Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

Much room for optimism on measuring diet, preventing cancer and cardiovascular disease, and correcting for measurement error – discussion of the paper by R. L. Prentice and Y. Huang

Donna Spiegelman

To cite this article: Donna Spiegelman (2018) Much room for optimism on measuring diet, preventing cancer and cardiovascular disease, and correcting for measurement error – discussion of the paper by R. L. Prentice and Y. Huang, Statistical Theory and Related Fields, 2:1, 14-20, DOI: 10.1080/24754269.2018.1493632

To link to this article: https://doi.org/10.1080/24754269.2018.1493632

Published online: 31 Jul 2018.

STATISTICAL THEORY AND RELATED FIELDS
2018, VOL. 2, NO. 1, 14–20
https://doi.org/10.1080/24754269.2018.1493632

Taylor & Francis
Taylor & Francis Group

Much room for optimism on measuring diet, preventing cancer and cardiovascular disease, and correcting for measurement error – discussion of the paper by R. L. Prentice and Y. Huang

Donna Spiegelman

Harvard T.H. Chan School of Public Health, Boston, MA, USA

I commend Drs. Prentice and Huang (PH) for their continued and longstanding commitment to the generation of knowledge concerning the nutritional causes of chronic diseases, and in their dedication to applying their extensive, and in the case of Dr. Prentice, virtually unparalleled, expertise towards the development of statistical methods in this regard, to which this paper is one of a long line of such contributions.

I begin this piece by pointing out that this article (PH) and my response continue a debate about the quality of data in nutritional epidemiology and about the interpretation of findings in nutritional epidemiology in the face of possible limitations of these data that goes back 20 years or more. On one side, has been Dr. Prentice and colleagues, particularly from Fred Hutchinson Cancer Research Institute and the National Cancer Institute, and on the other, have been myself, Dr. Walter Willett, and our colleagues based mostly at, or trained at, Harvard.

There is great interest among people everywhere and among funders of research as well, in particular the U.S. National Institutes of Health, about the relationship between diet and the leading causes of mortality and morbidity around the world (not just in high-income countries as PH write): cancer and cardiovascular disease (Network GBoDC, 2017). Nearly 30 years ago, Dr. Prentice and others successfully lobbied the U.S. Congress to earmark $625,000,000 for the 15-year Women's Health Initiative (WHI) (Howard et al., 2006; Prentice et al., 2006), whose primary goal was to establish through a randomised clinical trial once and for all whether or not dietary fat was a risk factor for breast cancer, as the ecological data discussed by PH (Prentice & Sheppard, 1990) might suggest. The answer: a resounding no, as consistent with a large body of prior epidemiologic data using imperfect measures to assess long-term dietary intake (Hunter et al., 1996; Willett, 2013). Despite the failure of WHI to confirm the hypothesis Dr. Prentice was so deeply committed to testing, and despite similar results from another well-conducted randomised trial of this issue (Martin et al., 2011), which found a non-significant protective effect of increased fat intake, it appears from this article that Dr. Prentice still believes this hypothesis. It is possible that the failure of WHI and the other trial could have been due to non-adherence, because lipid biomarkers that are sensitive to changes in dietary fat intake, which can be used as a measure, albeit imperfect, of adherence, did not differ between the intervention group and the control group in WHI at the study's end (Howard et al., 2010), demonstrating that the substitution of the challenge of measuring sustained exposures across the life course in long-term observational epidemiologic research, for another, perhaps equally vexing one, the challenge of maintaining adherence to a dietary intervention in a long-term individually randomised clinical trial, is not clearly advantageous.

In what follows, I will further discuss aspects of PH's paper concerning the epidemiology, followed by exposure measurement issues and then, the statistics. I will conclude with some summary remarks about each.

1. On the epidemiology

PH assert 'The specific drivers for observed risk elevations for specific chronic diseases are not well understood'. We agree that this is the case for some cancers, but not for most others, e.g., lung cancer and cervical cancer, which are leading sites of cancer in most countries around the world among men and women (Jacques et al., 2015), the former caused primarily by cigarette smoking, and the latter by infection with the human papilloma virus, for which we now have a vaccine. Secondary prevention has been shown to work magnificently for yet a third leading cause of cancer – colorectal – with screening for this disease having been shown to reduce mortality by approximately half (Elmunzer et al., 2012; Mandel et al., 1993). Breast cancer, the leading cause of cancer in women in most countries worldwide, is caused by alcohol consumption (Smith-Warner et al., 1998), with a 10% increase in risk per each increased serving/day, a rather high amount of

