Environment International 102 (2017) 30–41

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Causal inference in cumulative risk assessment: The roles of directed acyclic graphs

L. Elizabeth Brewer [a], J. Michael Wright [b,*], Glenn Rice [b], Lucas Neas [c], Linda Teuschler [d]

[a] Oak Ridge Institute for Science and Education (ORISE), U.S. Environmental Protection Agency, Office of Research and Development, Office of the Science Advisor, 1300 Pennsylvania Ave., NW, MC8195R, Washington, DC 20004, United States
[b] U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, 26 W. Martin Luther King Dr., MS-A110, Cincinnati, OH 45268, United States
[c] U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, B305-01, Research Triangle Park, NC 27711, United States
[d] LK Teuschler and Associates, St. Petersburg, FL 33707, United States

ARTICLE INFO

Article history:
Received 15 March 2016
Received in revised form 18 November 2016
Accepted 6 December 2016
Available online 14 December 2016

Keywords:
Cumulative risk assessment
Conceptual model
Directed acyclic graph
Causal inference
Confounding
Causal models

ABSTRACT

Cumulative risk assessments (CRAs) address exposures to multiple chemical and nonchemical stressors and often focus on characterization of health risks in vulnerable populations. Evaluating complex exposure-response relationships in CRAs requires the use of formal and rigorous methods for causal inference. Directed acyclic graphs (DAGs) are graphical causal models used to organize and communicate knowledge about the underlying causal structure that generates observable data. Using existing graphical theories for causal inference with DAGs, risk analysts can identify confounders and effect measure modifiers to determine if the available data are both internally valid to obtain unbiased risk estimates and are generalizable to populations of interest. Conditional independencies implied by the structure of a DAG can be used to test assumptions used in a CRA against empirical data in a selected study and can contribute to the evidence evaluations related to specific causal pathways. This can facilitate quantitative use of these data, as well as help identify key research gaps, prioritize data collection activities, and evaluate risk management alternatives. DAGs also enable risk analysts to be explicit about sources of uncertainty and to determine whether a causal effect can be estimated from available data. Using a conceptual model and DAG for a hypothetical community located near a concentrated animal feeding operation (CAFO), we illustrate the advantages of using DAGs for evaluating causality in CRAs. DAGs also can be used in conjunction with weight of evidence (WOE) methodology to improve causal analysis for CRA, which could lead to more effective interventions to reduce population health risks.

1. Introduction

Cumulative risk assessments (CRAs) examine, characterize, and quantify the combined risks to human health or the environment from exposures to multiple stressors, potentially including chemical, physical, biological, and social stressors (Lokke, 2010; Meek et al., 2011; NRC, 2009; U.S. EPA, 2003). Development of methods to conduct CRAs is responsive to the multifactorial nature of many human diseases and concerns that evaluation of multiple chemical and nonchemical stressors, including complex chemical mixtures, and vulnerabilities on a community relevant scale are incompletely addressed by traditional environmental risk assessment approaches (Callahan and Sexton, 2007; NEJAC, 2004; NRC, 2009; Sexton, 2012).

Deriving accurate and unbiased estimates of causal effects is critical to CRAs and other risk analyses. Experimental studies (e.g., randomized controlled trials) are considered the most reliable types of studies for causal inference and allow the researcher to interpret the effects of intervention in the treatment group as the true causal effect of the intervention. However, for quantifying risk to multiple stressors, data from observational epidemiological studies of environmentally relevant exposures are often more abundant. When observational epidemiological evidence is available, quantitative use of these data including extrapolation to other populations can be complicated by a variety of issues including variation in population characteristics and unmeasured alternative causal pathways caused by confounding or other sources of bias. Among modern causal inference methods, such as the Neyman-Rubin Model (i.e., Potential Outcomes Framework; Holland, 1986) and

Abbreviations: CRA, cumulative risk assessment; DAG, directed acyclic graphs; CAFO, concentrated animal feeding operation; SES, socioeconomic status; MOA, mode of action; AOP, adverse outcome pathway; WOE, weight of evidence; NRC, National Research Council; USEPA, U.S. Environmental Protection Agency.

* Corresponding author.
E-mail addresses: brewer.beth@epa.gov, bethbrew@gmail.com (L.E. Brewer), wright.michael@epa.gov (J.M. Wright).

http://dx.doi.org/10.1016/j.envint.2016.12.005

Pearl's (2009) Structural Causal Model, Pearl's approaches for graphical annotation of causal relationships, that unify graphs, counterfactuals, and the potential outcome framework, appear well-suited to CRA because they can depict complex relationships and serve as visual guides to address these sources of uncertainty.

When conducting CRAs, developing well-constructed graphical causal models is critical for evaluating causality and considering risk management opportunities (e.g., intervention points in the pathways from sources of stressors, to exposures, to health outcome/s). Directed acyclic graphs (DAGs), a type of graphical causal model, are comprised of nodes connected by unidirectional arrows containing no paths that form a cycle (Glymour, 2006; Greenland et al., 1999). Other types of graphical causal models, including conceptual or theoretical models, are DAGs if they consist of clearly directed causal paths without cycles.

DAGs provide opportunities for notating causal assumptions about statistical associations and determining whether relationships are causal. The available evidence conveyed in a DAG may vary from qualitative to quantitative; the arrows drawn between the nodes may represent expert opinions, statistical associations derived from epidemiological or toxicological studies, or a priori knowledge of cause-effect relationships including mode of action (MOA) processes. In conjunction with weight of evidence (WOE) approaches (Rhomberg et al., 2010; Weed, 2005), DAGs can be used in CRAs to enhance causal analysis by identifying sources of bias, enumerating the testable implications of the causal diagram, and identifying if causal effects and potential impacts of an intervention are estimable from available data (Pearl, 2009). Formal methods for evaluating causal hypotheses are recognized as important to support risk-based decisions (Linder et al., 2010). These include graphical causal modeling and conditional independence testing (described in this article), Granger causality tests, panel data analysis, and intervention analysis, among others (Cox, 2013).

