# Highlights

**Causal diagrams and mediation analysis for estimating the effect of technological interventions on cardiac arrest survival**

Jacob Hutton, Boris Sobolev, Joseph Puyat, Jim Christenson, Brian Grunau

- Causal diagrams can be used to outline the potential impact of an as-yet untested intervention.
- For interventions that may affect outcome through intervening on a mediator, combining a mediation analysis with causal diagrams is a powerful and intuitive methodology.

# Causal diagrams and mediation analysis for estimating the effect of technological interventions on cardiac arrest survival

Jacob Hutton, Boris Sobolev, Joseph Puyat, Jim Christenson, Brian Grunau

[a]*British Columbia Resuscitation Research Collaborative UBC Faculty of Medicine Vancouver BC Canada*
[b]*School of Population and Public Health UBC Faculty of Medicine Vancouver BC Canada*
[c]*School of Population and Public Health UBC Faculty of Medicine Vancouver BC Canada*
[d]*British Columbia Resuscitation Research Collaborative UBC Faculty of Medicine Vancouver BC Canada*
[e]*British Columbia Resuscitation Research Collaborative UBC Faculty of Medicine Vancouver BC Canada*

**Abstract**

A common structure for questions in public policy involves the asking of a ”what if” question. ”If we implemented intervention A in the population, what would be the impact?” However, such queries cannot be answered utilizing by data alone. Instead, researchers must consider counterfactuals to compute the likely effect of an intervention in the population of interest. However, there is a second problem that requires solving prior to being able to answer these queries. Many interventions impact the outcome indirectly. They act on a mediating variable between the exposure and outcome itself. Therefore we cannot make a purely data-based claim about the mechanism of effect a priori. Instead, we must reason about the likely impact of a proposed intervention in a transparent and accessible way, such that it can be evaluated by critical readers and debated among collaborators. This methods paper details an application of a methodological approach to solving these issues using a mediation analysis informed by causal diagrams. The approach is then applied in estimating the potential impact of technology interventions in out-of-hospital cardiac arrest (OHCA). Our approach has utility for researchers and developers seeking to estimate the potential impact of hypothetical interventions that seek to impact variables in the causal path between exposure and outcome.

*Keywords:* causal inference, mediation analysis, directed acyclic graphs, biotechnology

## 1. Introduction

In biomedical research, the highest standard of assessing causal effect in a hypothesized causal system remains direct intervention in an experimental framework. [1] However, for proposed interventions that have not yet reached sufficient technological maturation for real world experimentation, it is necessary to use existing observational data to estimate the likely effectiveness of these technologies.[2] Accordingly, a methodology has been developed to perform causal inference in datasets that do not include measurement of the intervention of interest. This methodology utilizes existing data-based relationships in conjunction with causal graphs to propose the likely mechanism of action of the intervention of interest.[3] These questions become more complex when the target action of the intervention in question is exerted on a variable within the causal chain between exposure and outcome, and an informed approach is necessary to transparently and reproducibly answer these questions.

In this paper, we provide necessary context on two pillars of causal inference methodology: graphical causation with Directed Acyclic Graphs (DAGs) and mediation analysis. We propose a method to fully eludicate the causal system of key factors in observational data, and how to apply structured abstractions to that system to approximate the contribution of several independent mediating pathways to the exposure-outcome relationship using a mediation analysis. Once this has been established, we apply this methodology to an example using observational data from a clinical registry for cardiac arrest.

## 2. Background

### *2.1. Primer on Directed Acyclic Graphs (DAGs)*

Directed Acyclic Graphs (DAGs) are a common tool for illustrating dependencies among variables in causal literature. In these graphs, "nodes" represent the variables, while "edges" depict dependencies and their directions. The simplest DAG involves two nodes of special importance dedicated to exposure and outcome, and an edge connecting the two (Figure 1).

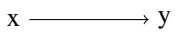


Figure 1: A Directed Acyclic Graph (DAG) of nodes $x$ and $y$.