CONTACT Donna Spiegelman spiegelman@tom.com Harvard T.H. Chan School of Public Health, Boston, MA, USA

consumption for most women around the world, and prolonged use of hormone replacement therapy (HRT) (Million Women Study Collaborators, 2003), and it has been estimated that approximately 35% of breast cancer can be prevented by reduction of exposure to known modifiable risk factors (Tamimi et al., 2016). A possibly unintended benefit of Dr. Prentice's WHI was the finding that the short-term use of combined oestrogen plus progestin HRT causes an increase in risk of breast cancer, confirming earlier findings from observational studies. As a result of this, U.S. women dramatically decreased their use of HRT, and breast cancer rates dropped and have remained at this lower level since (Ravdin et al., 2007). These are just a few examples demonstrating large population attributable risks of known causes for the most common cancers.

PH discuss the evidence for sodium intake in relation to cardiovascular disease risk, and assert that the evidence for this association is inconclusive. I returned to the source document they cited (Medicine, 2013), a detailed review commissioned and published by the National Academy of Science. In fact, I discovered that PH have misinterpreted the conclusions from this report: the report clearly affirmed the strength of the evidence for an adverse effect of salt intake greater than 2300 mg/day on the risks of CVD and all-cause mortality (Medicine, 2013, p. 90]. To put this into perspective, 2300 mg is equivalent to one teaspoon of salt per day, while the average intake of salt in the US is around 3400 mg/day, most of it coming from processed foods (Harnack et al., 2017). What remains in question is whether further reductions of daily salt intake below 1 tsp/day will be beneficial, harmful, or neutral, but given current levels of salt intake, the current guidelines are fully adequate to move on to preventive interventions and policy initiatives to reduce intake and lower CVD risk.

I am surprised that PH considers the evidence concerning the dietary and physical activity causes of cardiovascular disease to be inconclusive. In fact, the evidence is quite strong and it is widely accepted that physical activity and healthy diets decrease the risk of this disease and that obesity greatly increases its risk (Eckel et al., 2014a, 2014b; Sacks et al., 2017; Willett, 2013, chapter 19). In fact, the very same WHO document summarising the evidence for diet and chronic disease that PH cite affirms convincing evidence for many dietary factors along with physical activity, where in this document 'convincing' is the highest possible level of evidence ranking (WHO, 2003). Table 10 below excerpts the key table from this document on the evidence for lifestyle factors in relation to cardiovascular disease incidence.

In addition, cigarette smoking is a clearly established and major cause of lung and other cancers, cardiovascular disease, and several other major chronic diseases (Courtney, 2015).

2. On the quality of measurement of nutritional exposures

Dr. Prentice has been emphasising the limitations of self-reported measurements of dietary intake for more than 20 years; in fact, this concern was a major justification for the (failed) WHI. These concerns have been answered in prior numerous publications (Hebert et al., 2014; Satija, Yu, Willett, & Hu, 2015a; Satija, Yu, Willett, & Hu, 2015b). In brief, the remarkable concordance between epidemiologic evidence concerning diet and health, and that obtained from randomised trials clearly demonstrates the validity of self-reported dietary measures (Table 1; excerpted from Satija, Stampfer, Rimm, Willett, & Hu, in press). A second stream of evidence derives from the validation studies themselves, with the first wave validating self-reported questionnaire-based measures with detailed real-time recording of intake, (Chasan-Taber et al., 1996; Willett et al., 1985), and a more recent wave validating these same self-reported dietary measures with recovery biomarkers (Freedman et al., 2014, 2015; Yuan et al., 2017). It should be noted that there was no evidence for differential reporting by body mass index (BMI) or other pre-specified characteristics in these recent papers, co-authored by Dr. Prentice himself, and BMI has been routinely accounted for in epidemiologic analyses (Yuan et al., 2017). Correlated errors have thus far appeared empirically to be weak and thus possible over-estimation of the validity of self-reported measures in previous studies comparing different dietary assessment methods does not appear to be a credible claim. For example, total fat intake (percent of energy) from the food frequency questionnaire (FFQ) predicted blood triglyceride levels at least as strongly as expected from controlled feeding studies (Willett et al., 2001). In addition, the correlation between the FFQ energy-adjusted protein intake with the recovery biomarker, urinary nitrogen divided by energy measured by doubly labelled water, was very similar to the correlation using diet records as the gold standard (Yuan et al., 2018), suggesting a lack of correlated errors in the latter comparison. Given this weight of evidence, I must express my respectful but emphatic disagreement with PH that although it is theoretically possible that extreme versions of errors could 'thoroughly distort or even reverse observed associations', in the face of the large body of data available, this is unlikely, in fact, to be the case.

