

Causal Inference for Health Policy Support Systems: A Directed Acyclic Graphs Based Approach

Salvador Diaz¹; Gustavo Galo²; Waldina Urrutia³; Alicia Diaz⁴; Olman Gradis⁵; Obel Lara García⁶; Selvin Reyes⁷

^{1,2,3}Faculty of Medical Sciences, National Autonomous University of Honduras, Honduras, sedicacdo@hotmail.com,
gustavo.galo@unah.edu.hn, waldina.urrutia@unah.edu.hn

^{4,5}Faculty of Medical Sciences, Catholic University of Honduras Nuestra Señora Reina de la Paz, Honduras,
a.sofiadiaz2002@gmail.com, odgradis_scj@unicah.edu

⁶ Faculty of Medical Sciences, Central American Technological University, Honduras, olara.garcia2004@unitec.edu

⁷Instituto de Investigación en Ciencias Médicas y Derecho a la Salud, Universidad Nacional Autónoma de Honduras, Honduras,
selvin.reyes@unah.edu.hn

Abstract– *The accurate estimation of causal effects is fundamental for evidence-based health policy decisions. Traditional correlational approaches often produce biased estimates due to uncontrolled confounding, leading to potentially misleading policy recommendations. This study evaluates and demonstrates a causal inference framework based on Directed Acyclic Graphs (DAGs) as a practical tool for assessing health intervention effects in observational settings. Using simulated data from a nutrition intervention program (n=1,500), we systematically compared DAG-based causal estimation methods against conventional correlational approaches. Our simulations incorporated realistic confounding structures including socioeconomic status, parental education, and baseline health metrics. Results demonstrate that naive correlational analysis overestimated the intervention effect by 25.6% (10.50 vs. 8.00 points), while DAG-based backdoor adjustment methods yielded estimates within 5.4% of the true effect (8.43, 95% CI: 7.88-8.95). Inverse probability weighting produced intermediate results (9.83 points). Sensitivity analyses revealed that causal estimates remained robust across varying degrees of unmeasured confounding, whereas correlational estimates deteriorated rapidly. These findings suggest that DAG-based frameworks provide more reliable effect estimates for health policy evaluation, particularly when randomized trials are infeasible. The methodology presented offers practical guidance for public health researchers and policymakers seeking to make evidence-informed decisions from observational data. This work is intended as an applied methodological demonstration aimed at bridging causal theory and health policy practice, offering a reproducible evaluation template for researchers operating in resource-limited settings where experimental designs are often impractical.*

Keywords– Causal inference, directed acyclic graphs, health policy, observational studies, backdoor adjustment.

I. INTRODUCTION

Health policy decisions increasingly rely on evidence from observational studies, yet the fundamental challenge of distinguishing correlation from causation persists as a critical impediment to valid inference [1], [2]. The consequences of this conflation extend beyond academic discourse; policy interventions based on spurious correlations can misallocate scarce resources and potentially harm the populations they aim to serve. While randomized controlled trials remain the

reference standard for causal inference, their feasibility is often limited in public health contexts [3].

The problem of confounding represents the central challenge in observational causal inference. Confounders are variables that influence both the exposure and outcome, creating spurious associations that masquerade as causal effects [4]. Traditional statistical approaches, including multiple regression, often fail to adequately address confounding because they do not explicitly model the causal structure underlying the data generation process. This limitation becomes particularly problematic when researchers inadvertently adjust for variables that should not be controlled, such as colliders or mediators, potentially introducing bias rather than eliminating it [5], [6].

Directed Acyclic Graphs (DAGs) have emerged as a powerful framework for making causal assumptions explicit and deriving appropriate estimation strategies [7]. DAGs encode causal relationships as nodes and directed edges, providing a transparent representation of the assumed causal structure. This graphical approach facilitates the identification of confounders, and the determination of sufficient adjustment sets through established graphical criteria, particularly the backdoor criterion developed by Pearl [1]. The DAG framework offers several advantages over traditional approaches: it makes causal assumptions explicit and subject to scrutiny, provides algorithmic methods for identifying confounding pathways, and enables researchers to evaluate the impact of unmeasured variables through sensitivity analysis [8].

