

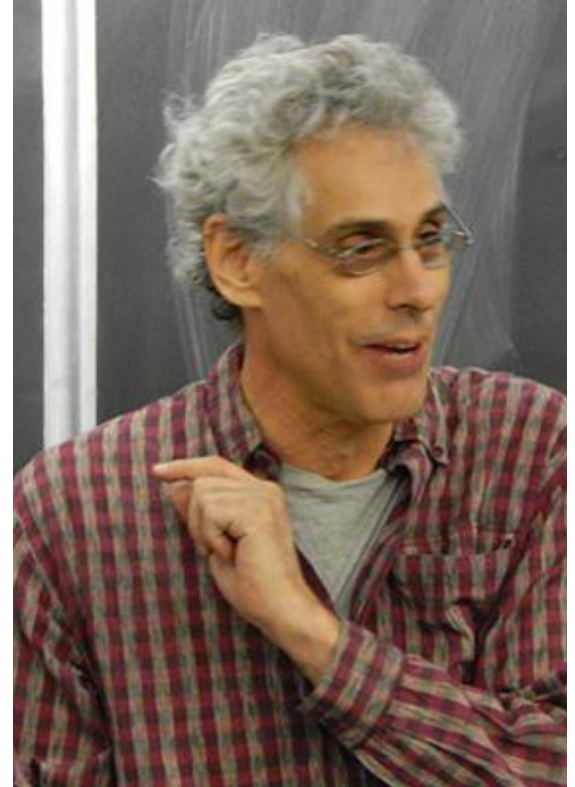
# ERD

Examine.com  
Research Digest

Issue 49, Vol 1 of 2 ♦ November 2018

# INTERVIEW:

Sander Greenland MS, DrPH



*Several observational studies in nutrition will often contradict one another, where one day coffee is good for you, and the next day it's not so good for you. In the last few years, [meta-researchers](#) have pointed out that studies that come out of this field (nutritional epidemiology) are not very useful because of problems such as the multiplicity of statistical models, high amounts of measurement and sampling error, and difficulty separating signal from noise. Do you think nutritional epidemiology has much utility in determining cause and effect relationships or utility in generating hypotheses?*

For nutrition, the lack of biochem sophistication among the trial designers leads to a lot of dubious and noncomparable studies, while the meta-analysts and reviewers do a lot of distortive lumping, e.g., talking about “vitamin E” as if that were a single entity - [a recent review by prominent authors](#) didn't even notice that almost all trials used the racemic synthetic mixture, dl-alpha-tocopheryl, (which they misidentify with “alpha tocopherol”) with vastly different and unjustified dosages, and which hardly resembles the eight or so natural d-tocopherols or d-tocotrienols that account for dietary intake. The literature on nutrition replication is even more appalling - heavily based on fallacies like thinking there is conflict because one study had  $P < 0.05$  and the other had  $P > 0.05$ ; and often claiming that RCTs found no effect based on fallacies like seeing all  $P > 0.05$  when the trials were far too underpowered to find the key long-term effects. See especially 3-8 and 15-17 in our [2016 TAS supplement](#).

So my answer is that the entire nutritional literature could have been very useful if it had been designed and put together with proper respect for both the biology and the stats. But it wasn't and still isn't - basically the “experts” did it in.

*Clinical trials or observational studies will often use some form of randomization, whether it be random assignment in an experiment or random sampling in a survey. Could you explain why using randomization is essential for reducing bias and why statistical tests often depend on them? Do these types of tests have much utility in studies that don't utilize some form of randomization?*

Randomization is not “essential” but is promoted that way by certain interests. [Here's](#) an example of a RCT using an insane dosing protocol reported by Medscape as if meaningful. Read the comments that note the dosing as well as other problems such as lack of attention to cofactors like vitamin K2.