As Figure 1 depicts, a connecting arrow between nodes $x$ and $y$ indicates that there is a causal connection between the two nodes that is directional and non-cyclic. In other words, $x$ is a direct cause of $y$, and the lack of any other nodes in the graph indicates that the relationship is not influenced by any other variables. In this simple DAG, we may infer that since $x$ is the only parent node of $y$, $x$ is necessary for $y$ to occur.

In addition to the two nodes of special importance depicting exposure and outcome, we may depict the influence of other variables relevant to the relationship of interest. Common causes, and common effects, are two particularly important types of nodes.

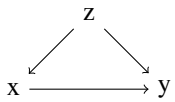


Figure 2: A Directed Acyclic Graph (DAG) of nodes $x$ and $y$ that are both influenced by the common cause (confounding variable) $z$.

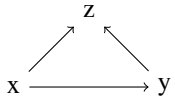


Figure 3: A Directed Acyclic Graph (DAG) of nodes $x$ and $y$ that both exert a causal effect on the common effect (collider) $z$.

In observational data, an association between exposure and outcome manifests as covariation. When the value of the exposure variable changes, the outcome variable is observed to also change. In causal inference, this covariance is a prerequisite but not sufficient for determining

causation. Pearl (1990) outlined that a process of causal attribution can be achieved by conditioning (adjusting, controlling for, stratifying on) on a set of factors that are sufficient for blocking all biasing influence between exposure and outcome. In this context, conditioning on the set z refers to holding values for the variables in set z fixed, either through stratification (if sufficient data is present) or regression to estimate the remaining association between exposure and outcome.

Identification of this adjustment set z is achieved using a causal diagram in the form of a directed acyclic graph (DAG). DAGs are dependency graphs with two nodes of particular importance (exposure and outcome). The network of dependencies in the graph encode our causal assumptions about the nature of the relationship between the exposure and outcome of interest and attempt to outline shared non-causal factors that may also contribute to covariance. We hypothesize a relationship between exposure and outcome using the DAG and use it to deduce the causal nature of a proposed relationship. To make this deduction, we must eliminate all non-causal relationships between exposure and outcome using the specific topology of the DAG as a guide. Once this is completed, the remaining effect can be construed as causal. Figure 1 shows the causal diagram with all known relationships among factors that influence the relationship of interest either directly or through a chain of causal dependencies.

### *2.2. Primer on Mediation*

In addition to direct influence, exposures can influence outcomes indirectly through multiple intermediate nodes in a causal chain.[4] When the effect of an exposure of interest is conveyed to an outcome of interest through impacting variables in the causal chain, this is an example of mediation.

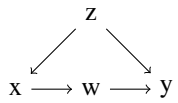


Figure 4: A Directed Acyclic Graph (DAG) of the mediator *w* conveying a causal effect between nodes *x* and *y* that are both influenced by the confounding variable *z*.

Mediation is a common phenomenon in both biological and socioeconomic relationships. In the relationship between *x* and *y*, a portion of the total effect of *x* on *y* is produced through the chain of influence of *x* on *z*, then *z* on *y*, and a separate portion of the total effect is due to the direct action of *x* on *y*.Taken in whole, the effects exerted through all mediating pathways represent the total effect of the exposure on the outcome. While for some analyses, computation of the total effect of an exposure on outcome is sufficient for the research goals, in other cases it is beneficial both for researchers and consumers of the end products of research to understand the mechanisms of causality within the total effect, which can be accomplished through mediation analysis.

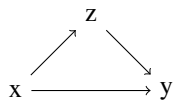


Figure 5: A Directed Acyclic Graph (DAG) of the mediator *w* conveying a causal effect between nodes *x* and *y* that are both influenced by the confounding variable *z*.

We can differentiate between direct effects (not mediated) and indirect effects (exert influence through a mediating chain).[5] The direct effect refers to the response to intervention in the absence of the mediator, while the indirect effect describes the response produced through the mediator. Assuming no unmeasured mediator-outcome confounding, the controlled direct effect can be estimated by conditioning the analysis on the mediator.[6] In the presence of meditator-outcome confounding, advanced statistical packages can be used to estimate a range of mediation parameters.