As for the promise of metabolomics and microbiomic measures of dietary intake, the jury is still out. Objective but biased measures of intake can be quite useful in validation and measurement error adjustment (Spiegelman, Zhao, & Kim, 2005). Although present metabolomic measures seem to be most appropriate for recent intake, the possibility of new markers that better reflect sustained intake over long periods of time in the past cannot be ruled out. I am quite a bit more

Table 1. Examples of effect estimates from prospective cohort studies and RCTs that examine similar diet-disease associations (Satija et al., in press).

Dietary exposure	Health outcome	Prospective cohort effect estimate	RCT effect estimate
Total fat (Cao, Hou, & Wang, 2016; Prentice et al., 2006)	Breast cancer	0.97 (0.92, 1.03)[a,c] (bottom vs top categories)	0.91 (0.83, 1.01)[b] (low fat vs control diet)
Total fat (Harcombe, Baker, & Davies, 2017; Howard et al., 2006)	Coronary heart disease	0.96 (0.91, 1.02)[a,c] (bottom vs top categories)	0.97 (0.90, 1.06)[b] (low fat vs control diet)
Saturated fat (mostly in place of carbohydrate intake) (Hooper, Martin, Abdelhamid, & Davey Smith, 2015; Siri-Tarino, Sun, Hu, & Krauss, 2010)	Coronary heart disease	0.93 (0.84, 1.04)[a,c] (bottom vs top categories)	0.87 (0.74, 1.03)[a] (low fat vs control diet)
Replacing saturated fat with polyunsaturated fat (Jakobsen et al., 2009; Mozaffarian, Micha, & Wallace, 2010)	Coronary heart disease	0.87 (0.77, 0.97)[a] (per 5% of energy replacement)	0.81 (0.70, 0.95)[a] (PUFA replacing SFA vs control diet)
Mediterranean diet (Fung et al., 2009; Martinez-Gonzalez & Bes-Rastrollo, 2014)	Cardiovascular disease	0.61 (0.49, 0.76)[b] (top vs bottom categories)	0.64 (0.53, 0.79)[a] (Mediterranean vs. control diet)
Mediterranean diet (Esposito et al., 2014; Salas-Salvado et al., 2014)	Type 2 diabetes	0.80 (0.68, 0.93)[a] (top vs bottom categories)	0.70 (0.54, 0.92)[b] (Mediterranean vs. control diet)
Potassium (Filippini, Violi, D'Amico, & Vinceti, 2017; Kieneker et al., 2014)	Hypertension	0.83 (0.73, 0.95)[b,c] (top vs bottom categories of urinary K excretion)	
	SBP (mmHg)		-6.22 (-8.82, -3.93)[a] (potassium vs control)
	DBP (mmHg)		−3.47 (−5.22, −1.73)[a] (potassium vs control)
Dietary fiber (Ascherio et al., 1992, 1996; Streppel, Arends, van 't Veer, Grobbee, & Geleijnse, 2005)	Hypertension	0.68 (0.51, 0.92)[b,c] (top vs. bottom categories)	
	SBP (mmHg)	−1.09 (-1.66, −0.52)[b] ($\geq$ 25 g/day vs $<$ 10 g/day)	−1.13 (−2.49, 0.23)[a] (fiber supplementation vs control)
	DBP (mmHg)	−1.11 (−1.50, −0.72)[b] ($\geq$ 25 g/day vs $<$ 10 g/day)	−1.26 (−2.04, −0.48)[a] (fiber supplementation vs control)
Sugar-sweetened beverages (Malik, Pan, Willett, & Hu, 2013)	Weight (kg), adults	0.22 (0.09, 0.34)[a] (per serving/day increase)	0.85 (0.50, 1.20)[a] (increasing SSB vs control)
	BMI (kg/m^2), children	0.06 (0.02, 0.10)[a] (per serving/day increase)	−0.17 (−0.39, 0.05)[a] (reducing SSB vs control)
DASH (Bai et al., 2017; Saneei, Salehi-Abargouei, Esmaillzadeh, & Azadbakht, 2014)	Hypertension	0.85 (0.73–0.98)[b] (top vs bottom categories)	
	SBP (mmHg)		−6.74 (−8.25, −5.23)[a]
	DBP (mmHg)		−3.54 (−4.29, −2.79)[a]

DASH: dietary approaches to stop hypertension; SBP: systolic blood pressure; DBP: diastolic blood pressure.
[a]Meta-analysis.
[b]Single study.
[c]RR inverted for ease of interpretation.

optimistic than PH about the future utility of such developments, and further statistical research on how to best use them could clearly be of interest as well.

