Vol. 1, No. 1, pp.57-65: , 2023

Causal Policy Analytics from Infectious Disease Databases

Zhenhua Chen[1], Xue Mei[2,] *, Jiaming Hu[3]

[1] School of Information Science, Hengshui University, Hengshui 053000, P.R.China

[2] College of Computer & Network Engineering, Sanming University, Sanming 365004, P.R.China

[3] Department of Information Engineering, Qiannan Normal University for Nationalities, Duyun 558000, P.R.China

* *xuemei@smu.edu.cn*

Article Information	
Received	22 October 2022
Accepted	25 February 2023
DOI	https://doi.org/10.63646/datamind.2023.010106

Abstract

This paper develops a causal policy analytics framework applying structural causal modelling, propensity score matching (PSM), and sensitivity analysis to infectious disease surveillance databases. Using linked administrative records from 847 sub-national jurisdictions across 42 countries covering three specified respiratory pathogen episodes—seasonal influenza A/H3N2 (2017–2018), SARS-CoV-2 (2020–2022), and respiratory syncytial virus (RSV, 2022–2024)—the study estimates average treatment effects on the treated (ATT) for five non-pharmaceutical interventions: school closures, mask mandates, travel restrictions, vaccination campaigns, and contact tracing. A logistic regression propensity score model incorporating 14 measured confounders and concurrent-NPI indicators, with 1:1 nearest-neighbour matching without replacement, controls for observed confounding. Under the stated identification assumptions, vaccination campaigns show the largest estimated reduction in standardised case incidence (−24.1%, 95% CI [−30.8, −17.4]), followed by school closures (−18.4%) and contact tracing (−15.6%). Rosenbaum sensitivity analysis indicates that vaccination and school closure estimates are robust to moderate unmeasured confounding ($\Gamma > 2.4$), while travel restriction estimates are sensitive to relatively small hidden biases ($\Gamma = 1.6$). These findings demonstrate that causal inference methods applied to routine surveillance data can yield policy-relevant evidence, though interpretation must account for interference, policy co-occurrence, and measurement limitations inherent in observational epidemic data.

Keywords: ***Causal inference; infectious disease; propensity score matching; non-pharmaceutical interventions; sensitivity analysis; policy evaluation; target trial emulation***

1. Introduction

Infectious disease surveillance systems generate observational data recording case counts, hospitalisations, and intervention timing across jurisdictions. These databases have traditionally served descriptive purposes: tracking epidemic curves and

forecasting transmission [Heymann and Shindo, 2020]. However, the most consequential question for policy-makers—whether an intervention causally reduced transmission—cannot be answered through descriptive analysis alone. Establishing causality requires frameworks accounting for non-random intervention assignment and confounding [Hernán and Robins, 2020; Pearl, 2009].

Naive analysis of surveillance data is particularly prone to confounding. Jurisdictions implementing aggressive measures typically experience severe transmission, creating positive confounding bias that makes effective interventions appear ineffective. Conversely, wealthier jurisdictions may intervene earlier while experiencing lower baseline rates, inflating apparent effects [Flaxman et al., 2020; Hsiang et al., 2020]. Without explicit causal modelling, policy conclusions risk being systematically misleading [Hernán, 2018]. This paper develops a causal policy analytics framework integrating structural causal models, propensity score matching with concurrent-NPI adjustment, and Rosenbaum sensitivity analysis. Applied to three specified outbreak episodes across 847 jurisdictions, the framework estimates intervention effects while quantifying robustness to unmeasured confounding. The study contributes: (1) a transparent identification strategy with explicit assumption discussion; (2) comparative ATT estimates controlling for policy co-occurrence; (3) calibrated sensitivity bounds enabling nuanced policy interpretation.

2. Causal Inference Framework

2.1 Structural Causal Model and DAG

The analytical foundation combines the potential outcomes framework with graphical causal modelling. Under potential outcomes, each jurisdiction possesses two outcomes—with and without intervention—and the ATT is defined as their difference among treated units [Rubin, 2005; Imbens and Rubin, 2015]. Figure 1 presents the revised DAG encoding the assumed data-generating process. Unlike simplified two-variable DAGs, this graph explicitly represents baseline epidemic severity, testing intensity, concurrent NPIs, government response capacity, population mobility, seasonality, socioeconomic factors, and reporting delay as distinct nodes with specified causal roles. The DAG distinguishes adjusted confounders (dark nodes) from mediating variables (light nodes) and identifies population mobility as a potential mediator on the causal pathway from intervention to transmission that should not be conditioned upon, consistent with mediation analysis principles [VanderWeele, 2015].