For a general introduction to mediation analysis, consider a study investigating the effect of an intervention (exposure) on an outcome, and the role of a variable (mediator). In this case, we could estimate several different facets of mediation.[7]

1. **ACME (control)**: The average causal mediation effect for the control group.
2. **ACME (treated)**: The average causal mediation effect for the treated group.
3. **ADE (control)**: The average direct effect for the control group.
4. **ADE (treated)**: The average direct effect for the treated group.
5. **Total Effect**: The combined effect of the direct effect of the intervention and the effect mediated by the mediator on the outcome.
6. **Proportion Mediated (control)**: The proportion of the total effect that is mediated by the mediator in the control group.
7. **Proportion Mediated (treated)**: The proportion of the total effect that is mediated by the mediator in the treated group.
8. **ACME (average)**: The average of the ACMEs for the control and treated groups.
9. **ADE (average)**: The average of the ADEs for the control and treated groups.
10. **Proportion Mediated (average)**: The proportion of the total effect that is mediated by the mediator for the general population, taking into account both control and treated groups.

## 3. Example: Linking the effect of biosensors to the effect of a bystander witness: a mediation analysis approach

### *3.1. Background*

Our study sought to evaluate the potential impact of increasing recognition of out-of-hospital cardiac arrest (OHCA) using biosensor technologies such as wearable devices or home-based devices that would detect the presence of cardiac arrest and alert health authorities. While it has been a goal of public health investments to increase the proportion of OHCAs that are bystander-witnessed, 75% of OHCA remains unwitnessed.[8] Accordingly, sensor technologies have been proposed as a solution to reduce this proportion and alert responders in a timelier manner. However, at the time of this study, it was as yet unclear how these technologies may impact survival.

However, it is well established that individuals who experience out-of-hospital cardiac arrest (OHCA) benefit greatly from the presence of a bystander witness who is available to alert authorities and provide time-sensitive interventions. Using this existing relationship in our observational data, we constructed a causal model of this relationship between a bystander witness and survival, and posited that a biosensor may be similar (but not equivalent to) the effect of a bytstander witness on survival.

In a recently published study, We sampled 2 years of cases from 2019 to 2020 from a clinical registry of cardiac arrest in our region (British Columbia, Canada) [9]. To estimate the potential

effect of recognition that would be provided by a biosensor, we conducted a mediation analysis using the relationship between a bystander witness and survival. The following sections of this paper will detail the specific methodology used in our study.

### *3.2. Applying Causal Framework with DAGs*

In the case of biosensors, we may hypothesize that the effect of a biosensor is similar to the recognition component of a bystander witness. To estimate this effect using data, we may rely on the impact of a bystander witness and attempt to decompose the total effect into its component effects to isolate the component attributable to recognition only. This is a mediation analysis. Figure 6 outlines this conceptually and characterizes two exclusive pathways through which bystanders may impact survival.

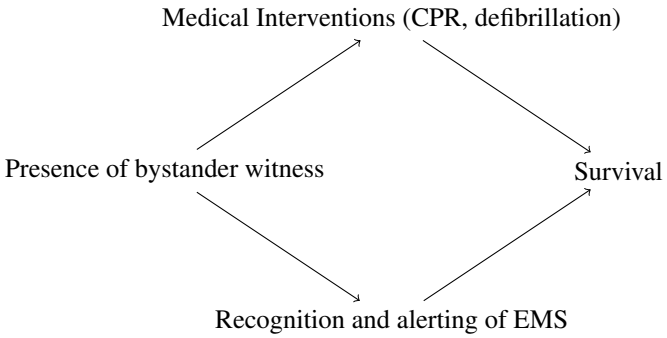


Figure 6: Two pathways for the effect of a bystander witness on out-of-hospital cardiac arrest survival

Figure 7 formalizes this conceptual framework into an exhaustive DAG outlining the exposure, outcome, common causes, common effects, and mediators relevant to the relationship of interest. While this initial DAG explicitly outlines a complex causal chain in the mediating pathway attributable to recognition (path 2), we can apply levels of abstraction to the DAG to allow for easier treatment with existing mediation analysis methods.