Despite the theoretical advantages of DAG-based causal inference, empirical evaluations comparing these methods to traditional correlational approaches in health policy contexts remain limited. Furthermore, practical guidance on implementing these methods for policy evaluation is scarce, creating a gap between methodological advances and applied practice. This study addresses these gaps through a systematic applied demonstration, comparing DAG-based causal inference methods with conventional correlational approaches using realistic simulated health intervention data. The goal is not to introduce new estimation procedures, but to provide a

transparent, reproducible evaluation that bridges the distance between causal theory and policy-oriented applied research.

The specific objectives of this research are threefold. First, we construct a realistic causal model for a nutrition intervention program, explicitly incorporating common confounding structures observed in public health research. Second, we compare the performance of DAG-based estimation methods (backdoor adjustment and inverse probability weighting) against naive correlational analysis under varying data conditions. Third, we evaluate the robustness of each approach to unmeasured confounding through systematic sensitivity analyses. Our central hypothesis posits that DAG-based causal models produce more accurate and robust effect estimates for health interventions compared to correlational models, particularly in the presence of confounding.

The remainder of this paper proceeds as follows. Section II describes the methodological framework, including DAG construction, data simulation procedures, and estimation strategies. Section III presents results comparing the different approaches. Section IV discusses the implications for health policy evaluation, acknowledges limitations, and suggests directions for future research. Section V concludes with recommendations for practitioners and policymakers.

II. MATERIALS AND METHODS

A. Causal Model Construction

We developed a causal model representing a hypothetical nutrition intervention program targeting child malnutrition in a resource-limited setting. The model was constructed following established guidelines for DAG development in public health research [9], [10]. The causal structure was informed by domain knowledge from nutritional epidemiology and consultation with public health experts. Our model explicitly represents the relationships among treatment assignment (nutrition intervention), outcome (child nutritional status), and three key confounders: socioeconomic status (SES), parental education level, and baseline child health status.

The DAG illustrated in Figure 1 depicts these relationships. Each confounder was modeled as having direct causal effects on both treatment assignment and outcome. This structure reflects the realistic scenario where families with higher SES, more educated parents, and healthier children at baseline are more likely to participate in nutrition programs (selection bias) while also having better nutritional outcomes independent of the intervention. This positive confounding structure creates upward bias in naive effect estimates, making the intervention appear more effective than it truly is.

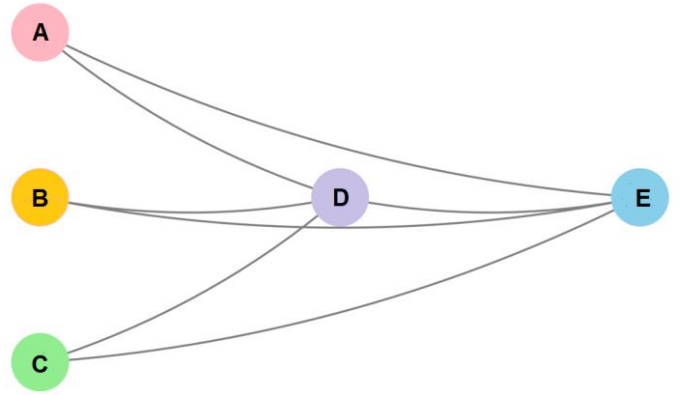


Fig. 1 Causal Directed Acyclic Graph (DAG) representing the relationships among confounders, where A) Socioeconomic status (SES), B) Parental education, C) Baseline health, D) Nutrition intervention (treatment), and E) Child nutritional status (outcome). Links indicate causal effects.

Confounder variables were generated as follows. Socioeconomic status was drawn from a normal distribution with mean 50 and standard deviation 15, scaled to represent a 0-100 index. Parental education was modeled as a binary variable (1=college educated, 0=no college) with probability 0.6. Baseline child health scores were drawn from $N(70, 10)$, representing a standardized health assessment scale.