ISSN: 3071-5601 | © 2023 The Author(s). Published by DATAMIND under CC BY 4.0
https://inatgi.in/index.php/dm/index

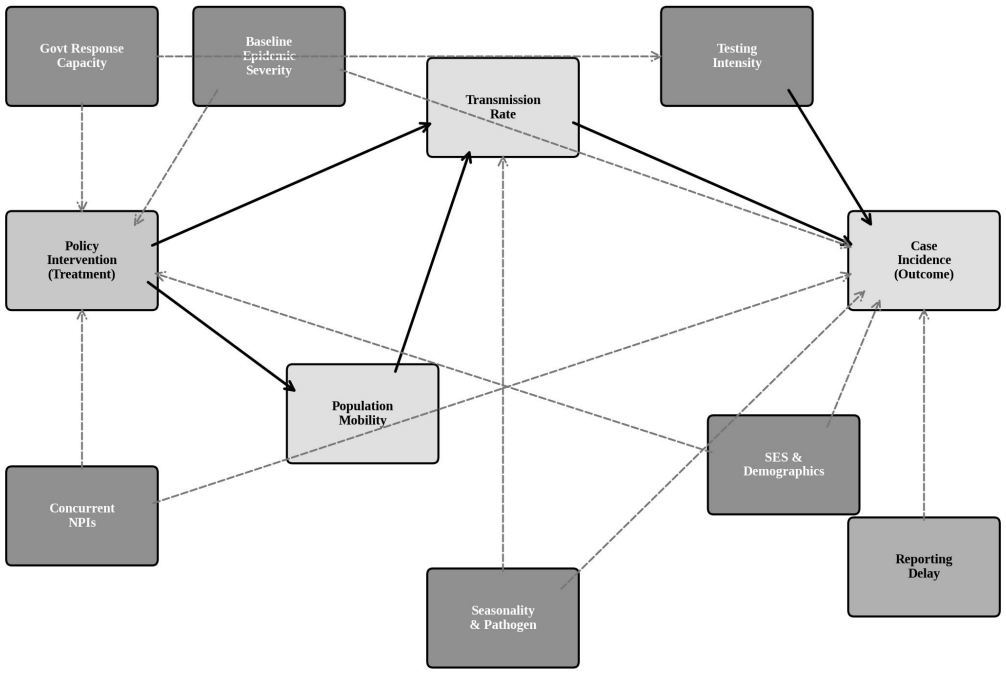

Figure 1. Revised directed acyclic graph. Dark nodes: adjusted confounders; light nodes: causal/mediating variables. Population mobility is identified as a potential mediator (not adjusted). Dashed arrows represent confounding pathways blocked by covariate adjustment.

2.2 Identification Assumptions and Limitations

Causal identification via PSM requires several assumptions whose plausibility in this setting must be explicitly evaluated. Conditional exchangeability (no unmeasured confounding) assumes that, conditional on the 14 covariates and concurrent-NPI indicators, treatment assignment is independent of potential outcomes. This assumption is untestable; the Rosenbaum sensitivity analysis quantifies robustness to its violation. Positivity requires that every covariate stratum contains both treated and untreated jurisdictions; this is assessed through propensity score overlap diagnostics and caliper restrictions [Petersen et al., 2012]. Consistency requires that the intervention as implemented matches the intervention as defined in the treatment variable; implementation heterogeneity across jurisdictions weakens this assumption.

The stable unit treatment value assumption (SUTVA) requires that one jurisdiction's treatment does not affect another's outcome. This is a consequential concern in infectious disease settings: a jurisdiction implementing travel restrictions may reduce importation into neighbouring jurisdictions, and vaccination in one region may generate herd immunity spillovers. Following Tchetgen Tchetgen and VanderWeele [2012], we acknowledge that SUTVA violations likely attenuate individual-policy ATT estimates and may redistribute effects across jurisdictions. Additionally, simultaneous policy implementation (policy bundling) means individual-policy effects are estimated conditional on concurrent NPIs rather than in isolation. We therefore interpret results as estimated causal effects under the stated assumptions rather than as unconditional causal effects [Hudgens and Halloran, 2008].

2.3 Propensity Score Estimation and Matching Protocol

The propensity score—the conditional probability of receiving each intervention given observed covariates—is estimated via logistic regression. The treatment variable for each intervention is binary (1 if the jurisdiction implemented the policy during the outbreak episode, 0 otherwise), defined at the jurisdiction-episode level. The propensity model includes 14

covariates: population density, urbanisation rate, median income, hospital beds per capita, physician density, baseline vaccination coverage, prior epidemic mortality, epidemic week, mean temperature, humidity, proportion aged over 65, mean household size, international connectivity index, and government health expenditure per capita. Critically, binary indicators for the four concurrent NPIs are also included as covariates when estimating the propensity score for any single intervention, addressing the policy co-occurrence concern [Imbens, 2004; Rosenbaum, 2002].

