Our Estimates *are* Uncertain, but that is ok.

Carl V. Phillips, MPP, PhD

University of Minnesota School of Public Health
Division of Environmental and Occupational Health

phill047@tc.umn.edu

See also Posters #160 and #161. Suggested order for reading: 162, 161, 160.

Manuscripts available for 161, 162, and half of 160. The other half of 160 (food safety) will be available within a few months.

The Policy Information Vicious Circle

What policy makers need for optimal health policy decisions:
Probability Distributions for Parameters of Interest, regardless of how precise the point estimate is.

What policy makers believe that health researchers can provide:
Precise Point Estimates, with No Way to Substantively Analyze Uncertainty.

Why the disconnect?

What health researchers think policy makers want/need/desire?
Precise Point Estimates.

What epidemiology (and most other health) researchers cannot actually honestly provide most of the time:
Precise Point Estimates.

What epidemiologists (et al.) can provide:
Probability Distributions for Parameters of Interest, regardless of how precise the point estimate is.

Reporting Only Point Estimates -- Problem #1

Epidemiology cannot actually provide accurate point estimates.

By creating the fiction that it can, we create a market for the estimates and eliminate the market for better quantifying uncertainty.

Uncertainty is much greater than generally acknowledged, since the quantified uncertainty is based on only one out of the many sources. (See also my other posters.)

Reporting Only Point Estimates -- Problem #2

<u>They keep us from making effective decisions in many cases.</u>

Real-world loss functions from making suboptimal policy choices are seldom symmetrical about the estimated mean of the parameter of interest, and so knowing only whether that point estimate produces positive or negative value is insufficient.

We care a lot about probabilities of various outcomes.

One manifestation (of many) is that many decisions that seem quite likely to provide small benefits may be terrible choices because of the chance of huge losses. (Consider the example of the Chernobyl nuclear power plant. The 95% confidence interval for number of people killed by another accident is [0,0], but Clinton still -- sensibly -- persuaded Ukraine to shut it down.)

To partially make up for this, we have badly-performing kluges like the ill-defined "precautionary principle" and other makeshift precautions include rules that exposures be ten or a thousand times lower than the level of exposure at which negative impacts have been "proven."

Reporting Only Point Estimates -- Problem #3

Status quo bias.

If decision makers expect precise point estimates and updates when they change, no one will think in terms of making the best decision with the available limited information.

The dominant practice is "don't do anything until we are sure that something else is better, whereupon switch to the alternative." This is reflected in epidemiology textbooks, medical research handbooks, policy guidelines, and the actual decisions made.

But, of course, "don't do anything" just means keep doing whatever we are doing, which is a decision like any other, which deserves no particular deference.

This leads to both doing too little to protect the public health in some cases, and doing too much (because the costs outweigh the benefits) in others.

But in the vicious circle, researchers come to believe that only point estimates will be used in decisions (and the status quo will be maintained, even when it looks inferior), and so report accordingly.

Reporting Only Point Estimates -- Problem #4

Confusing lay audiences (including the public and most policy makers)...

Most people do not understand how truly uncertain our "precise" estimates are.

They take each new finding as the new Truth (because researchers portray it that way!), displacing the old knowledge.

When this changes frequently, people think health researchers are hopelessly confused.

They quickly learn to treat all health claims as highly uncertain.

...and researchers themselves.

The tendency to devote our research to finding associations where we can reject the null hypothesis (using our incorrectly narrowed confidence intervals) contributes to data-mining (to create "provable" hypotheses based on the data) and publication bias (reporting only the "proven" results).

Science and public health benefit from learning about associations when they are real. But we apparently forget the lesson from first-semester statistics, that trolling for p=.05 will generate an article from one out of every twenty sets of random numbers (or more, if multiple functional forms are tried) if we do not account for multiple-hypothesis testing.

Low p-values and narrow confidence intervals are *not* important decision parameters.

They are not necessary.

-see below

They are not sufficient.

-a 1% chance that something will lead to disaster is a good reason not to do it, even if you are 99% sure it generates some modest value

They are not true anyway.

-see Poster 161

Health Policy Decision Making

Health policy decisions[*] will always involve some costs and some benefits that must be traded off and some uncertainty that should be formally considered.

The expected net value of an action should be used to determine its worth.

But instead, status quo bias is the dominant rule of thumb.

We need to start by realizing that,

- an optimized decision can be made based on whatever information is available,

- and, indeed, a decision always *will be* made no matter the quality of the current information.

  (staying with the status quo is a decision that implicitly says that the current information favors that course of action)

---

[*]Defined broadly, to include widespread behavioral recommendations and setting standards of practice, as well as government regulations and other formal public policy.