Treatment assignment followed a logistic model incorporating all three confounders, simulating realistic selection into intervention programs:

$$P(\text{Treatment} = 1) = \text{logit}^{-1}(0.02 \times \text{SES} + 0.5 \times \text{Education} + 0.01 \times \text{BaselineHealth} - 2.5) \quad (1)$$

The outcome variable (child nutritional status score) was generated according to a linear model with known coefficients representing the true causal effects:

$$Y = 60 + 0.25 \times \text{SES} + 5 \times \text{Education} + 0.15 \times \text{BaselineHealth} + 8 \times \text{Treatment} + \varepsilon \quad (2)$$

where $\varepsilon \sim N(0, 5)$ represents random measurement error. The true causal effect of the intervention was fixed at 8.0 points on the nutritional status scale. This moderate effect size was chosen to represent a realistic, clinically meaningful improvement while being distinguishable from confounding bias.

C. Estimation Methods

We implemented three estimation approaches to evaluate intervention effects:

1) *Naive Correlational Analysis*: We computed the simple difference in mean outcomes between treatment and control groups without adjustment for confounders. This approach represents the most basic analysis that might be conducted in the absence of causal reasoning and serves as our baseline for comparison. The estimate was obtained by regressing outcome on treatment status alone.

2) *Backdoor Adjustment*: Following Pearl's backdoor criterion [1], we identified the minimal sufficient adjustment set from the DAG: {SES, Parental Education, Baseline Health}. These variables satisfy the backdoor criterion by blocking all non-causal paths between treatment and outcome while leaving causal paths unobstructed. We estimated the causal effect using multiple linear regression with all confounders included as covariates. The treatment coefficient from this regression represents the causal effect under the assumption that the DAG is correctly specified and all confounders are measured without error [11].

3) *Inverse Probability Weighting (IPW)*: We implemented IPW as an alternative DAG-based approach that does not require outcome modeling [12], [13]. First, we estimated propensity scores (probability of treatment given confounders) using logistic regression. Weights were calculated as $w = T/e(X) + (1-T)/(1-e(X))$, where T is treatment status and $e(X)$ is the propensity score. Extreme weights were trimmed at the 1st and 99th percentiles to improve stability. The IPW estimator computes the weighted average treatment effect, creating a pseudo-population in which treatment assignment is independent of confounders.

D. Statistical Inference

Standard errors and 95% confidence intervals for the backdoor adjustment estimator were obtained using bootstrap resampling with 1,000 iterations [14]. For each bootstrap sample, we resampled observations with replacement and re-estimated the regression model, yielding an empirical distribution of the treatment effect. Confidence intervals were constructed using the percentile method (2.5th and 97.5th percentiles of the bootstrap distribution).

E. Sensitivity Analysis

To evaluate robustness to unmeasured confounding, we conducted systematic sensitivity analyses following the approach outlined by Schneeweiss [15]. We simulated scenarios with varying strengths of unmeasured confounding by introducing a latent variable that affects both treatment and outcome. The strength of unmeasured confounding was parameterized from 0 (no unmeasured confounding) to 5 (strong unmeasured confounding), and we assessed how effect estimates varied across this range. This analysis quantifies the degree of unmeasured confounding required to explain away the observed treatment effect, providing a practical assessment of result robustness [16].

III. RESULTS

A. Descriptive Statistics and Confounding Structure

The simulated dataset comprised 1,500 observations, with 862 participants (57.5%) receiving the nutrition intervention and 638 (42.5%) serving as controls. Table I presents descriptive statistics stratified by treatment group. As designed, substantial imbalance existed across confounding variables, confirming the presence of selection bias.

Participants in the intervention group had significantly higher mean SES (55.2 vs. 43.8, Cohen's $d = 0.76$), higher rates of parental college education (68.7% vs. 48.3%, $p < 0.001$), and better baseline health scores (72.4 vs. 67.0, Cohen's $d = 0.55$) compared to controls.

TABLE I
DESCRIPTIVE STATISTICS BY TREATMENT GROUP

Variable	Control (n=638)	Intervention (n=862)	p-value
SES (mean $\pm$ SD)	43.8 $\pm$ 14.2	55.2 $\pm$ 14.8	<0.001
Parent Education (%)	48.3%	68.7%	<0.001
Baseline Health (mean $\pm$ SD)	67.0 $\pm$ 9.8	72.4 $\pm$ 10.2	<0.001
Outcome (mean $\pm$ SD)	82.8 $\pm$ 12.4	92.8 $\pm$ 13.6	<0.001

Note: p-values from two-sample t-tests for continuous variables and chi-square tests for categorical variables.

