Interventions, i.e. deliberate actions to drive change, are fundamental to improve people’s health and well-being, to advance the economy, to protect the environment, or to better any other aspect of life. As long as humans have existed, they have gained understanding of the world (and the universe!) through observations. At the micro level, a child learns about the (potential) consequences of actions through continuing observations, complemented by trial and error to consolidate its own imaginary ecosystem. In the long history of science, observational data have played a driving role in establishing fundamental laws and theorems relevant to human life and nature. In the post-2nd World War period, the role of observational studies was supplemented as the randomized trial design took prominence for establishing causation (Bothwell and Podolsky, 2016; Marks, 2000). However, randomized trials are not always feasible or ethical. For example, studying the health effect of radon exposure through an experimental study would be an impractical and infeasible endeavor. Further, when the research question concerns the long-term effects of an exposure, randomized trials would take many years thus precluding implementation of effective interventions in the near future. Using historically obtained or surveillance observational data to provide meaningful information on causal effects would address these important limitations.

Observational data have always been considered appropriate to make predictions about the world we are in as long as the world doesn’t change too quickly. For example, most athletes with a knee injury would like to know how long it will take before they can get back to training and sport. This information can usually be obtained from appropriate injury surveillance studies. In contrast, one may also want to use observational data to address “causal questions”. Unlike predictive questions where the world doesn’t change, a causal
question is specifically about predicting the future after we change the world through an intervention such as an economic policy change, health initiative, political strategy or so on. This is a much more difficult problem. Causal inference requires assumptions from outside of the observed data. Therefore, one’s training, expertise and previous experience will greatly influence what assumptions are considered plausible, and how they are used in the modelling and inference.

When we study how the world changed after observed real-life actions in most contexts, we recognize that other processes (aside from the actions of interest) could have led to the observed changes. How do we “control” for these other possible causes to determine if the real-life actions were indeed responsible for the observed changes? Formalizing causal inference in observational studies only began in the 1970s and developed simultaneously mainly in four fields: Statistics, Epidemiology, Computer Science, and Econometrics. Because causal inference methods were developed in several disciplines that did not often communicate with each other, differences in approaches may sometimes appear ideologically divisive. This has created some confusion.

From our perspective, the apparent conflicts were likely related to fundamental differences between the respective fields’ philosophical approaches to science, and about the types of data observed within each discipline. To gain further insight, we launched a project to better understand the different perspectives. This special issue brings together four interview-style papers from a leader in each of the four fields where causal inference developed.

Don Rubin was formally trained in statistics. He approached causal inference from a Bayesian Decision Theoretic perspective, where one characterizes uncertainty in the data sources via the most useful parsimonious models. His contributions to causal inference include the Rubin Causal Model, propensity scores and principal stratification.

Jamie Robins was formally trained as a physician. He started with data and wanted to develop methods using as few assumptions as possible, even when treatment and confounding factors varied over time. He created a unified theory of semiparametric estimation for missing data and causal inference that changed the way we approach observational studies with time-varying confounding. In 2022, he shared the inaugural Rousseeuw Prize for Statistics for his body of work and collaborations in causal inference.

Judea Pearl (computer science) won the ACM A.M. Turing Award for fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning. Pearl begins with the structure of the data generating process (as opposed to observed data), and asks what aspects of that process (usually in the form of directed acyclic graph) and what types of data would allow identification of causal or counterfactual questions a scientist may wish to answer.

James Heckman (econometrics) won the Nobel Memorial Prize in Economics for his work on developing and promoting causal inference in the field of economics. The Heckman correction approached confounding bias as a model specification problem with respect to the model errors (rather than model predictors). This allows one to make only moment assumptions, rather than conditional independence assumptions.

These interviews provide a behind-the-scenes understanding of each person’s perspective of the history of causal inference, some personal experiences in their scientific career, underlying motivation for their research studies, as well as a description of the importance of their
different achievements. We believe the stories and insights told by the interviewees in this series will help us learn from the past to improve the future. There are important commonalities in the interviews with respect to general approaches to the scientific method. Each interview also has specific lessons given the nature of the interviewee’s research interests. Together, the four interviews provide a breadth of experience that forms the foundation from which causal inference continues to gain strength and momentum. Finally, the references provided in each interview reflect an excellent bibliography for those interested in specific topics, or from a historical / academic perspective in general. We hope you enjoy reading the interviews and learn as much from them as we did.
References
