Eurocim 2025

Data fusion, generalisability, transportability

Causal Record Linkage: Critical Issues and Novel Approaches to False Discovery Propagation

From the idea of assembling a book of life for individuals to support public health decisions, record linkage emerged. It evolved to empower research by integrating information from diverse sources, such as electronic patient records and observational studies.

Today, it offers an opportunity to control for confounding variables in causal inference, to study long term outcomes by linking longitudinal data, or else to explore secondary outcomes which were not measured in the first place.

In this investigation, we review the perspectives taken on causal record linkage. Specifically, we examine uncertainty propagation from the linkage to the treatment effect estimation and we survey the two stage and joint approaches which are being used in applications to economics, banking, medicine. We give an account on the criteria addressed in recent developments and we evaluate the potential of a novel algorithm to control the rate of false links for causal effect estimation. Finally, we discuss the remaining challenges and outline potential directions for future causal research using linked data.