2013/10/01

science 2.0

Predicting The Past Isn't As Easy As It Sounds

http://www.science20.com/steve_hentges/predicting_past_isn%E2%80%99t_easy_it_sounds-121533

A recent article discussed the question of causation versus statistical association in cross-sectional epidemiology studies that evaluate the potential for chemicals to cause health effects. In this type of study, health effect and chemical exposure data are collected at the same point in time, which means there is no way to know, based on the data evaluated, if the exposure preceded the isease. Without this temporal information, statistical associations between exposure and health effects may be derived, but it is not possible to establish causation.

This issue is of particular current relevance due to the ready availability of databases that contain large volumes of cross-sectional exposure and health effect data. A good example is the NHANES (National Health and Nutrition Examination Survey) database from CDC. It's relatively easy to mine the database for statistical associations and, depending on the chosen level of statistical ignificance, it is almost certain that some statistical associations will be found by chance alone.

What's not clear though is which, if any, of these statistical associations have any biological meaning (e.g., causation). Nevertheless, given the ready availability of data to analyze, many researchers have fallen prey to temptation and their resulting statistical associations have frequently been uncritically, even sensationally, reported in the popular media as evidence of causation.

Image credit: BevNet

The recent article discussed a study on the common chemical bisphenol A (BPA), which is known to have a short physiological half-life of only a few hours. Studies have shown that BPA levels in urine, where it is excreted in the form of a metabolite, are highly variable even within a day. This information alone suggests that measurement of BPA in single urine spot samples is unlikely to be indicative of long-term exposure. Accordingly, cross-sectional epidemiology studies on BPA that rely on single urine spot samples are unlikely to provide any information on causation, regardless of what statistical associations might be found.

However, there are some circumstances when a cross-sectional study might provide more information than just statistical associations. For example, if it were known that habitual exposure to a chemical consistently occurred over long periods of time at a particular time of day (e.g., exposure occurs with dinner every day), and urine spot samples for analysis were consistently collected at the same time of day (e.g., first-morning voids), it might be possible to predict past exposures based on current measurements.

Experimental data for specific chemicals is needed to assess whether this is possible, and for BPA a new study provides just the sort of data needed. In this study, two spot urine samples were collected from each of 80 women over a 1-3 year time period. Almost all of the urine samples were first-morning voids. The temporal variability of BPA in urine was assessed by calculating an intraclass correlation coefficient (ICC), which reflects the relationship between within- and between-person variance. The ICC can have a value between 0 and 1 with higher values indicating low within-person variance, which is what would be needed for spot urine sample measurements to have any chance at predicting past (or future) exposures.

The study found high within-person variability of urinary BPA levels with an ICC value of 0.14, meaning there is little correlation between BPA levels in spot urine samples collected 1-3 years part. Removing the few samples that were not first morning voids to improve consistency on timing of sample collection resulted in little improvement with an ICC value of 0.15. The study found slightly less, but still high, variability for samples taken <25 months apart (ICC = 0.23) compared to samples taken >25 months apart (ICC = 0.06), the latter samples showing almost no correlation at all.

The implications of high variability for epidemiologic studies are quite significant. As stated by the authors, “investigation of associations between a single urinary bisphenol A measurement and disease risk may be challenging in epidemiologic studies.” Given the low ICC values, “challenging” may be somewhat of an understatement.

As appropriately noted by the authors, their results are specific to their study participants and may not be generalizable to other populations. Although this new study examined variability over a longer time period, other studies that examined shorter time periods found only slightly better results with ICC values ranging from 0.11 to 0.43. In a CDC study that comprehensively measured BPA levels in urine over the course of a week, within-day variation was the main contributor to total variation (70%), with between-day (21%) and between-person (9%) variability being less significant.

With such high variability demonstrated in multiple studies over shorter time periods, it's not likely that other populations will show significantly lower variability over longer time periods than the population examined in the new study.

