Revisiting Sequential Attributable fractions

CURRENT STATUS: UNDER REVISION

Archives of Public Health BMC

John Ferguson
National University of Ireland Galway

john.ferguson@nuigalway.ie*Corresponding Author*
ORCiD: https://orcid.org/0000-0001-8475-398X

Maurice O'Connell
National University of Ireland Galway

Martin O'Donnell
National University of Ireland Galway

DOI:
10.21203/rs.2.22051/v1

SUBJECT AREAS
Statistical Epidemiology

KEYWORDS
attributable fraction, causal DAG, do-operator, Bayesian network, causal inference

Abstract

Background: Eide and Gefeller [1] introduced the concepts of sequential and average attributable fractions as methods to partition the risk of disease to differing exposures. In particular, sequential attributable fractions are interpreted in terms of an incremental reduction in disease prevalence associated with removing a particular risk factor from the population, having removed other risk factors. Clearly, both concepts are causal entities, but are not usually estimated within a causal inference framework.

Methods: We propose causal definitions of sequential and average attributable fractions using the potential outcomes framework. To estimate these quantities in practice, we model exposure-exposure and exposure-disease interrelationships using a causal Bayesian network, assuming no unobserved variables. This allows us to model not only the *direct* impact of removing a risk factor on disease, but also the *indirect* impact through the effect on the prevalence of causally downstream risk factors that are typically ignored when calculating sequential and average attributable fractions. The procedure for calculating sequential attributable fractions involves repeated applications of Pearl's do-operator over a fitted Bayesian network, and simulation from the resulting joint probability distributions.

Results: The methods are applied to the INTERSTROKE study, which was designed to quantify disease burden attributable to the major risk factors for stroke. The resulting sequential and average attributable fractions are compared to results to a prior estimation approach which uses a single logistic model and which does not properly account for differing causal pathways.

Conclusions: In contrast to estimation using a single regression model, the proposed approaches allow consistent estimation of sequential, joint and average attributable fractions under general causal structures.

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the manuscript can be downloaded and accessed as a PDF.