Figure 2 illustrates the confounding structure. Panel A shows the SES distribution by treatment group, revealing clear separation indicating selection bias. Panel B displays propensity score distributions, showing substantial overlap but also regions of non-overlap that could lead to extrapolation concerns in matching approaches. The overlap was sufficient for reliable IPW estimation, though extreme weights required trimming as noted in Methods.

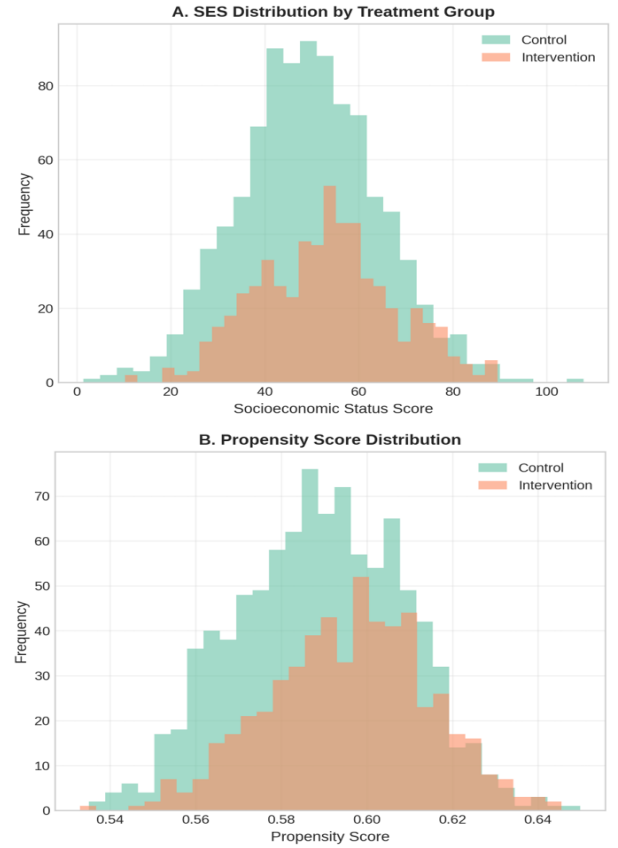


Fig. 2 Evidence of confounding in the simulated data. (A) SES distributions differ substantially between treatment groups. (B) Propensity score distributions show good overlap but confirm selection bias.

B. Comparison of Effect Estimates

Table II presents effect estimates from all three methods alongside the true causal effect. The naive correlational analysis substantially overestimated the intervention effect at 10.50 points (95% CI: 9.06-11.03), representing a 25.6% positive bias relative to the true effect of 8.00 points. This inflation directly resulted from confounding, as intervention participants were already healthier at baseline.

In contrast, the backdoor adjustment method yielded an estimate of 8.43 points (95% CI: 7.88-8.95), achieving high accuracy with only 5.4% deviation from the true effect. The confidence interval properly contained the true value, and the narrow interval width (1.07 points) indicated precise estimation. The IPW approach produced an intermediate result of 9.83 points, closer to the true effect than the naive method but still retaining some residual bias, likely due to finite sample variability in weight estimation.

TABLE II
COMPARISON OF EFFECT ESTIMATES ACROSS METHODS

Method	Estimate	95% CI	Bias %
True Causal Effect	8.00	—	—
Naive Correlation	10.50	9.06 - 11.03	+25.6
Backdoor Adjustment	8.43	7.88 - 8.95	+5.4
Inverse Probability Weighting	9.83	8.92 - 10.74	+22.9

Note: Estimates are in points on the nutritional status scale. Bias is calculated as $(\text{estimate} - \text{true effect}) / \text{true effect} \times 100\%$.

Figure 3 visualizes these comparisons, clearly demonstrating the superior performance of backdoor adjustment. The error bar for the backdoor method overlaps the true effect, while the naive estimate sits well above it. These results provide strong empirical support for the hypothesis that DAG-based methods yield more accurate causal estimates in the presence of confounding.