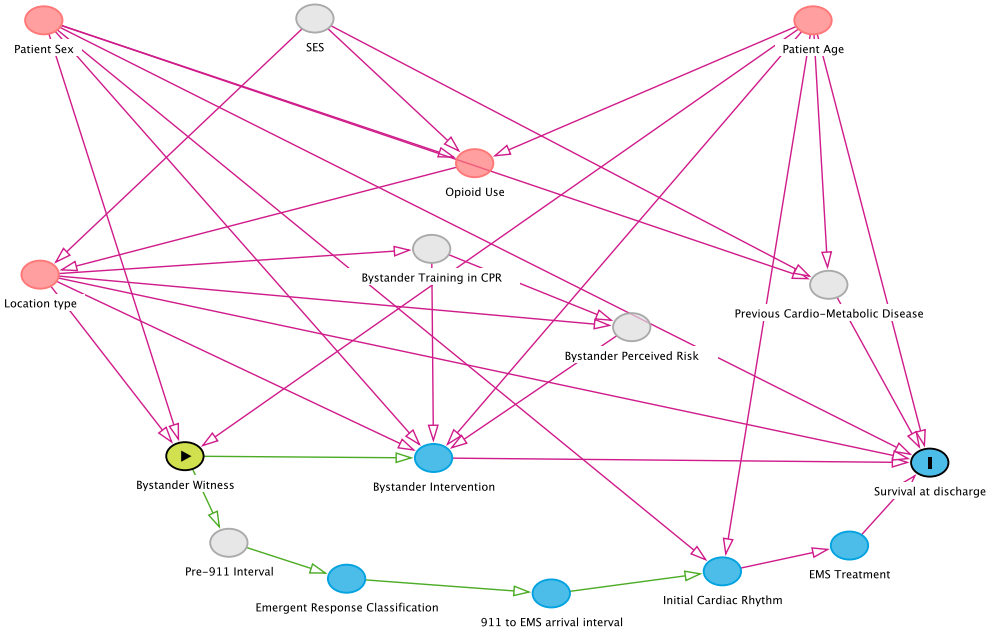


Figure 7: Dependencies among factors involved in producing the association between the presence of a bystander witness and survival to hospital discharge following OHCA. Pink nodes representing the following factors that influence both the exposure and outcome through chains of dependencies (pink edges): sex, location, age. While the full effect of a bystander witness considers all the causal pathways between exposure and outcome, we can also infer from the DAG that there are two mediating causal pathways through which the effect is exerted (1: bystander witness >bystander intervention >survival at discharge; 2: bystander witness >pre-911 interval >emergent MPDS >911-to-EMS interval >Initial cardiac rhythm >EMS treatment >survival at discharge). We can differentiate these two pathways as representing the portion of the effect that is exerted through direct bystander intervention (path 1) and the portion of the effect that is exerted through bystander alerting of EMS (path 2). Note that the pre-911 interval node in path 2 is grey, indicating that it is an unobserved quantity (latent variable) in the BC Cardiac Arrest Registry.

Figure 8 presents an abstracted version of the DAG, collapsing the variables in the recognition path into a single conceptual quantity – recognition.