Figures

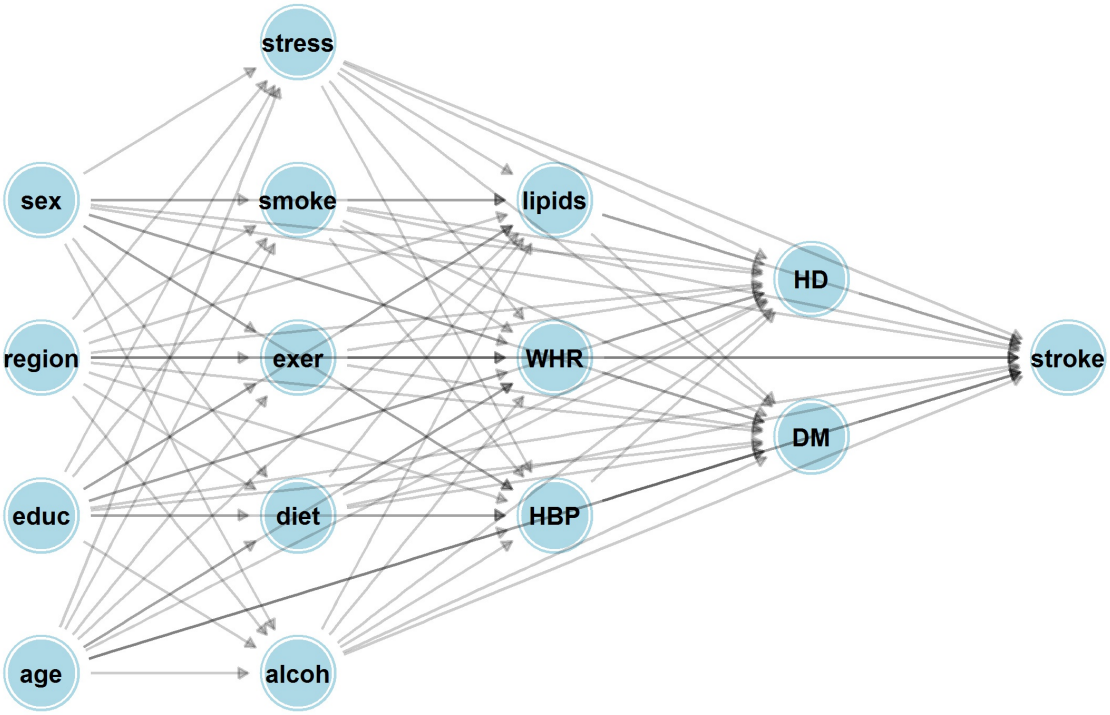

Figure 1

Hypothesized causal Bayesian network describing causal risk factors of stroke

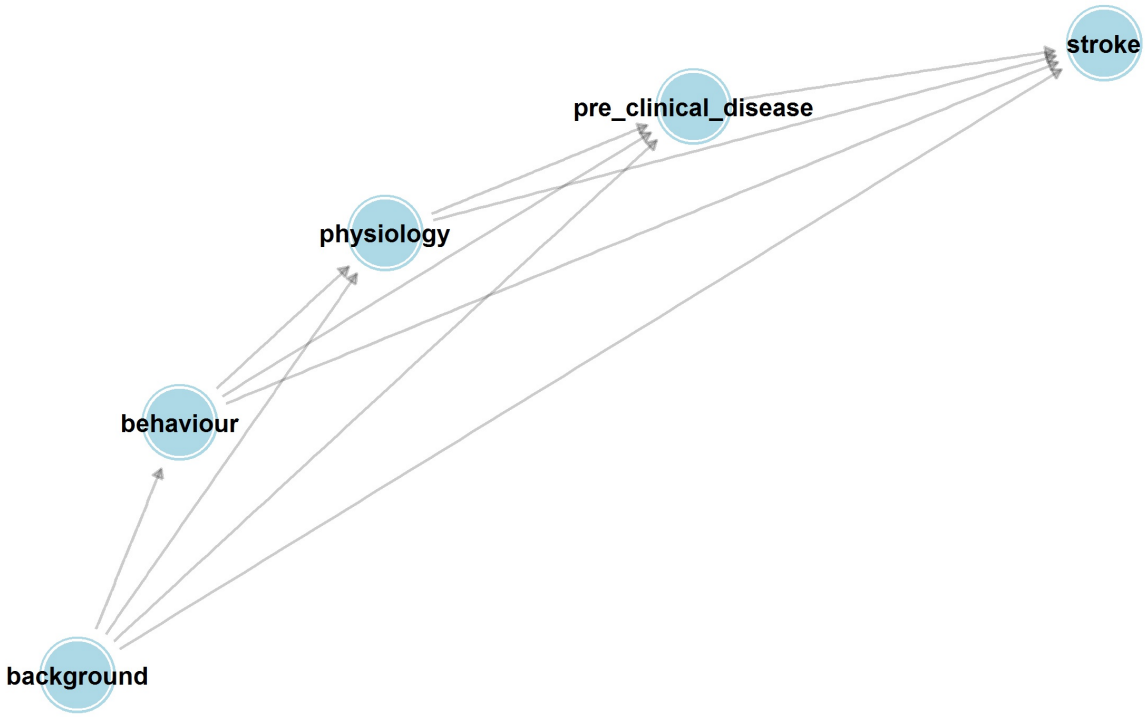

Figure 2

Simplication of network from Figure 1, showing it's layered structure

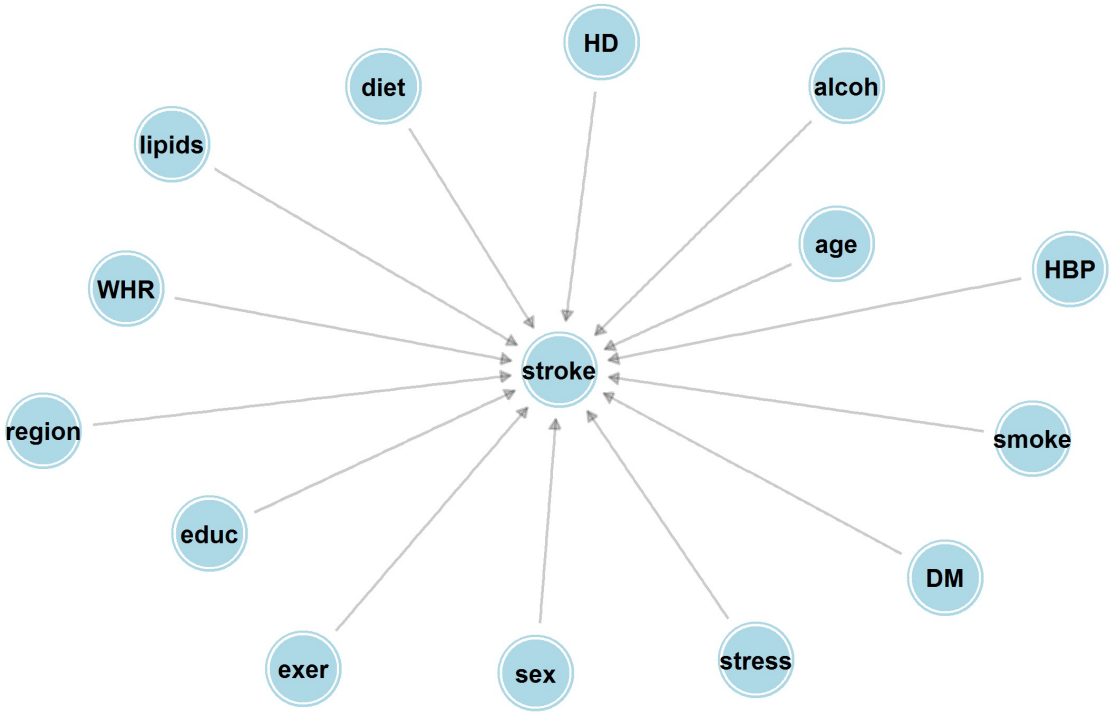

Figure 3

Bayesian network with only direct effects

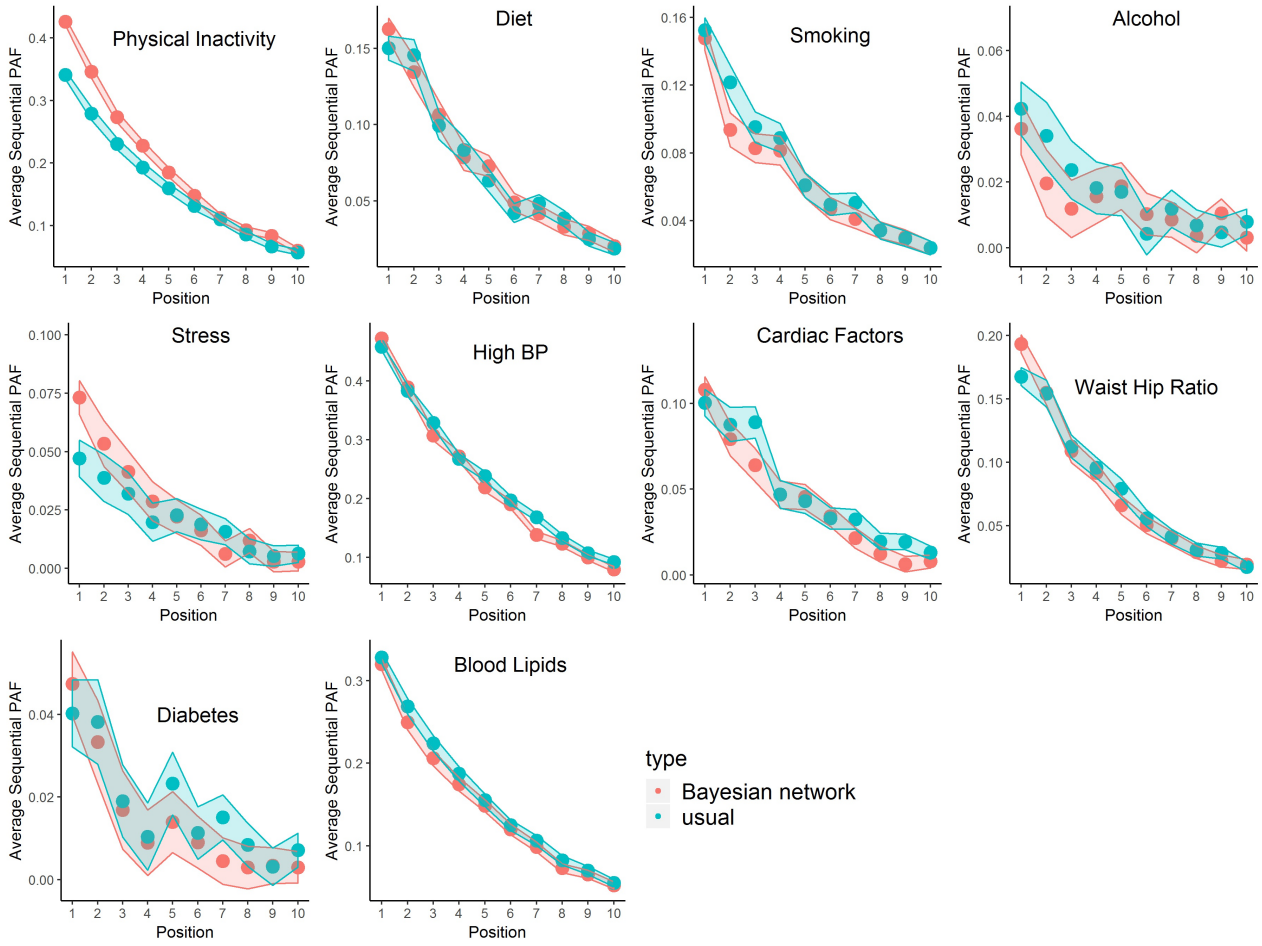

Figure 4

Estimated sequential attributable fractions, by position in elmination order. We can be 95% condent the true estimate (that would be calculated from the procedure when the number of simulations m → ∞) lies in the Monte Carlo interval around the point estimate. The estimates shaded red correspond to the Bayesian network in Figure 1, whereas the estimates shaded blue correspond to the Bayesian network in Figure 3. Note that the Monte Carlo error at position k incorporates variation due to random selection of the set of risk factors/exposures that are intervened on in stages 1,...k - 1, and also variation based on the recursive simulation of the disease response described in the main text