3. On the statistics

In 1982, Dr. Prentice published the seminal paper on covariate measurement error in survival data analysis (1982). Since then, a substantial body of literature has developed on this topic, with much of it contributed by Prentice and his colleagues. PH mention that the approximation given by the second unnumbered equation on p. 4 is accurate only when the disease is rare. Further work from Prentice's group extended this initial methodology to allow for valid estimation

Table 10
Summary of strength of evidence on lifestyle factors and risk of developing cardiovascular diseases

Evidence	Decreased risk	No relationship	Increased risk
Convincing	Regular physical activity Linoleic acid Fish and fish oils (EHA and DHA) Vegetables and fruits (including berries) Potassium Low to moderate alcohol intake (for coronary heart disease)	Vitamin E supplements	Myristic and palmitic acids Trans fatty acids High sodium intake Overweight High alcohol intake (for stroke)

even when the disease is not rare using a 'risk set regression calibration approach' (Xie, Wang, & Prentice, 2001). This initial extension was further developed in additional work published by my own group, where the methodology was subsequently generalised to allow not only simple measurement error models (PH Equation (1)) but also models including time-varying covariates as in PH Equations (4) and (5) (Liao, Zucker, Li, & Spiegelman, 2011) where the necessary functional form of the exposure data is validated in an external or internal sample. Most importantly, our recent work developed the methodology for applications where the exposure of interest in the outcome model is a function of the exposure history, e.g., the cumulative average (Hu et al., 1999) or cumulative total, but only the point exposure is validated, as is almost always the case (Xiaomei et al., 2018). It should be noted that all of these methods have been implemented in a single user-friendly publicly available SAS macro that is freely and publicly available (Spiegelman, 2013). Asymptotic variances were derived for all cases of interest, obviating the computationally intensive bootstrap, in contrast to what PH assert. We agree with PH that more work is needed to further extend these methods to handle non-linear exposure–response relationships. Although these occur infrequently they can be important, such as in the case of the J-shaped relationship between alcohol intake and cardiovascular disease (Corrao, Bagnardi, Zambon, & Arico, 1999), or in situations, perhaps more common, where there is a window of susceptibility of exposure after and before which points the exposure has little or no effect (Wang, Liao, Laden, & Spiegelman, 2016).

Next, PH raised the issue of survival data analysis with exposure variable misclassification. It is unfortunate that the authors seemed to have been unaware of a series of papers addressing this topic (Zucker & Spiegelman, 2004, 2008), as it would have been of interest to have understood their views on what further needs to be done in this area, beyond what has already been accomplished. We have some ideas about this, however. We have found that with the typical amounts of misclassification occurring in dietary data, with the complexities of correlations of both underlying variables and the errors themselves, and with the relatively small sizes of most existing validation studies, further work is needed to strengthen the robustness of the existing methods. The regression calibration option may be one such approach (Spiegelman, Rosner, & Logan, 2000). In our own extensive simulation studies of a likelihood-based logistic regression model with multivariable misclassification and measurement error published in 2000 (Spiegelman et al., 2000), we found that despite its approximations, regression calibration outperformed maximum likelihood estimation for the misclassified variables, similar to what is reported for survival data analysis here. However, it is surprising that PH's simulations found that performance of the regression calibration approximation was excellent even when the rare disease assumption was violated, as this was not the case in their previous paper (Xie et al., 2001) or in several of ours.

It was shown by Carroll, Ruppert, Stefanski, and Crainiceanu (2006) and others that regression calibration applies to general measurement error models, i.e., $E(Z|W) = \mu(W;\ \alpha)$ as long as $Var\ (Z|W)$ is constant. We found, however, that a second-order extension of regression calibration which theoretically should largely address the heteroscedasticity issue did not perform better than standard regression calibration in finite samples (Spiegelman, Logan, & Grove, 2011). Although this issue has not specifically been studied in survival data analysis, all results found for regression calibration in generalised linear models have been found to also apply to Cox regression, at least under the rare disease assumption, and it is unlikely that there would be any difference here.