This paper explores potential uses of DAGs for causal inference within CRAs and provides a starting point for their practical application by risk analysts. We describe existing usage of causal models in CRA, and then describe how to integrate DAGs as causal models throughout a CRA. To highlight the types of reasoning and information needed to develop DAGs for CRAs, we initially constructed a conceptual model and then developed a DAG for a hypothetical community near a concentrated animal feeding operation (CAFO) using associations identified from the published literature. While this article focuses on the uses of DAGs in CRAs, we realize that other analyses may utilize DAGs for other purposes and have different considerations (e.g., DAGs supporting epidemiological studies). Causal modeling and its applications in social science, psychology, economics, and epidemiology has been described elsewhere (Glymour, 2001; Greenland et al., 1999; Morgan and Winship, 2014; Pearl, 2009; Spirtes et al., 2001).

2. Existing usage of graphical causal models in CRA

Graphical causal models are graphs that depict background knowledge, previously established theories, hypotheses about causal structures, and mechanisms that may be altered through an external change or intervention. These graphs may be accompanied by a set of mathematical equations derived from empirical data (Pearl, 2009; Russo, 2009). This section describes their current and potential usage in the context of CRA.

The U.S. Environmental Protection Agency (USEPA) (U.S. EPA, 2003) describes the following three phases of CRAs: 1) planning, scoping, and

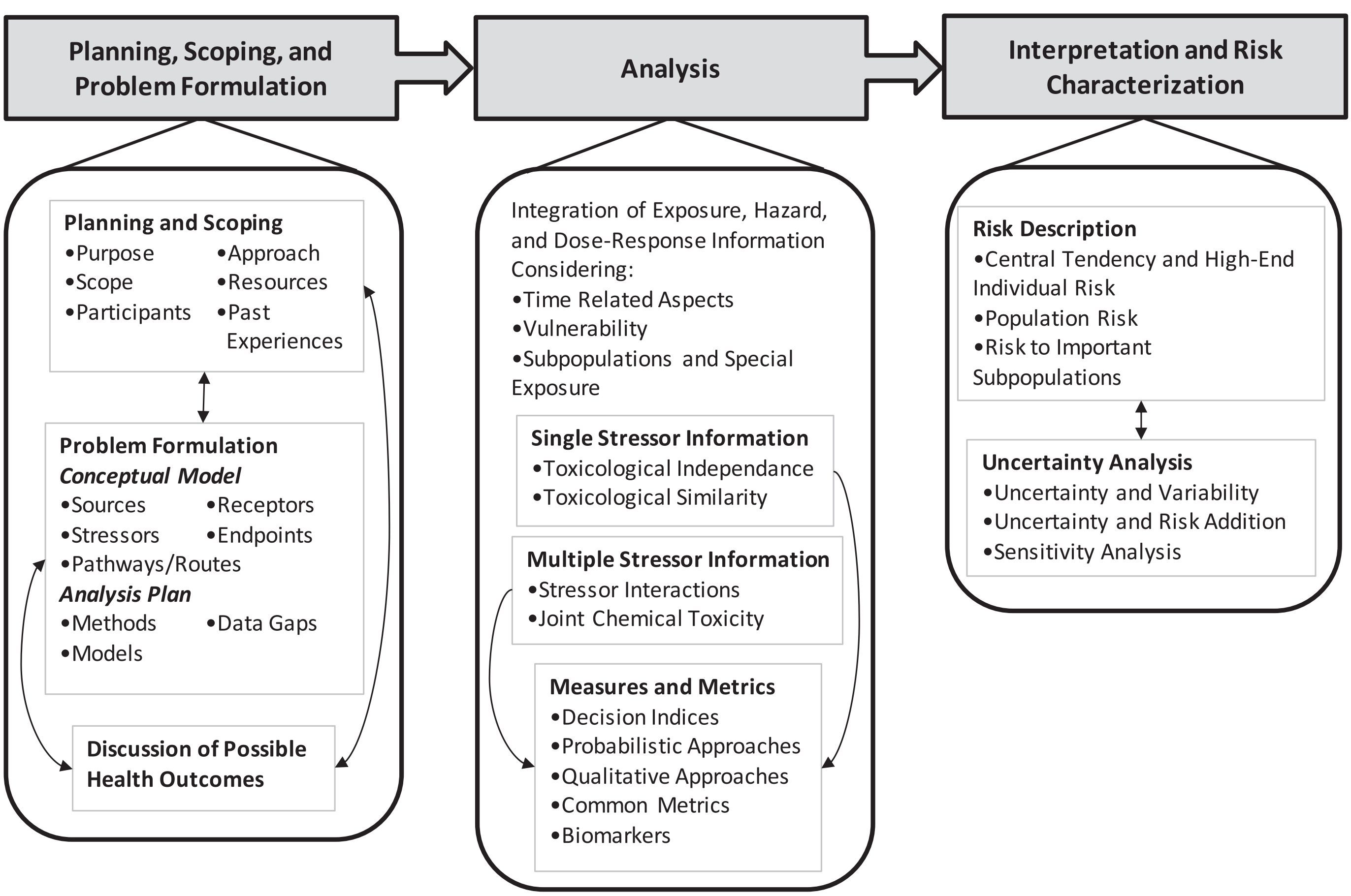

Fig. 1. Framework detailing three main phases of CRAs.
Adapted from U.S. EPA (2003).