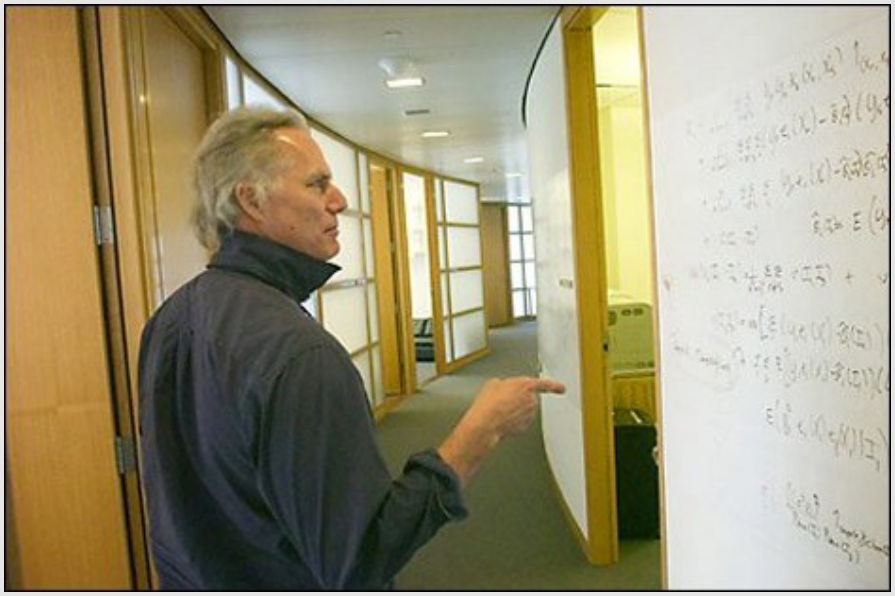
Getting Real

Does air pollution cause people to die? Do different teacher-compensation systems affect students' educational outcomes? Are middle-aged women who take hormone replacement therapy more or less susceptible to heart attacks than those who don't? In these and multiple other real-world scenarios, causal inference research can be instrumental, even life-altering.

Relatively new as a field of study, causal inference (CI) systematically analyzes data that link cause and effect, teasing them out from data that only identify associations that may not be due to a cause-and-effect relationship. The aim is to move from conclusions based on association to conclusions based on causation, in the process translating the inferences into probability calculations.



2013 Nathan Mantel Lifetime Achievement Award in Statistics and Epidemiology recipient, Dr. James Robbins

An exploding field, CI in epidemiology was kick started into existence about three decades ago by the maverick thinking of James Robins, MD, the Mitchell L. and Robin LaFoley Dong Professor of Epidemiology at Harvard School of Public Health's Departments of Epidemiology and Biostatistics. The field's eminence gris, Robins was the first to systematically apply statistical models to data on real exposures that may have downstream effects over time, which accounting methods had not previously factored. Traditional statistical analyses, not designed to track treatment changes over time, weren't able to calculate the effect of confounding variables–– a doctor in a study independently adjusting medication dosage, for example. Counterfactual 'what if' scenarios ––what would happen, for instance, if a particular therapy was given continuously rather than intermittently––also were left out of the equation.

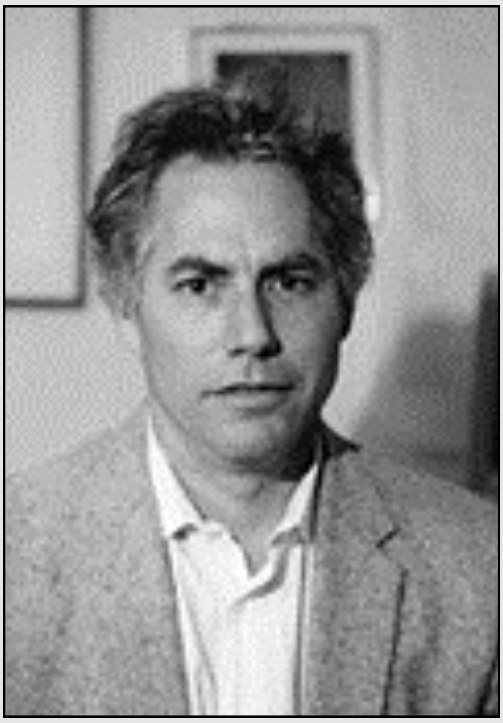
These shortcomings clouded researcher's ability to discern cause and effect. Robins' causal inference work has lifted the clouds dramatically. The complex methodology, which gives researchers more accurate estimates of uncertainty and more precise analytics, has since skyrocketed, affecting how researchers in vastly disparate disciplines do their research to devise better real-world decisions.