4. Conclusion

Standard methods for the assessment of long- and short-term dietary intake in nutritional epidemiology have led to the discovery of many actionable and reproducible associations, and when comparable randomised trial data are available, most of these have been consistent with findings from these designs. Metabolomics holds much promise for further dietary exposure validation, most feasibly through the use as main study/validation study designs (Spiegelman & Gray, 1991) and two-stage designs (Breslow & Cain, 1988; Cain & Breslow, 1988). Well-developed methods exist for survival data analysis with covariate measurement error of a great deal of complexity and flexibility. Further work may be needed to extend these methods to allow for non-linear outcome model structures, especially those that empirically estimate the impact of timing of prolonged exposures.

Rather than developing new statistical methods, the greatest need is to foster widespread application of the existing methods in nutritional epidemiology, so that valid estimation and inference can proceed no longer biased by measurement error and misclassification. Validation studies large enough to provide adequate power to apply these methods are now available, and presently under-utilised for measurement error correction in scientific publications (Yuan et al., 2017; Park et al., 2018). We encourage investigators to make regular use of the rich set of existing methods to address and largely remove bias in estimation and inference due to exposure measurement error.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Ascherio, A., Hennekens, C., Willett, W. C., Sacks, F., Rosner, B., Manson, J., … Stampfer, M. J. (1996). Prospective study of nutritional factors, blood pressure, and hypertension among US women. *Hypertension*, *27*(5), 1065–1072.

Ascherio, A., Rimm, E. B., Giovannucci, E. L., Colditz, G. A., Rosner, B., Willett, W. C., … Stampfer, M. J. (1992). A prospective study of nutritional factors and hypertension among US men. *Circulation*, *86*(5), 1475–1484.

Bai, G., Zhang, J., Zhao, C., Wang, Y., Qi, Y., & Zhang, B. (2017). Adherence to a healthy lifestyle and a DASH-style diet and risk of hypertension in Chinese individuals. *Hypertension Research*, *40*(2), 196–202.

Breslow, N. E., & Cain, K. C. (1988). Logistic-regression for 2-stage case-control data. *Biometrika*, *75*(1), 11–20.

Cain, K. C., & Breslow, N. E. (1988). Logistic regression analysis and efficient design for two-stage studies. *American Journal of Epidemiology*, *128*, 1198–1206.

Cao, Y., Hou, L., & Wang, W. (2016). Dietary total fat and fatty acids intake, serum fatty acids and risk of breast cancer: A meta-analysis of prospective cohort studies. *International Journal of Cancer*, *138*(8), 1894–1904.

Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models* (2nd ed.). London: Chapman & Hall.

Chasan-Taber, S., Rimm, E. B., Stampfer, M. J., Spiegelman, D., Colditz, G. A., Giovannucci, E., … Willett, W. C. (1996). Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology*, *7*(1), 81–86.

Corrao, G., Bagnardi, V., Zambon, A., & Arico, S. (1999). Exploring the dose-response relationship between alcohol consumption and the risk of several alcohol-related conditions: A meta-analysis. *Addiction*, *94*(10), 1551–1573.

Courtney, R. (2015). The health consequences of smoking-50 years of progress: A report of the surgeon general, 2014. *Drug & Alcohol Review*, *34*(6), 694–695.

Eckel, R. H., Jakicic, J. M., Ard, J. D., de Jesus, J. M., Houston Miller, N., Hubbard, V. S., … Yanovski, S. Z. (2014a). 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Journal of the American College of Cardiology*, *63*(25 Pt B), 2960–2984.

Eckel, R. H., Jakicic, J. M., Ard, J. D., de Jesus, J. M., Houston Miller, N., Hubbard, V. S., … Tomaselli, G. F. (2014b). 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*, *129*(25 Suppl 2), S76–S99.

Elmunzer, B. J., Hayward, R. A., Schoenfeld, P. S., Saini, S. D., Deshpande, A., & Waljee, A. K. (2012). Effect of flexible sigmoidoscopy-based screening on incidence and mortality of colorectal cancer: A systematic review and meta-analysis of randomized controlled trials. *PLoS Medicine*, *9*(12), e1001352.

Esposito, K., Chiodini, P., Maiorino, M. I., Bellastella, G., Panagiotakos, D., & Giugliano, D. (2014). Which diet for prevention of type 2 diabetes? A meta-analysis of prospective studies. *Endocrine*, *47*(1), 107–116.

Filippini, T., Violi, F., D'Amico, R., & Vinceti, M. (2017). The effect of potassium supplementation on blood pressure in hypertensive subjects: A systematic review and meta-analysis. *International Journal of Cardiology*, *230*, 127–135.