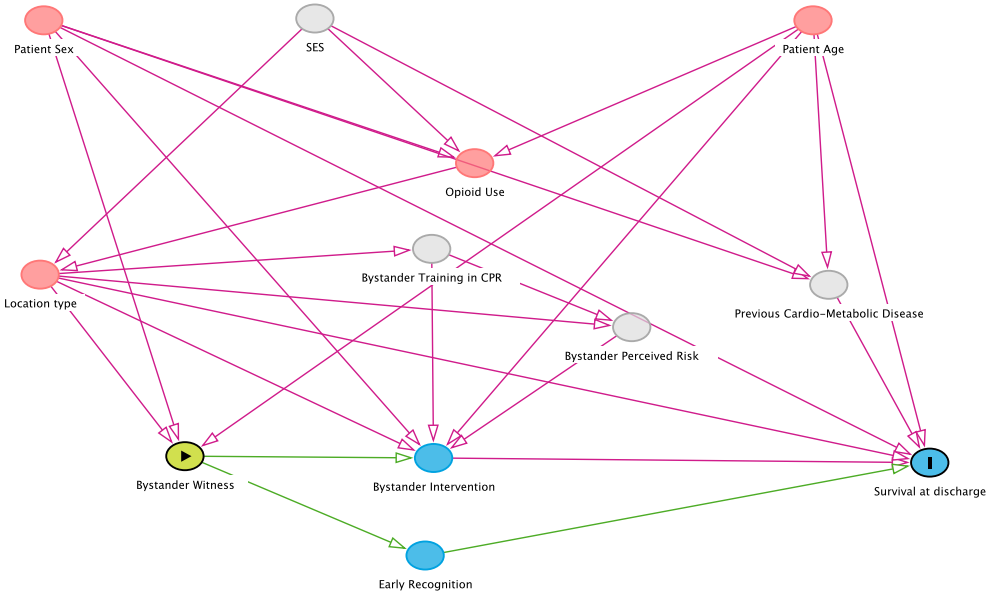


Figure 8: An abstracted version of Figure 2. Mediating factors in the recognition chain have been collapsed into one quantity: early recognition. This represents the factors influencing survival that are attributable to early recognition and response by professional responders to the scene of the cardiac arrest event.

Finally, Figure 9 presents the effect of recognition as a direct effect, with the indirect effect mediated through bystander intervention.

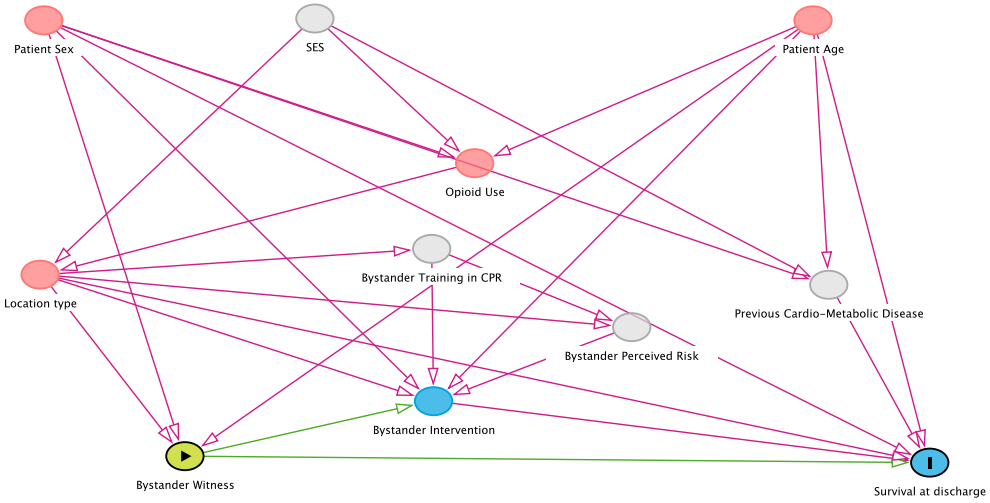


Figure 9: Further abstracting the DAG to represent the effect of recognition as a direct effect. This approach allows for simpler estimation of the direct effect using existing software packages for mediation analysis. By conditioning on the minimally sufficient set of confounders (sex, age, location type), as well as the mediator (bystander intervention), we can estimate the direct effect of a bystander witness. This direct effect represents the portion of the total effect that is not attributable to bystander intervention.