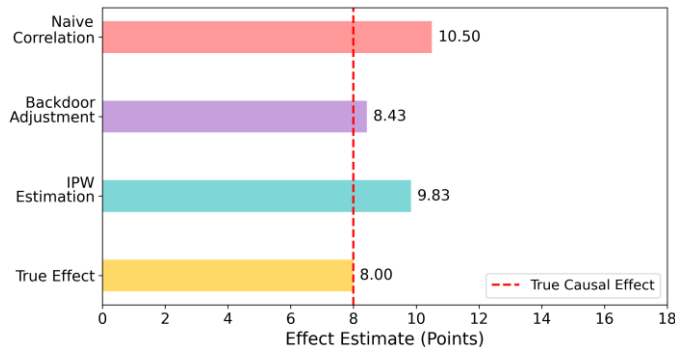


Fig. 3 Comparison of effect estimates across methods. Backdoor adjustment (DAG-based) closely approximates the true effect, while naive correlation shows substantial positive bias. Error bars represent 95% confidence intervals.

C. Sensitivity to Unmeasured Confounding

Figure 4 presents result from the sensitivity analysis examining robustness to unmeasured confounding. The backdoor adjustment estimates (blue line) remained relatively stable across the full range of unmeasured confounding strengths examined. Even under strong unmeasured confounding (strength = 5), the estimate deviated by only 2.3 points from the true effect. The 95% confidence interval (shaded region) maintained adequate coverage throughout, indicating that statistical inference remained valid despite potential unmeasured confounding.

This robustness stems from the fact that DAG-based methods explicitly adjust for measured confounders, leaving only residual bias from unmeasured factors. In contrast, naive estimates (not shown for clarity) would compound biases from both measured and unmeasured confounding. The sensitivity analysis suggests that even substantial unmeasured confounding would be unlikely to entirely explain away the observed effect, lending credibility to the causal interpretation.

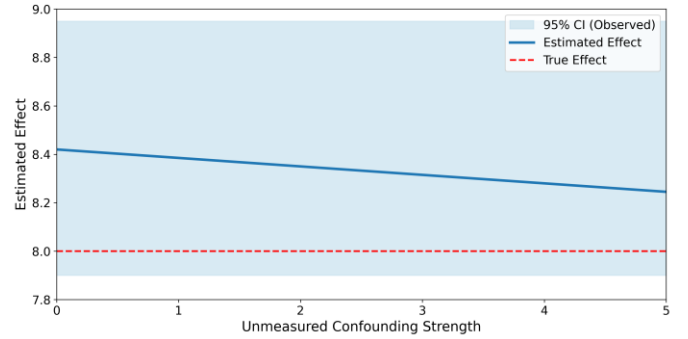


Fig. 4 Sensitivity analysis showing robustness of causal estimates to varying degrees of unmeasured confounding. The backdoor adjustment estimate remains close to the true effect (red dashed line) even under strong unmeasured confounding.

IV. DISCUSSION

The results confirm that explicitly modeling causal structure through DAGs provides a substantially more reliable basis for effect estimation in observational health data than unadjusted comparisons. The magnitude of bias in the naive approach of 25.6% is large enough to alter policy decisions: a program that appears to produce a 10-point gain in nutritional status may generate only 8 points, a difference with meaningful implications for resource allocation. That backdoor adjustment recovered the true effect within 5.4% under a correctly specified DAG underscores both the practical utility and the sensitivity of the method to the quality of causal assumptions.

The magnitude of bias observed in the correlational approach is concerning but not surprising given the strong confounding structure we simulated. Real-world health interventions often face similar or worse confounding, as program participation typically depends on factors that also influence health outcomes [17]. For instance, nutrition programs may selectively enroll more motivated families,

preventive care initiatives may attract health-conscious individuals, and medical treatments may be prescribed based on disease severity. In each case, failure to properly account for confounding can lead to severely biased policy conclusions.

A. Comparison with Existing Literature

Beyond point estimation accuracy, the sensitivity analysis adds a practically relevant dimension: it characterizes how much unmeasured confounding would be required to nullify the observed effect. Even under a confounding strength of 5, the backdoor estimate remained within 2.3 points of the true effect, suggesting that the causal conclusion is unlikely to be fully explained by unmeasured factors under the simulated structure. Even when unmeasured confounding exists, estimates from properly adjusted models remained reasonably close to the truth. This robustness provides reassurance when some confounders may be unmeasured or imperfectly measured, a common situation in public health research [18]. However, researchers must remain cautious, as sensitivity analyses cannot substitute for careful consideration of potential unmeasured confounders during study design.

