health outcomes
- Main studies in chronic disease prevention have been:
  - Case-control studies
  - Cohort studies
  - Randomized dietary intervention trials

Background II

The role of dietary intake measurement is very different in these studies:

- Case-control studies & Cohort studies:

  Primary role, since it is the main exposure measure

- Randomized dietary intervention trials

  Secondary role, since it is a measure of compliance and a potential mediator

Background III

- Dietary intake has mainly been measured through self-report instruments

  e.g., Food Frequency Questionnaire (FFQ)
- Inconsistent results across studies, and failure to find evidence for apparently strong hypotheses (e.g. fat and breast cancer)
- Consequent doubts about the accuracy of measurements from such instruments
- Search for biomarkers of dietary intake

Background III

Role of nutrition biomarkers

(i) Assessment of compliance in randomized trials

(ii) Mediation analyses in randomized trials

(iii) Validation of self-report instruments for cohort studies

(iv) Adjusting for the bias in estimated risks in cohort studies

(v) Recovering lost power in cohort studies

Background IV

Types of nutrition biomarker

- **Recovery biomarkers**

Give an essentially unbiased estimate of intake over a given period:

(i) Doubly labelled water – energy intake

(iii) 24 hour urinary nitrogen – protein intake

(iv) 24 hour urinary potassium – potassium intake

Useful for validation of self-report instruments.
Very expensive or difficult to collect.

Background V

Types of nutrition biomarker II

- **Concentration biomarkers**

  Subject to complex metabolic pathways in their regulation: correlated to intake but not an unbiased measure:

  e.g., serum carotenoids, lipids, vitamins, etc.

  Less useful for validation, but could be useful for prediction of some dietary intakes.

  Often less expensive and easier to collect.

Dietary measurement error I

Setting: Cohort study of diet and disease

Exposure: Usual dietary intake, X

Outcome: Disease (often quite rare), Y

Interest: To estimate the relationship between X and Y

$$h(E(Y)) = \alpha_0 + \alpha_1 X$$

Problem: We observe not X, but a self-report, W, that has some error δ:

$$W = X + \delta$$

Dietary measurement error II

The measurement error in W causes two major problems:

- Bias in the estimated relationship
- Loss of statistical power to detect the relationship

Dietary measurement error III

Example: classical measurement error model

$$W = X + \delta$$

$$E(\delta) = 0$$

$$\delta \perp X$$

Then ,

$$E(\hat{\alpha}_{1W}) = \lambda \alpha_1$$

where

$$\lambda = \mathrm{var}(X) / \mathrm{var}(W) < 1$$

so the estimated coefficient is attenuated

Dietary measurement error IV

Statistical power

The effective sample size is reduced from **n** to ρ^2**n**, where ρ is the correlation between W and X.

Validation studies with recovery biomarkers allow us to estimate λ and ρ, and thereby gauge how serious is the problem.

The first large validation study with recovery biomarkers

The OPEN Study

- Conducted by the National Cancer Institute, 1999- 2000
- 261 men, 223 women
- Dietary instruments: 24HR (twice), FFQ (twice)

- Biomarkers: Doubly Labeled Water (for Energy)
  Urinary Nitrogen (for Protein)
  Urinary Potassium (for Potassium)

- These biomarkers have been shown in previous studies to give unbiased measures of these intakes

How serious are the problems? I

Biased Estimation

FFQ attenuation factors, λ, for selected nutrients (OPEN):

Nutrient	Men	Women
Energy	0.08	0.04
Protein	0.16	0.14
Protein Density	0.40	0.32

Note:

The attenuation improves after adjustment for energy

How serious are these problems? II

Average estimated RRs when true RR = 2:

Nutrient	Men	Women
Energy	1.06	1.03
Protein	1.12	1.10
Protein Density	1.32	1.25

Note:

It is generally thought that uncontrolled confounding precludes reliably detecting RRs <1.25 in a cohort study

How serious are these problems? IV

FFQ correlations with true usual intake, ρ, for selected nutrients (OPEN):

Nutrient	Men	Women
Energy	0.08	0.04
Protein	0.16	0.14
Protein Density	0.40	0.32

Conclusions from the OPEN Study

1. Even after adjustment for energy there is serious attenuation of estimated RRs due to the measurement error in a FFQ.

2. There is also serious loss of power to detect diet-disease relationships.

Caveat: OPEN and similar studies can examine only protein, energy and potassium. We can only extrapolate to other nutrients and foods.

Traditional remedy to the problem of RR attenuation

Regression Calibration:

Y = disease; X = true dietary intake; W = self-reported intake

Instead of regressing Y on W (leading to bias) we can regress Y on E(X|W).

(How we determine E(X|W) is another story!)

This leads to (nearly) unbiased estimates, but no gain in power

Remedy to these problems using (concentration) biomarkers

Enhanced Regression Calibration:

We also have a biomarker M for X;

So regress Y on E(X|W,M) instead of E(X|W)

Under certain circumstances, we can obtain unbiased estimates and **gains in power**.

Predicted sample size reduction is by factor:

var(E(X|W) / var(E(X|W,M)

Remedy to these problems using markers

Example:

Carotenoids and Eye Disease Study (CAREDS)

Relation between dietary lutein/zeaxanthin and eye cataracts

- X = log (true usual lutein/zeaxanthin intake)
- Self-report instrument, W = log (FFQ)
- Biomarker, M= log (serum lutein/zeaxanthin)
- Outcome, Y = Eye cataracts (yes/no)

- Analysis = logistic regression of Y on explanatory variables

  Models: 1: W

  2: E(X|W)

  3: E(X|W, M)

Remedy to these problems using markers

Example:

To evaluate E(X|W) and E(X|W, M), we need a measurement error model relating these measures to true intake.

We developed such a model based on feeding studies and validation studies reported in the literature (Freedman et al, Epidemiol Persp Innov, 2010):

$$W = 0.35 + 0.71\,X + \varepsilon_W;\ \mathrm{var}(\varepsilon_W) = 0.36$$

$$M = 5.29 + 0.60\,X + \varepsilon_M;\ \mathrm{var}(\varepsilon_M) = 0.15$$

$$\mathrm{var}(X) = 0.25$$

var(E(X|W)) / var(E(X|W,M)) = 0.53

Combining self-report with biomarker

Carotenoids in Eye Disease Study (WHI)
Analysis of relationship between dietary lutein and eye cataracts

Model	log OR*	(SE)	z-value	Sample size ratio
W=log(FFQ)	-0.165	(0.080)	-2.07	1.00
E(X|W)	-0.464	(0.225)	-2.07	1.00
E(X|W,M)	-0.506	(0.161)	-3.15	0.43

* Adjusted for age and smoking

Combining self-report with biomarker

Caveats

1. Using E(X|W,M), although the **Wald test is valid**, the log OR estimate (-0.506) is **biased** (inflated). However, there is a way to obtain an unbiased estimate.

2. There is always the lurking danger of unknown **confounders** that are involved in the complex metabolic pathways that determine the biomarker level and are also associated with the disease.

Future

1. Prentice et al (Stat Biosci, 2009)

- Large feeding study (150 women)
- Provided with a personalized diet for several weeks
- Measure many biomarkers in blood and urine at end of period
- Develop regression prediction equations for true intakes based on the array of biomarkers
- Use these in place of reported intake

2. Challenge: to incorporate self-report into these prediction equations

Summary

1. Dietary biomarkers have proved extraordinarily useful in quantifying the extent of the problems caused by measurement error in self-reports.

2. The time has now come to invest in developing their use to solve the problems.

3. This development should go hand-in-hand with attempts to improve self-report instruments for cohort studies.