Freedman, L. S., Commins, J. M., Moler, J. E., Arab, L., Baer, D. J., Kipnis, V., … Willett, W. (2014). Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. *American Journal of Epidemiology*, *180*(2), 172–188.

Freedman, L. S., Commins, J. M., Moler, J. E., Willett, W., Tinker, L. F., Subar, A. F., … Prentice, R. L. (2015). Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. *American Journal of Epidemiology*, *181*(7), 473–487.

Fung, T. T., Rexrode, K. M., Mantzoros, C. S., Manson, J. E., Willett, W. C., & Hu, F. B. (2009). Mediterranean diet and incidence of and mortality from coronary heart disease and stroke in women. *Circulation*, *119*(8), 1093–1100.

Harcombe, Z., Baker, J. S., & Davies, B. (2017). Evidence from prospective cohort studies does not support current dietary fat guidelines: A systematic review and meta-analysis. *British Journal of Sports Medicine*, *51*(24), 1743–1749.

Harnack, L. J., Cogswell, M. E., Shikany, J. M., Gardner, C. D., Gillespie, C., Loria, C. M., … Steffen, L. M. (2017). Sources of sodium in US adults from 3 geographic regions clinical perspective. *Circulation*, *135*(19), 1775–1783.

Hebert, J. R., Hurley, T. G., Steck, S. E., Miller, D. R., Tabung, F. K., Peterson, K. E., … Frongillo, E. A. (2014). Considering the value of dietary assessment data in informing nutrition-related health policy. *Advances in Nutrition (Bethesda, Md)*, *5*(4), 447–455.

Hooper, L., Martin, N., Abdelhamid, A., & Davey Smith, G. (2015). Reduction in saturated fat intake for cardiovascular disease. *Cochrane Database of Systematic Reviews*, *6*, CD011737.

Howard, B. V., Van Horn, L., Hsia, J., Manson, J. E., Stefanick, M. L., Wassertheil-Smoller, S., … Kotchen, J. M. (2006). Low-fat dietary pattern and risk of cardiovascular disease: The women's health initiative randomized controlled dietary modification trial. *JAMA*, *295*(6), 655–666.

Howard BV, Curb JD, Eaton CB, Kooperberg, C., Ockene, J., Kostis, J. B., … Van Horn, L. Low-fat dietary pattern and lipoprotein risk factors: The Women's Health Initiative Dietary Modification Trial. *The American Journal of Clinical Nutrition*. 2010;91(4):860–874.

Hu, F. B., Stampfer, M. J., Rimm, E., Ascherio, A., Rosner, B. A., Spiegelman, D., & Willett, W. C. (1999). Dietary fat and coronary heart disease: A comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. *American Journal of Epidemiology*, *149*(6), 531–540.

Hunter, D. J., Spiegelman, D., Adami, H. O., Beeson, L., van den Brandt, P. A., Folsom, A. R., … Marshall, J. R., et al. (1996). Cohort studies of fat intake and the risk of breast cancer–a pooled analysis. *New England Journal of Medicine*, *334*(6), 356–361.

Jacques, F., Isabelle, S., Rajesh, D., Sultan, E., Colin, M., Marise, R., … Freddie, B. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, *136*(5), E359–E386.

Jakobsen, M. U., O'Reilly, E. J., Heitmann, B. L., Pereira, M. A., Balter, K., Fraser, G. E., … Ascherio, A. (2009). Major types of dietary fat and risk of coronary heart disease: A pooled analysis of 11 cohort studies. *The American Journal of Clinical Nutrition*, *89*(5), 1425–1432.

Kieneker, L. M., Gansevoort, R. T., Mukamal, K. J., de Boer, R. A., Navis, G., Bakker, S. J., & Joosten, M. M. (2014).