*In one of your upcoming articles, you discuss an article by psychologist Daniel Lakens that promotes designing studies with an emphasis on controlling false positives and false negatives using background information to judge how important these types of errors are. Why do you think such methods may not be suitable for research?*

Part of the divergence (maybe all of it) is that he works in an area (experimental psychology) with many small experiments and pressure to publish positive results, whereas I work in an area (hazard surveillance) dominated by large nonexperimental cohorts and databases with pressures to publish negative results on some researchers.

*A few months ago, a form of statistical inference known as magnitude-based inference (MBI) was criticized by several statisticians. This is a form of inference often used by exercise scientists, and several statisticians exclaimed that it produces a high rate of false positives. Could you give us your take on this form of inference?*

MBI as fought over here is not discussed in Modern Epidemiology 3rd ed (2008, Ch. 10), B&H 2006 [a foundational paper proposing MBI] is only cited as one among many papers complaining about NHST. We only

advised looking at the estimates (including the confidence limits) before drawing inferences, advice which no one seriously contests (that I know of). But if some passage looks supportive B&H please point it out, as there's no question that MBI is not a well-founded method for forcing binary decisions out of data (as you can see by all the comments from statisticians including Gelman).

Going back to the B&H 2006 paper we (actually Ken; I never looked at it before this controversy blew up) cited in 2008, Figure 2 isn't even correct in its labeling (declaring ambiguous intervals as if they showed something); and within the frequentist literature the entire scope of the figure is addressed by topics like P-values for alternatives, equivalence testing, inferiority testing, etc. Figure 3 displays utter confusion of frequentist and Bayesian ideas, which they try and rationalize with the "flat prior" argument - never noting that a flat prior is generally nonsensical in scientific terms and suboptimal for both frequentist and Bayesian decision and inference; at most flat priors only serve to bound results from optimized or sensible priors. The figures and text might have been harmless with more purely descriptive

“I think single studies should be presented with emphasis on motivation (background), design, conduct, and the resulting data - in other words, they should be descriptive narrations, like The Voyage of the Beagle, not On the Origin of Species.”

labeling, but it really does take some immersion in stat theory to meld frequentist and Bayesian ideas properly (it can be done and I think should be done routinely; for most users however, doing so would take far more extensive retraining than one can reasonably expect). Instead however MBI was expanded to include all the details that Sainani complained about.

MBI is not the only dubious proposal promoted (sometimes even statisticians do that, the latest example being the “2nd generation P-value”), it just happened to catch on and thus become a problem. But despite my jaundiced view of their MBI, my reaction to the fight is a pox on both sides. I don’t think it’s sound science to claim presence or absence of effects or infer their sizes based on single studies. I think single studies should be presented with emphasis on motivation (background), design, conduct, and the resulting data - in other words,

they should be descriptive narrations, like *The Voyage of the Beagle*, not *On the Origin of Species*. Were I in charge, conclusions (as well as policy recommendations) would be banned from single-study reports; conclusions would instead be reserved for research syntheses, and policy recommendations would be reserved for real policy studies (which would be headed by policy scientists). I’ve explained that view at length in many papers now...

*For folks who are trying to become better acquainted with statistics and study designs, what would you recommend they read?*

I can’t say I know much of what is available. Obviously I am biased toward [Modern Epidemiology](#), which has much on design and analysis! Supplemented by later cautionary writings like the [ASA TAS supplement](#) and my [2017 AJE article](#). ♦

---

Sander Greenland is Professor of Epidemiology and Statistics at the University of California, Los Angeles. He received Bachelor’s and Master’s degrees in mathematics and Master’s and Doctoral degrees in Epidemiology from the University of California. Since then he has become a leading contributor to epidemiologic statistics, theory, and methods. His focus has been the limitations and misuse of statistical methods in observational studies. He has authored or co-authored over 300 articles in epidemiology, statistics, and medical journals, and co-authored the textbook *Modern Epidemiology*. He is a Fellow of the American Statistical Association and the Royal Statistical Society. He has served as an associate editor for several statistics and epidemiology journals, as an advisor for the Food and Drug Administration, the Environmental Protection Agency, the Centers for Disease Control, the State of California, and the National Academy of Sciences, and has been an invited speaker at universities and conferences throughout the world.