DAGs depict the causal system based on our understanding. Levels of abstraction may vary, contingent upon our specific objectives. Therefore it is critical to be explicit about the meaning of an abstracted effect. When comparing Figure 9 with the original DAG (Figure 7), we observe that five mediating factors have been condensed into a single node, equated to the effect of recognition. In the context of this study, the ”effect of recognition” refers to the cumulative impact of these mediators in path 2. This level of abstraction enables us to utilize existing methods for estimating the indirect effect of bystander intervention as the mediator. This leaves us with the direct effect, which is the portion of the total effect not attributed to the mediator. [10] The core assumption required in this framework is that a bystander witness does not have any causal influence on survival to hospital discharge beyond the influence attributed to either intervention or recognition.

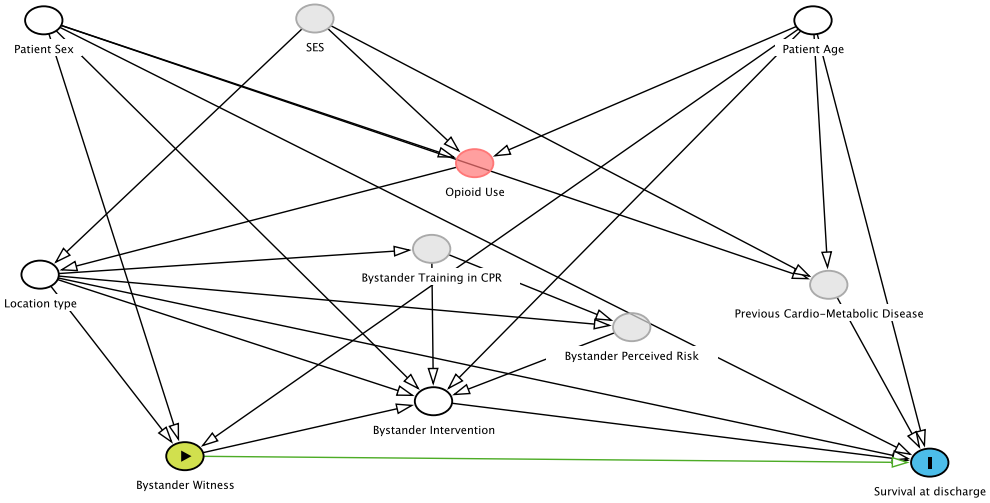


Figure 10: Conditioning on the minimally sufficient adjustment set (patient age, patient sex, location type), as well as the mediator (bystander intervention) to estimate the direct effect of a bystander witness on survival to hospital discharge following cardiac arrest. Note that this approach utilizes a two-model method to account for the presence of treatment-mediator and mediator-outcome confounding

### *3.3. Applying Mediation Analysis*

To estimate the mediated effect in the presence of treatment-mediator and mediator-outcome confounding, we fit a mediation model using the mediation package in R.[11] Following the relationships proposed in the DAG, we assume a common set of confounders for the treatment-mediator relationship and the mediator-outcome relationship. The mediation model consists of two separate models: a mediator model, where we regress bystander intervention on the presence of a bystander witness, adjusting for confounders; and an outcome model, where we regress the outcome (survival) on the presence of a bystander witness, adjusting for the same confounders. We then estimate the causal mediation model and produce quasi-Bayesian confidence intervals for the estimated effects. We can estimate the natural direct effect (NDE) as the portion of the total effect attributable to bystander recognition. The NDE allows for estimation of the direct effect of the exposure on the outcome while allowing the mediator to take its natural course under different exposure conditions.[12] We estimate the NDE to understand the causal effect of exposure on outcome without intervening on the mediator. In other words, the NDE estimates the change in outcome that would occur as a result of changing the exposure value from 0 to 1 while holding the mediator at the level it would have otherwise taken had the exposure not been manipulated. In this case, we assume that the intervention of a biosensor does not impact the mediator (bystander intervention is not on the causal path between recognition and outcome) and thus the NDE is the appropriate estimate for this study. Using formula 1, we can estimate the portion of the total effect on survival that is attributable to bystander interventions. The estimates with their associated 95% quasi-Bayesian confidence intervals are presented in Table 1. The

partial effect attributable to recognition is the ADE (average), which we use as an approximation for the natural direct effect.