Optimized Decisions

The idealized decision method is to integrate (or sum) the net value of a policy across possible values of the parameter of interest based on our current knowledge (however imperfect) of parameters of interest.

That is, we want to compare the expected net value across all policy options. That value is, in a rather stylized representation,

$$\int f(\boldsymbol{x})\,(Benefit(P,\boldsymbol{x}) - Cost(P,\boldsymbol{x}))\;\; d\boldsymbol{x} \tag{1}$$

where the *Benefit* and *Cost* functions are the social outcomes for policy option *P* given state of the world ***x***, the parameters of interest.

(Dynamic issues, like irreversibilities and what options will provide more information should be considered too, but this does not change the main point.)

Key Point: Function *f* is our probability density on the parameter of interest, ***x***, given current knowledge, considering all sources of error, random and systematic. (The function, like the integral itself, is a notational shorthand, since ***x*** will generally be multi-dimensional and complicated.)

E.g., the parameter of interest might be the relative risk of cancer due to exposure to a pesticide, with one policy option being banning the pesticide, which has an expected benefit of reducing cancer (according to f) and a cost of the value of using the pesticide.

Ideally, we choose the policy with the highest calculated value based on available information, even if it is not "proven" to be better than the status quo.

Which implies:

The chosen (best) policy may have a substantial (even majority) chance of being inferior if its net benefits are high enough for some portion of the density of $\boldsymbol{x}$.

A policy that is probably a bit better than the status quo, but has a small positive chance of being a disaster would be shown to be inferior. (no need for "precautionary principle")

E.g., a chemical that provides a bit of benefit compared to the alternatives but might be causing widespread reproductive damage could have a net negative expected value, even though it is impossible to "prove" it is unsafe.

A new policy could displace the status quo even if there is a substantial chance that it actually has lower net benefits, as long as the net losses are low should they occur, while the net benefits are high.

As daunting as it seems, epidemiology can estimate ***f(x)***.

(for more details, see my other posters)

In so doing, we can

Improve policy (rescuing it from status quo bias and precautionary principle type rules).

and

Increase the contribution and influence of epidemiologic research.

Increasing the Contribution of Epidemiology

Epidemiology, with its many sources of uncertainty, is <u>simply not good at providing evidence that something is absolutely true</u>. With notable exceptions like the link between smoking and lung cancer, population studies seldom demonstrate causal effects that we can believe beyond any reasonable doubt. Instead, they produce a range of plausible values. Since decisions are necessarily made, and should be made based on our assessment of what seems to be better -- rather than what is Proven To Be True, this is ok. Understanding the true uncertainty does not detract from the value of epidemiology, it enhances it.

Environmental epidemiologists are sometimes loath to admit the true uncertainty in their results because they believe it will give even more policy influence to toxicology with its great precision. Indeed, as long as we allow health policy discussions to focus on being able to reject *some* null hypothesis that is *somehow* related to the policy in question (even rather tangentially), epidemiology -- despite its huge contributions -- will remain a poor cousin to bench science and the occasional RCT.

Other health research technologies usually offer more precise measurements of substitutes (possibly not very good ones) for what we actually want to know. Toxicology offers us very controlled and precise measures. But since mice receiving a huge dose of toxin in a controlled setting are seldom our target population of interest, it is obvious that epidemiology offers some major advantages.

Epidemiology offers the possibility of measuring exactly what we are interested in -- such as the effect of a real-life exposure on a variety of people in an everyday setting -- with all the variability and uncertainties introduced by real people in the real world, and the possibility of measuring the variability and uncertainty themselves.

Conclusions

Epidemiology must learn to embrace and properly respond to the inevitable uncertainty, lest it surrender too much of the study of human health to those who can precisely measure things we are not actually interested in.

To become a more honest science, epidemiology needs to start accurately reporting the substantial uncertainty in its results.

To simultaneously keep itself positioned as a critical science in health policy making, it needs to recognize how to properly use known uncertainty in decision making and to educate others on that point.

Addressing the vicious circle will require attention to the optimal methods for making tough policy decisions and an active effort to feed that process the information it really needs. Until this occurs, epidemiology will fall short of its true potential contribution to scientific knowledge while simultaneously overstating its contribution by overselling the certainty of results.

The expected value calculation is difficult, but it is the gold standard and worth pursuing. The impact of health research is frequently many lives and many dollars, and the research itself is expensive, so it is impossible to justify not even *attempting* to estimate the missing values. When we fail to quantify, decisions are wholly given over to a political process, and we are doing science and policy no favors by being too modest about our ability to quantify.

Our estimates are uncertain, but that is ok.