Urinary potassium excretion and risk of developing hypertension: The prevention of renal and vascular end-stage disease study. *Hypertension*, *64*(4), 769–776.
Liao, X., Zucker, D. M., Li, Y., & Spiegelman, D. (2011). Survival analysis with error-prone time-varying covariates: A risk set calibration approach. *Biometrics*, *67*, 50–58.
Malik, V. S., Pan, A., Willett, W. C., & Hu, F. B. (2013). Sugar-sweetened beverages and weight gain in children and adults: A systematic review and meta-analysis. *The American Journal of Clinical Nutrition*, *98*(4), 1084–1102.
Mandel, J. S., Bond, J. H., Church, T. R., Snover, D. C., Bradley, G. M., Schuman, L. M., & Ederer, F. (1993). Reducing mortality from colorectal cancer by screening for fecal occult blood. *New England Journal of Medicine*, *328*(19), 1365–1371.
Martin, L. J., Li, Q., Melnichouk, O., Greenberg, C., Minkin, S., Hislop, G., & Boyd, N. F. (2011). A randomized trial of dietary intervention for breast cancer prevention. *Cancer Research*, *71*(1), 123–133.
Martinez-Gonzalez, M. A., & Bes-Rastrollo, M. (2014). Dietary patterns, Mediterranean diet, and cardiovascular disease. *Current Opinion in Lipidology*, *25*(1), 20–26.
Medicine, I. (2013). *Sodium intake in populations: Assessment of evidence*. Washington, DC: The National Academies Press.
Million Women Study Collaborators. (2003). Breast cancer and hormone-replacement therapy in the Million Women Study. *The Lancet*, *362*(9382), 419–427.
Mozaffarian, D., Micha, R., & Wallace, S. (2010). Effects on coronary heart disease of increasing polyunsaturated fat in place of saturated fat: A systematic review and meta-analysis of randomized controlled trials. *PLoS Medicine*, *7*(3), e1000252.
Network GBoDC. (2017). Global burden of disease study 2016. In: (IHME) nfHMaE, ed. Seattle, USA.
Park, Y., Dodd, K. W., Kipnis, V., Thompson, F. E., Potischman, N., Schoeller, D. A., … Subar, A. F. (2018). Comparison of self-reported dietary intakes from the automated self-administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. *American Journal of Clinical Nutrition*, *107*(1), 80–93.
Prentice, R. L. (1982). Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika*, *69*, 331–342.
Prentice, R. L., Caan, B., Chlebowski, R. T., Patterson, R., Kuller, L. H., Ockene, J. K., … Henderson, M. M. (2006). Low-fat dietary pattern and risk of invasive breast cancer: The Women's Health Initiative randomized controlled dietary modification trial. *JAMA*, *295*(6), 629–642.
Prentice, R. L., & Sheppard, L. (1990). Dietary fat and cancer: Consistency of the epidemiologic data, and disease prevention that may follow from a practical reduction in fat consumption. *Cancer Causes & Control*, *1*(1), 81–97. ; discussion 99–109.
Ravdin, P. M., Cronin, K. A., Howlader, N., Berg, C. D., Chlebowski, R. T., Feuer, E. J., … Berry, D. A. (2007). The decrease in breast-cancer incidence in 2003 in the United States. *New England Journal of Medicine*, *356*(16), 1670–1674.
Sacks, F. M., Lichtenstein, A. H., Wu, J. H. Y., Appel, L. J., Creager, M. A., Kris-Etherton, P. M., … American Heart Association. (2017). Dietary fats and cardiovascular disease: A presidential advisory from the American Heart Association. *Circulation*, *136*(3), e1–e23.
Salas-Salvado, J., Bullo, M., Estruch, R., Ros, E., Covas, M. I., Ibarrola-Jurado, N., … Martinez-Gonzalez, M. A. (2014). Prevention of diabetes with Mediterranean diets: A subgroup analysis of a randomized trial. *Annals of Internal Medicine*, *160*(1), 1–10.
Saneei, P., Salehi-Abargouei, A., Esmaillzadeh, A., & Azadbakht, L. (2014). Influence of dietary approaches to stop hypertension (DASH) diet on blood pressure: A systematic review and meta-analysis on randomized controlled trials. *Nutrition, Metabolism, and Cardiovascular Diseases*, *24*(12), 1253–1261.
Satija, A., Stampfer, M. J., Rimm E. B., Willett, W. C., & Hu, F. B. (in press). Are large simple trials the solution for nutritional research? *Advances in Nutrition*.
Satija, A., Yu, E., Willett, W. C., & Hu, F. B. (2015a). Objective measures are complementary to, rather than a replacement for, self-reported methods. *International Journal of Obesity (London)*, *39*(7), 1179.
Satija, A., Yu, E., Willett, W. C., & Hu, F. B. (2015b). Understanding nutritional epidemiology and its role in policy. *Advances in Nutrition (Bethesda, MD)*, *6*(1), 5–18.
Siri-Tarino, P. W., Sun, Q., Hu, F. B., & Krauss, R. M. (2010). Meta-analysis of prospective cohort studies evaluating the association of saturated fat with cardiovascular disease. *The American Journal of Clinical Nutrition*, *91*(3), 535–546.
Smith-Warner, S. A., Spiegelman, D., Yaun, S. S., van den Brandt, P. A., Folsom, A. R., Goldbohm, R. A., & Hunter, D. J. (1998). Alcohol and breast cancer in women: A pooled analysis of cohort studies. *JAMA*, *279*(7), 535–540.
Spiegelman, D. 2013. %rcc Macro. 2013. Retrieved from http://www.hsph.harvard.edu/donna-spiegelman/software/rrc-macro/
Spiegelman, D., & Gray, R. (1991). Cost-efficient study designs for binary response data with Gaussian covariate measurement error. *Biometrics*, *47*(3), 851–869.
Spiegelman, D., Logan, R., & Grove, D. (2011). Regression calibration with heteroscedastic error variance. *International Journal of Biostatistics*, *7*(1), 4.
Spiegelman, D., Rosner, B., & Logan, R. (2000). Estimation and inference for logistic regression with covariate misclassification and measurement error in main study/validation study designs. *Journal of the American Statistical Association*, *95*, 51–61.
Spiegelman, D., Zhao, B., & Kim, J. (2005). Correlated errors in biased surrogates: Study designs and methods for measurement error correction. *Statistics in Medicine*, *24*(11), 1657–1682.
Streppel, M. T., Arends, L. R., van 't Veer, P., Grobbee, D. E., & Geleijnse, J. M. (2005). Dietary fiber and blood pressure: A meta-analysis of randomized placebo-controlled trials. *Archives of Internal Medicine*, *165*(2), 150–156.
Tamimi, R. M., Spiegelman, D., Smith-Warner, S. A., Wang, M., Pazaris, M., Willett, W. C., … Hunter, D. J. (2016). Population attributable risk of modifiable and nonmodifiable breast cancer risk factors in postmenopausal breast cancer. *American Journal of Epidemiology*, *184*(12), 884–893.
Wang, M., Liao, X., Laden, F., & Spiegelman, D. (2016). Quantifying risk over the life course - latency, age-related susceptibility, and other time-varying exposure metrics. *Statistics in Medicine*, *35*(13), 2283–2295.
WHO. (2003). Diet, nutrition and the prevention of chronic diseases. *World Health Organization Technical Report Series*, *916*, i–viii. 1–149, backcover.
Willett, W. (2013). *Nutritional epidemiology* (3rd ed.). Oxford: Oxford University Press.