$$(1 - \frac{NDE}{TE}) \times 100, \tag{1}$$

| | Estimate | 95% CI Lower | 95% CI Upper |
|---|---|---|---|
| ACME (control) | 0.00784 | 0.00562 | 0.01 |
| ACME (treated) | 0.03529 | 0.02684 | 0.04 |
| ADE (control) | 0.10291 | 0.08942 | 0.12 |
| ADE (treated) | 0.13036 | 0.11548 | 0.15 |
| Total Effect | 0.13820 | 0.12307 | 0.15 |
| Prop. Mediated (control) | 0.05661 | 0.03959 | 0.08 |
| Prop. Mediated (treated) | 0.25605 | 0.19592 | 0.31 |
| ACME (average) | 0.02156 | 0.01627 | 0.03 |
| ADE (average) | 0.11664 | 0.10312 | 0.13 |
| Prop. Mediated (average) | 0.15633 | 0.11747 | 0.19 |

Table 1: The results of a confounder-adjusted mediation analysis for the relationship between a bystander witness. ACME: average causal mediated effect (bystander intervention); ADE: average direct effect (recognition).

The mediation analysis estimates can be contextualized in terms of the study exposure, outcome, and mediator as follows:

1. ACME (control): The average causal mediation effect for the control group, representing the average change in the outcome (survival) through the mediator (bystander intervention) for individuals not exposed to the exposure (bystander witness).
2. ACME (treated): The average causal mediation effect for the treated group, representing the average change in the outcome (survival) through the mediator (bystander intervention) for individuals exposed to the exposure (bystander witness).
3. ADE (control): The average direct effect for the control group, representing the average change in the outcome (survival) due to the exposure (bystander witness) independent of the mediator (bystander intervention) for individuals not exposed to the exposure.
4. ADE (treated): The average direct effect for the treated group, representing the average change in the outcome (survival) due to the exposure (bystander witness) independent of the mediator (bystander intervention) for individuals exposed to the exposure.
5. Total Effect: The combined effect of the direct effect of the exposure (bystander witness) and the effect mediated by the mediator (bystander intervention) on the outcome (survival).
6. Proportion Mediated (control): The proportion of the total effect that is mediated by the mediator (bystander intervention) in the control group.
7. Proportion Mediated (treated): The proportion of the total effect that is mediated by the mediator (bystander intervention) in the treated group.
8. ACME (average): The average of the ACMEs for the control and treated groups, representing the overall average change in the outcome (survival) through the mediator (bystander intervention) for individuals in the general population.
9. ADE (average): The average of the ADEs for the control and treated groups, representing the overall average change in the outcome (survival) due to the exposure (bystander witness) independent of the mediator (bystander intervention) for individuals in the general population.

10. Proportion Mediated (average): The proportion of the total effect that is mediated by the mediator (bystander intervention) for the general population, taking into account both control and treated groups.

The Total Effect (TE) is 0.1382. This implies that the total treatment effect of a witness on survival to discharge was 138 per 1000 patients. In other words, out of every 1000 patients experiencing OHCA, an additional 138 patients are expected to survive until hospital discharge if they are witnessed by a bystander. The Average Direct Effect (average) is estimated to be 0.11664. This means that the unmediated effect of a witness on survival to discharge was 117 per 1000 patients. We estimate the proportion of the total treatment effect that is owed to the ADE (average) as 84.3%. We estimate (1 - ADE(average)/TE) * 100% = 15.7%, indicating that 15.7% of the effect on survival to discharge is owed to the effect of a witness on bystander intervention.

## 4. Conclusion

We have demonstrated an application of DAG-informed mediation analysis with the goal of estimating the effect of a hypothetical technology intervention on a clinical outcome through intervening on a mediator in the causal path. Using registry data, we can estimate the attributable fraction owed to recognition in the causal relationship between a bystander witness and survival to hospital discharge following cardiac arrest, which allows us to make claims regarding the incremental benefit of scaling such an intervention in the population. We believe this approach is relevant and useful to other researchers seeking to model the potential benefit of technology interventions in a similar manner.