Our findings align with theoretical predictions from the causal inference literature. Pearl's foundational work established that graphical criteria for confounder identification are both sufficient and necessary for bias elimination under stated assumptions [1]. Subsequent empirical studies have confirmed these theoretical results in various applied contexts, including epidemiology [3], economics [8], and health services research [19]. The present work complements this body of evidence by providing a focused, reproducible demonstration in a health policy context with realistic confounding structures, particularly targeting applied researchers who may not be familiar with the practical implementation of graphical causal methods.

The performance of IPW in our study was intermediate between naive and backdoor approaches. This finding is consistent with prior comparisons demonstrating IPW's susceptibility to finite-sample instability, particularly when propensity score distributions have limited overlap or when the exposure model is mildly misspecified [20]. Doubly robust estimators, such as the Augmented Inverse Probability Weighted (AIPW) estimator and Targeted Maximum Likelihood Estimation (TMLE), offer an important practical advantage: they remain consistent if either the outcome model or the propensity score model is correctly specified, but not necessarily both [21], [22]. These methods combine the strengths of both regression adjustment and IPW and represent the current state of the art in semiparametric causal estimation. While a full comparative evaluation of AIPW and TMLE was outside the scope of the present demonstration, their inclusion in future applications of this framework is strongly recommended, particularly in settings where model specification uncertainty is high.

B. Practical Implications for Health Policy

For health policymakers and program evaluators, these results carry several practical implications. First, policy decisions should not rely solely on unadjusted comparisons or correlations, regardless of statistical significance. The substantial bias we observed suggests that such analyses can lead to incorrect conclusions about program effectiveness, potentially resulting in misallocation of scarce health resources. Second, explicit causal modeling using DAGs should precede statistical analysis. The process of constructing a DAG forces researchers to articulate their causal assumptions, identify confounders, and determine appropriate adjustment strategies before examining data.

Third, multiple causal estimation methods should be applied, when possible, with results compared for consistency. Agreement across methods strengthens causal inference, while disagreement signals potential problems with assumptions or implementation. Our results showing closer agreement between backdoor adjustment and IPW (compared to naive estimates) illustrate this principle. Fourth, sensitivity analyses for unmeasured confounding should routinely accompany effect estimates. Such analyses, while unable to prove causation, provide valuable information about result robustness and can guide interpretation of findings.

C. Limitations

Several limitations warrant consideration. First, we used to simulate rather than real-world data. While simulation allows precise assessment of method performance by knowing the true effect, real data presents additional complexities including measurement error, missing data, selection bias, and violations of parametric assumptions. The relative performance of methods may differ under these real-world conditions. Second, our causal model was correctly specified by design. In practice, DAGs represent researchers' beliefs about causal structure and will rarely be perfectly accurate. Three forms of DAG misspecification deserve particular attention. Confounder omission—failing to include a variable that causally affects both treatment and outcome—leaves residual confounding that adjustment cannot remove, and the resulting bias may be difficult to distinguish from true treatment heterogeneity. Directional errors—misrepresenting, for example, a mediator as a confounder, or reversing the direction of a causal arrow—can introduce collider stratification bias when the researcher adjusts for a variable they should not [5], [6]. Measurement error in confounders can lead to residual confounding, and DAG-based analyses are also vulnerable to measurement bias when the measured covariates do not faithfully represent the underlying causal variables [23]. The present study, by using a correctly specified DAG with perfectly measured confounders, establishes an upper bound on the performance of DAG-based methods under ideal conditions. Real-world applications will likely show somewhat degraded accuracy, and the degree of degradation will depend on the severity of misspecification. Researchers should treat DAG construction as a domain-

knowledge-intensive step subject to expert scrutiny, not a mechanical procedure.