Willett, W. C., Sampson, L., Stampfer, M. J., Rosner, B., Bain, C., Witschi, J., ... Speizer, F. E. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *American Journal of Epidemiology*, *122*, 51–65.

Willett, W., Stampfer, M., Chu, N.-F., Spiegelman, D., Holmes, M., & Rimm, E. (2001). Assessment of questionnaire validity for measuring total Fat intake using plasma lipid levels as criteria. *American Journal of Epidemiology*, *154*(12), 1107–1112.

Xiaomei, L., Xin, Z., Molin W., Jaime, E. H., Francine, L., & Donna, S. (2018). Survival analysis with functions of mismeasured covariate histories: The case of chronic air pollution exposure in relation to mortality in the nurses' health study. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *67*(2), 307–327.

Xie, S. X., Wang, C. Y., & Prentice, R. L. (2001). A risk set calibration method for failure time regression by using a covariate reliability sample. *Journal of the Royal Statistical Society, Series B*, *63*, 855–870.

Yuan, C., Spiegelman, D., Rimm, E. B., Rosner, B. A., Stampfer, M. J., Barnett, J. B., ... Willett, W. C. (2018). Relative validity of nutrient intakes assessed by questionnaire, 24-hour recalls, and diet records as compared With urinary recovery and plasma concentration biomarkers: Findings for women. *American Journal of Epidemiology*, *187*(5), 1051–1063.

Yuan, C. Z., Spiegelman, D., Rimm, E. B., Rosner, B. A., Stampfer, M. J., Barnett, J. B., ... Willett, W. C. (2017). Validity of a dietary questionnaire assessed by comparison With multiple weighed dietary records or 24-hour recalls. *American Journal of Epidemiology*, *185*(7), 570–584.

Zucker, D. M., & Spiegelman, D. (2004). Inference for the proportional hazards model with misclassified discrete-valued covariates. *Biometrics*, *60*(2), 324–334.

Zucker, D. M., & Spiegelman, D. (2008). Corrected score estimation in the proportional hazards model with misclassified discrete covariates. *Statistics in Medicine*, *27*, 1911–1933.