*Acknowledgement.* We appreciate the dedication and service of BCEHS paramedics, Fire-rescue personnel, and hospital-based medical staff for their commitment to providing exceptional patient care to OHCA patients. The BC Resuscitation Research Collaborative would like to acknowledge funding and/or in-kind support and support from the Heart and Stroke Foundation, Providence Health Care, and BC Emergency Health Services.

*Funding Statement.* This project was supported by MITACS Accelerate Grant IT26063 as well as a UBC grant for catalyzing research clusters.

*Competing Interests.* None

## References

[1] P. B. Burns, R. J. Rohrich, K. C. Chung, The levels of evidence and their role in evidence-based medicine, Plastic and reconstructive surgery 128 (1) (2011) 305–310. doi:10.1097/PRS.0b013e318219c171.

[2] L. Lin, M. Sperrin, D. A. Jenkins, G. P. Martin, N. Peek, A scoping review of causal methods enabling predictions under hypothetical interventions, Diagnostic and prognostic research 5 (1) (2021) 1–16.

[3] G. Moffa, G. Catone, J. Kuipers, E. Kuipers, D. Freeman, S. Marwaha, B. R. Lennox, M. R. Broome, P. Bebbington, Using directed acyclic graphs in epidemiological research in psychosis: An analysis of the role of bullying in psychosis, Schizophrenia bulletin 43 (6) (2017) 1273–1279. doi:10.1093/schbul/sbx013.

[4] K. F. Arnold, W. J. Harrison, A. J. Heppenstall, M. S. Gilthorpe, Dag-informed regression modelling, agent-based modelling and microsimulation modelling: a critical comparison of methods for causal inference, International journal of epidemiology 48 (1) (2019) 243–253.

[5] J. M. Albert, S. Nelson, Generalized causal mediation analysis, Biometrics 67 (3) (2011) 1028–1038. doi:10.1111/j.1541-0420.2010.01547.x.

[6] L. Richiardi, R. Bellocco, D. Zugna, Mediation analysis in epidemiology: methods, interpretation and bias, International Journal of Epidemiology 42 (5) (2013) 1511–1519. doi:10.1093/ije/dyt127.

[7] D. Tingley, T. Yamamoto, K. Hirose, L. Keele, K. Imai, mediation: R package for causal mediation analysis, Journal of Statistical Software 59 (5) (sep 2014).

[8] C. W. Tsao, A. W. Aday, Z. I. Almarzooq, A. Alonso, A. Z. Beaton, M. S. Bittencourt, A. K. Boehme, A. E. Buxton, A. P. Carson, Y. Commodore-Mensah, M. S. V. Elkind, K. R. Evenson, C. Eze-Nliam, J. F. Ferguson, G. Generoso, J. E. Ho, R. Kalani, S. S. Khan, B. M. Kissela, K. L. Knutson, S. S. Martin, Heart disease and stroke statistics-2022 update: A report from the american heart association, Circulation 145 (8) (2022) e153–e639. doi:10.1161/CIR.0000000000001052.

[9] J. Hutton, J. H. Puyat, M. Asamoah-Boaheng, B. Sobolev, S. Lingawi, M. Khalili, C. Kuo, B. Shadgan, J. Christenson, B. Grunau, The effect of recognition on survival after out-of-hospital cardiac arrest and implications for biosensor technologies, Resuscitation 190 (2023) 109906. doi:https://doi.org/10.1016/j.resuscitation.2023.109906.

[10] B. Sobolev, L. Kuramoto, Mediation analysis to inform policy on coronary revascularization by expected time to treatment: analytical framework, medRxiv (2022).
URL https://www.medrxiv.org/content/10.1101/2021.12.30.21268556v1

[11] D. Tofighi, D. P. MacKinnon, Rmediation: an r package for mediation analysis confidence intervals, Behavior research methods 43 (3) (2011) 692–700. doi:10.3758/s13428-011-0076-x.

[12] J. Pearl, Direct and indirect effects, in: Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence, Morgan Kaufmann, 2001, pp. 411–420.