Given the high variability in urinary BPA levels over time, the value of cross-sectional epidemiology studies based on single urine spot samples is certainly questionable. Perhaps the studies could be useful for hypothesis generation, even though they have no capability to establish causation, but even hypotheses based on such poor quality data are of questionable value. Since the source of data for many of these studies is the NHANES database, guidance from CDC on the most appropriate use of the data, of which there are many excellent uses, could be helpful.

But guidance is apparently lacking and even CDC researchers have recently indulged in a cross-sectional study on BPA and other short half-life compounds, with only a brief mention in the discussion section of the issues discussed here

Comments Helen Barratt (http://www.science20.com/profile/helen_barratt)
This is all very interesting Steve, its important that scientists don't publish and claim causal links when there is no real evidence in the data they are using to support these claims. I had a quick look at the NHANES database which I hadn't heard of before and I will examine it a lot more closely when I have time. It says there that :-

> *As in past health examination surveys, data will be collected on the prevalence of chronic conditions in the population. Estimates for previously undiagnosed conditions, as well as those known to and reported by respondents, are produced through the survey. Such information is a particular strength of the NHANES program.*
>
> *Risk factors, those aspects of a person's lifestyle, constitution, heredity, or environment that may increase the chances of developing a certain disease or condition, will be examined. Smoking, alcohol consumption, sexual practices, drug use, physical fitness and activity, weight, and dietary intake will be studied. Data on certain aspects of reproductive health, such as use of oral contraceptives and breastfeeding practices, will also be collected.*
> *The diseases, medical conditions, and health indicators to be studied include:*
>
> *Anemia / Cardiovascular disease / Diabetes / Environmental exposures / Eye diseases*
> *Hearing loss / Infectious diseases / Kidney disease / Nutrition / Obesity / Oral health*
> *Osteoporosis / Physical fitness and physical functioning /*

Reproductive history and sexual behavior /Respiratory disease (asthma, chronic bronchitis, emphysema) / Sexually transmitted diseases Vision

Do you know whether the NHANES database holds any data about environmental exposure to a toxin known as β-N-methylamino-l-alanine (BMAA), which is found in blue-green algae and can bioaccumulate in our food chain, particularly fish and crustaceans as well as polluting our water supply?

Research recently published in the open access journal PLOS ONE and reported here by the Motor Neurone Disease Association (mnda) claims that BMAA causes proteins inside cells to clump together and cause cell death. This finding suggests that BMAA may be a potential cause of up to 90% of MND and Lou Gehrig's ALS and may also be linked to other neurodegenerative diseases like Alzheimer's and Parkinson's and could lead to the development of new treatments, which is both good and bad news for people suffering from these terrible debilitating neurodegenerative disorders!

My article about researchers identifying a potential blue green algae cause & L-Serine treatment for Lou Gehrig's ALS, MND, Parkinsons & Alzheimers is at http://www.science20.com/forums/medicine
Helen Barratt | 10/01/13 | 16:44 PM

Reply Comments by Steve Hentges

I've never heard of BMAA but will read your article to learn more now that you've caught my interest. I did take a quick look at an NHANES summary document (http://www.cdc.gov/nchs/data/nhanes/survey_content_99_12.pdf) to look for BMAA, but didn't find it. Take a look yourself since you might find other parameters of interest in the long list of things that are measured or collected in the survey. I'll bet there is also a way to nominate new items for future surveys. It might also be worth taking at look at what's measured in the similar Canadian program called the Canadian Health Measures Survey (http://www.statcan.gc.ca/start-debut-eng.html). The Canadian program was probably modelled after NHANES but might have some useful differences.

The list of risk factors and diseases that you found on the NHANES website highlights what I think is one of the very excellent uses of NHANES data, which is to improve our understanding of disease incidence in the US population and subpopulations. Some accomplishments are listed further down the NHANES webpage that you visited.

Steve