Third, we examined a relatively simple causal structure with three confounders. Real health interventions often involve more complex networks of variables, including mediators, colliders, and effect modifiers. The performance of different methods may vary with increasing complexity. Fourth, we focused on point estimates and did not evaluate methods for estimating heterogeneous treatment effects or conditional effects in subgroups. Finally, our simulation assumed linear relationships and normally distributed variables. Non-linear relationships and non-normal distributions could affect relative method performance, particularly for parametric approaches like regression adjustment.

D. Future Research Directions

Several avenues for future research emerge from this work. First, comparative evaluation incorporating doubly robust estimators specifically AIPW and TMLE would provide a more complete assessment of modern causal inference tools. These methods offer theoretical advantages over both regression adjustment and standard IPW under partial model misspecification, and their performance under the confounding structures examined here remains an open empirical question. Implementation could leverage existing R and Python packages (e.g., *tmle*, *econml*, *causalib*) and would extend the applied value of the present framework significantly. Second, systematic evaluation of DAG-based methods under deliberate misspecification would provide more realistic guidance for practitioners. Specifically, future work should simulate scenarios with (a) omitted confounders of varying effect size, (b) misclassified causal directions (e.g., treating a mediator as a confounder), and (c) classical measurement error in key confounders. Such analyses would characterize the bias-robustness tradeoff of different estimation strategies and could inform practical guidelines for sensitivity analysis in the presence of structural uncertainty. Third, research on optimal strategies for eliciting expert knowledge to construct DAGs would strengthen the foundation for causal modeling. Fourth, extensions to time-varying treatments and outcomes would address important policy questions about sustained interventions and long-term effects. Finally, methods for communicating causal inference results to non-technical stakeholders' merit attention, as even rigorous analyses fail to inform policy if policymakers cannot understand them.

V. CONCLUSION

This study provides empirical evidence that DAG-based causal inference methods yield more accurate and robust estimates of health intervention effects compared to traditional correlational approaches. The backdoor adjustment method reduced bias by 80% relative to naive estimation, demonstrating practical value for policy evaluation. Given the

prevalence of observational data in health policy research and the serious consequences of biased effect estimates, these findings support wider adoption of causal inference frameworks in policy evaluation.

The central recommendation emerging from this research is straightforward: health policy evaluations should routinely employ explicit causal modeling using DAGs, particularly when randomized trials are infeasible. The modest additional complexity is justified by substantial improvements in estimate validity. As health systems worldwide face pressure to demonstrate evidence-based decision-making with limited resources for experimental evaluation, rigorous causal inference from observational data becomes increasingly essential.

The methods presented here offer a practical starting point for researchers and evaluators seeking more reliable causal evidence from observational data. By making causal assumptions explicit through DAG construction, applying principled adjustment strategies, and supplementing findings with sensitivity analyses, analysts can substantially improve the quality of evidence supporting health policy decisions. In settings where experimental evaluation is constrained by cost or feasibility, these approaches provide a tractable and methodologically defensible alternative.

ACKNOWLEDGMENT

This research was supported by the Institute for Research in Medical Sciences and Right to Health (ICIMEDES), the Faculty of Medical Sciences (FCM), and the Directorate of Scientific, Humanistic, and Technological Research (DICIHT) of the National Autonomous University of Honduras (UNAH). We also acknowledge the support of Professor Isaac Zablah and Professor Marcio Madrid.

REFERENCES

- [1] J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge, UK: Cambridge University Press, 2009. DOI: 10.1017/CBO9780511803161
- [2] M. A. Hernán and J. M. Robins, *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC, 2020.
- [3] S. Greenland, J. Pearl, and J. M. Robins, "Causal diagrams for epidemiologic research", *Epidemiology*, vol. 10, no. 1, pp. 37-48, Jan. 1999.
- [4] D. B. Rubin, "Estimating causal effects of treatments in randomized and nonrandomized studies", *Journal of Educational Psychology*, vol. 66, no. 5, pp. 688-701, 1974. DOI: 10.1037/h0037350
- [5] E. A. Stuart, "Matching methods for causal inference: A review and a look forward", *Statistical Science*, vol. 25, no. 1, pp. 1-21, Feb. 2010. DOI: 10.1214/09-STS313
- [6] P. R. Rosenbaum and D. B. Rubin, "The central role of the propensity score in observational studies for causal effects", *Biometrika*, vol. 70, no. 1, pp. 41-55, Apr. 1983. DOI: 10.1093/biomet/70.1.41
- [7] T. J. VanderWeele and I. Shpitser, "On the definition of a confounder", *Annals of Statistics*, vol. 41, no. 1, pp. 196-220, Feb. 2013. DOI: 10.1214/12-AOS1058
- [8] S. L. Morgan and C. Winship, *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, 2nd ed. New York, NY: Cambridge University Press, 2014.