Dr. Miguel Hernan MD, DrPH, Chair-Elect, Section on Statistics in Epidemiology, American Statistical Association, 2014

One pointed example from the late 1990s involves anti-retroviral therapy. Robins's thinking critically informed the study, which was led by HSPH colleague Miguel Hernán, MD, DrPH, an internationally known CI leader who collaborates closely with Robins. Randomized studies had proven that combined anti-retroviral therapy should be the treatment choice for HIV-infected people. Observational studies, however, gave conflicting results.

Observational research, the vast majority of epidemiology studies, folds *a priori* assumptions into its statistical calculations. The more variables, the more difficult to pinpoint what's causing what. Randomized trials compare outcomes in control and test groups who are randomly assigned treatment. Though considered the gold standard, they often are compromised, says Robins. People may not obey their medication regimens. They may be subjected to harmful exposures over time, which are difficult to monitor or control. The trials take a long time to organize and are often unethical when one of the treatments, test or control, are considered suboptimal, as would be the case in withholding HIV therapy for some.

In the anti-retroviral therapy analyses that Hernán and Robins conducted, findings from observational studies indicated the therapy was only minimally effective in slowing the pace of AIDS and death. Robins' statistical wizardry solved the conundrum of conflicting results: When the observational data were analyzed with causal methods, the conflict with the results of the randomized trials disappeared. His methods are now essential in HIV research.



James Robins, MD

Robins' innovations "put causal inference on the map," says Hernán. CI has transformed how questions are framed––the greater the precision in how questions are posed, the more accurate the results––and how observational studies are analyzed, elevating them to a new standard in biomedical research.

Allen Wilcox, MD, PhD and Tyler VanderWeele, PhD at Dr. Wilcox's Department of Epidemiology seminar event

That standard, nationally and internationally, has been boosted by the first-rate work HSPH's CI group's. Among others, it features Professor of Epidemiology Tyler VanderWeele, PhD, a leader in developing methods to assess causal inference mechanisms, and Eric Tchetgen Tchetgen, PhD, Associate Professor of Biostatistics and Epidemiologic Methods, who has made seminal advances in longstanding biostatistical problems in causality.

Moyses Szklo, MD, DrPH, MPH, Editor-in-Chief of the American Journal of Epidemiology, speaking at the Spring 2013 Cutter Lecture

"[The team] is without doubt one of the leading groups in the world working on [CI]," says Allen Wilcox, MD, PhD, who helped establish the National Institutes of Health epidemiology research program and serves as Editor-in-Chief of Epidemiology, one of the discipline's flagship journals. The sentiment is echoed by Professor of Epidemiology and Medicine Moyses Szklo, MD, DrPH, MPH, of The Johns Hopkins University and Editor-in-Chief of another flagship journal, the American Journal of Epidemiology: The team's "landmark papers represent a benchmark for all clinical and population researchers interested in translating results of observational studies to public health and medical policies or treatment protocols."

HSPH's CI group is the largest in the world, offering the most extensive training of any epidemiology department anywhere. It gives courses not only on how to conduct research, but also on how to do research on research. A host of disciplines beyond statistical epidemiology––economics, psychology, law, computer science, sociology and more––are honing their own CI methods in line with the work done by the HSPH team.

To spread the understanding and reach of CI's complex methodology, Hernán and Robins are comprehensively collecting CI material, now scattered among numerous journals and confined to technical articles, into a book, Causal Inference. Though the book is not yet complete, universities across the world already are using the draft, publicly available online, to aid their teaching. Likewise with the material collected in VanderWeele's Explanation in Causal Inference: Methods for Mediation and Interaction. Both titles are triggering dozens of workshops, short courses, and lecture invitations.

The team's landmark work is all over the map, subject-wise. Whether focused on gene-environment interactions on disease rates or determining risk factors for ovarian cancer and low birth-weight infants, all the research is aimed at improving peoples' lives.

Dr. Miguel Hernán at the Fall 2013 Department of Epidemiology Cutter Symposium

One of Hernán's projects––revealing the efficacy of a costly dialysis treatment medication––has done just that. He and his collaborators at the Medical Technology Practice and Patterns Institute turned the spotlight on the drug Epoetin (Epo), given to dialysis patients with chronic kidney failure. Perhaps best known as the drug Lance Armstrong took to boost his red blood cell level, Epo was clearly effective in combatting renal disease and approved without many efficacy studies.

Medicare essentially covers all dialysis costs, including Epo. By 2007 Medicare was spending $2 billion a year on the drug ––the single most expensive medication on Medicare's roster. Until recently, Epo's manufacturer offered dialysis centers a hefty discount for Epo use, while dialysis centers were allowed to charge Medicare the list price, giving dialysis centers a substantial profit on every Epo unit administered. "We showed that for-profit dialysis centers were administering 30% more Epo than non-profit centers," says Hernán.

Observational data were difficult to analyze, since dialysis treatments change weekly or monthly based on a patient's hematocrit levels, and varying Epo exposure affects those levels. "When we applied causally-explicit methods, we found that high Epo doses were not helping patients. It wasn't clear they were deleterious, but they weren't beneficial," reports Hernán.

He and his colleagues published numerous papers and testified before the FDA. In addition, subsequent randomized trials on cancer patients proved Epo to be harmful. The upshot, says Hernán: "Medicare changed the way Epo is reimbursed. Now dialysis centers are paid a fixed amount for Epo use, regardless of dose."

Eric Tchetgen Tchetgen, PhD

A disparate study on HSPH's CI spectrum is the AIDS research in Africa spearheaded by Tchetgen Tchetgen, who is particularly concerned about the inferences laced into decisions doctors make on behalf of patients and policymakers. "I look at data that are rich and messy. Often they aren't collected for research and incorporate a lot of biases. I want to tease apart signal from noise and discover the actual scientific message," he explains.

A case in point centered on a household survey conducted as part of the Zambia Demographic and Health Survey to determine the country's HIV prevalence––information international aid agencies use to determine resource allocation. Zambia's surveyors had gone house to house asking people 16 years and older about their willingness to be tested for HIV. "It wasn't a good design," says Tchetgen Tchetgen. "The refusal rate was as high as 30%, which could have been due to their HIV status, stigma, or other reasons, and varied from one interviewer to the next." Understanding the data collection process and other available information, his team applied CI statistical methods to account for selection bias. They discovered Zambia's actual HIV incidence was 22%, as opposed to the 12.5% previously reported, a disparity that could have a large impact on resources given to fight HIV/AIDS. Says Tchetgen Tchetgen: "The work we do might be mathematical and statistical, but we care passionately about epidemiological issues," adding that "as an African, where the prevalence of AIDS can be as high as 30% in certain communities, I feel particularly responsible to do something about it with whatever skills I have."

In other HIV-related work, Tchetgen Tchetgen looked at analytic methods used when data is missing due to death, with a specific focus on assessing the impact of HIV on birth outcomes. HIV leads to an increased risk of stillbirths, and if still births are not appropriately accounted for in birth outcomes research, the data become skewed. His studies showed the effect of HIV on birth outcomes is much worse than previously found. The novel methodology he devised offers a solution to the difficulties of survival bias research and enables the data and analyses to be handled effectively.

Tyler VanderWeele, PhD

In other arenas, VanderWeele focuses on causal mechanisms––what, in a complex world with numerous variables, drives a specific effect? In one study, he is investigating former child soldiers in Sierra Leone. Not surprisingly, their struggles with mental health, employment, and community acceptance are rife. Collaborating with Theresa Betancourt, ScD, of HSPH's Department of Global Health and Population, VanderWeele was asked to help determine causality. "What's driving these problems? If we had to change something, what would most powerfully affect outcomes in later life? Would it be focusing on emotional problems? On acceptance from former family or community? School interventions?" The interim finding, says VanderWeele, points to changing the child/soldiers' self-perception and how they handle their emotions internally. "Unfortunately," he says, "this is one of the more difficult things to change."

Another of his studies involved a randomized trial on cognitive behavior therapy (CBT), which found that CBT had beneficial effects on depression. But the researchers also noticed that after three months, the use of anti-depressants was higher in the group practicing CBT. Skeptics said the reduction in depression was due not to CBT, but to compliance with the medication and that treatment should focus on compliance, not CBT.

VanderWeele was asked to help determine CBT's role. He saw that while CBT use was randomized, anti-depressant use was not. The study analysis had introduced confounding variables, since anti-depressant use could be caused by other factors. Analyzing data from multiple trials to handle several confounding relationships, VanderWeele found CBT to be beneficial not just because it encouraged medication compliance, but also because it was useful in and of itself. The result provides important therapeutic evidence to people considering the treatment.

In another project, VanderWeele has been looking at how religion and spirituality affect health, both in the general population and in end-of-life care. He is assessing whether the causal mechanisms include social support, lifestyle and behavior, prayer, optimism, belief, and self-discipline. In the end-of-life care setting, he, along with researchers at the Dana Farber Cancer Institute, have found that spiritual care within medicine often strongly shapes end-of-life decisions, but is rarely given. His research provides important guidance on how spirituality should be incorporated into care for the terminally ill.

Whatever the focus––depression, dialysis, AIDS or end-of-life––the CI group has seen its methodologies become increasingly important to research in epidemiology and beyond. The CI mindset propelled by the team––how to think, how to ask questions, how to embrace life's messy complexity–– how, in essence, to get more real–– has helped CI research become more systematized and structured across disciplines. Social sciences, life sciences, computer sciences, artificial intelligence, economics, law and many other fields now speak a unified language, allowing the best brains from each to cross pollinate. The cross talk has expanded by quantum leaps how to understand the causal forces affecting peoples' lives and how, ultimately, to make better real-world decisions.

-Orna Feldman

HARVARD
SCHOOL OF PUBLIC HEALTH