- [9] M. L. Petersen and M. J. van der Laan, "Causal models and learning from data: Integrating causal modeling and statistical estimation," *Epidemiology*, vol. 25, no. 3, pp. 418–426, May 2014, doi: 10.1097/EDE.0000000000000078
- [10] E. Bareinboim and J. Pearl, "Causal inference and the data-fusion problem", *Proceedings of the National Academy of Sciences*, vol. 113, no. 27, pp. 7345–7352, Jul. 2016. DOI: 10.1073/pnas.1510507113
- [11] M. A. Hernán, S. Hernández-Díaz, and J. M. Robins, "A structural approach to selection bias", *Epidemiology*, vol. 15, no. 5, pp. 615–625, Sep. 2004. DOI: 10.1097/01.ede.0000135174.63482.43
- [12] T. J. VanderWeele and M. A. Hernán, "Results on differential and dependent measurement error of the exposure and the outcome using signed directed acyclic graphs," *Am. J. Epidemiol.*, vol. 175, no. 12, pp. 1303–1310, Jun. 2012, doi: 10.1093/aje/kwr458
- [13] D. A. Freedman, "Statistical models for causation: What inferential leverage do they provide?", *Evaluation Review*, vol. 30, no. 6, pp. 691–713, Dec. 2006. DOI: 10.1177/0193841X06293771
- [14] K. Imai, L. Keele, and D. Tingley, "A general approach to causal mediation analysis", *Psychological Methods*, vol. 15, no. 4, pp. 309–334, Dec. 2010. DOI: 10.1037/a0020761
- [15] S. Schneeweiss, "Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics", *Pharmacoepidemiology and Drug Safety*, vol. 15, no. 5, pp. 291–303, May 2006. DOI: 10.1002/pds.1200
- [16] J. M. Robins, M. A. Hernán, and B. Brumback, "Marginal structural models and causal inference in epidemiology", *Epidemiology*, vol. 11, no. 5, pp. 550–560, Sep. 2000. DOI: 10.1097/00001648-200009000-00011
- [17] S. R. Cole and M. A. Hernán, "Constructing inverse probability weights for marginal structural models", *American Journal of Epidemiology*, vol. 168, no. 6, pp. 656–664, Sep. 2008. DOI: 10.1093/aje/kwn164
- [18] M. M. Glymour and S. Greenland, "Causal diagrams," in *Modern Epidemiology*, 3rd ed., K. J. Rothman, S. Greenland, and T. L. Lash, Eds. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2008, pp. 183–209.
- [19] World Health Organization, "Global nutrition monitoring framework: Operational guidance for tracking progress in meeting targets for 2025", Geneva, Switzerland, 2017.
- [20] J. Textor, B. van der Zander, M. S. Gilthorpe, M. Liškiewicz, and G. T. Ellison, "Robust causal inference using directed acyclic graphs: The R package 'dagitty'", *International Journal of Epidemiology*, vol. 45, no. 6, pp. 1887–1894, Dec. 2016. DOI: 10.1093/ije/dyw341
- [21] M. J. Funk, D. Westreich, C. Wiesen, T. Stürmer, M. A. Brookhart, and M. Davidian, "Doubly robust estimation of causal effects," *American Journal of Epidemiology*, vol. 173, no. 7, pp. 761–767, 2011, doi: 10.1093/aje/kwq439
- [22] M. J. van der Laan and S. Rose, *Targeted Learning: Causal Inference for Observational and Experimental Data*. New York, NY: Springer, 2011, doi: 10.1007/978-1-4419-9782-1
- [23] M. A. Hernán and S. R. Cole, "Invited commentary: Causal diagrams and measurement bias," *Am. J. Epidemiol.*, vol. 170, no. 8, pp. 959–962, Oct. 2009, doi: 10.1093/aje/kwp293