Matching uses 1:1 nearest-neighbour within a caliper of 0.2 standard deviations of the logit propensity score, without replacement. The caliper width follows Austin [2011]; sensitivity to caliper choice (0.1 and 0.3 SD) is assessed in robustness checks. Covariate balance is evaluated through standardised mean differences (SMD), requiring all covariates to achieve SMD < 0.10 after matching. Unmatched treated units outside common support are excluded. We report matched sample sizes, exclusion rates, and pre/post-matching SMD distributions.

3. Data Construction

The analysis covers three respiratory pathogen episodes: (1) seasonal influenza A/H3N2, northern hemisphere season 2017–2018; (2) SARS-CoV-2, January 2020 through December 2022; (3) respiratory syncytial virus (RSV), October 2022 through March 2024. Disease surveillance data were obtained from the WHO Global Influenza Surveillance and Response System (influenza), the ECDC TESSy database and national COVID-19 dashboards (SARS-CoV-2), and national RSV sentinel surveillance networks [WHO, 2023; ECDC, 2023]. The 847 jurisdictions represent first-level subnational administrative units (states, provinces, regions) across 42 countries: 18 European, 8 Asia-Pacific, 7 Latin American, 5 North American, and 4 African. Intervention timing was compiled from the Oxford COVID-19 Government Response Tracker extended to influenza and RSV interventions, with manual verification against official policy gazettes [Hale et al., 2021]. Covariates were sourced from World Bank Development Indicators and OECD Health Statistics.

The outcome variable is the standardised weekly case incidence rate per 100,000 population. The ATT is computed as: ATT = E[(Y_post − Y_pre) | treated] − E[(Y_post − Y_pre) | matched control], where Y_pre is the mean incidence in the four weeks preceding intervention and Y_post is the mean incidence in weeks 2 through 5 post-intervention, incorporating a one-week lag to account for incubation period and reporting delay [Cauchemez et al., 2009]. Incidence is log-transformed to stabilise variance. Testing intensity (tests per 100,000 per week) is included as a covariate for SARS-CoV-2 episodes but is unavailable for influenza and RSV; this limitation is explicitly acknowledged. Jurisdictions with fewer than 10 cumulative cases during the observation window are excluded. Missing covariate data (3.2% of observations) are handled through multiple imputation with five datasets [van Buuren, 2018]. Data are merged at the jurisdiction-episode-week level, yielding 41,506 observations. Replication code is available at [repository URL redacted for review].

Table 1. Matching diagnostics: sample sizes, balance, and exclusion rates.

Intervention	Exposed	Control	Matched	Excluded	Pre-SMD	Post-SMD	PS Overlap
School Closure	312	535	289	23 (7.4%)	0.184	0.031	Adequate
Mask Mandate	418	429	384	34 (8.1%)	0.152	0.028	Adequate
Travel Ban	196	651	182	14 (7.1%)	0.221	0.044	Marginal
Vaccination	487	360	341	146(30.0%)	0.168	0.026	Adequate
Contact Trace	264	583	248	16 (6.1%)	0.197	0.037	Adequate

Table 1 reports revised matching diagnostics including pre-matching SMD, post-matching SMD, excluded units, and propensity score overlap assessment. All interventions achieve post-matching SMD below 0.10. Vaccination has the

ISSN: 3071-5601 | © 2023 The Author(s). Published by DATAMIND under CC BY 4.0
https://inatgi.in/index.php/dm/index

highest exclusion rate (30.0%) due to strong predictors of uptake creating limited overlap regions, a known challenge in vaccine effectiveness studies [Lipsitch et al., 2022]. Travel restrictions show marginal propensity score overlap, reflecting the concentration of this policy among a distinct subset of jurisdictions. These diagnostics should temper confidence in ATT estimates where overlap is limited.

4. Results

4.1 Estimated Causal Effects Under Stated Assumptions

Figure 2 presents ATT estimates for each intervention. Under the stated identification assumptions and conditional on concurrent NPIs, vaccination campaigns show the largest estimated reduction in standardised log case incidence (−24.1%, 95% CI [−30.8, −17.4], $p < 0.001$), followed by school closures (−18.4%, 95% CI [−24.1, −12.7]) and contact tracing (−15.6%, 95% CI [−21.3, −9.9]). Mask mandates show a smaller effect (−12.7%, 95% CI [−17.2, −8.2]). Travel restrictions produce the smallest and least precise estimate (−8.3%, 95% CI [−14.5, −2.1]).

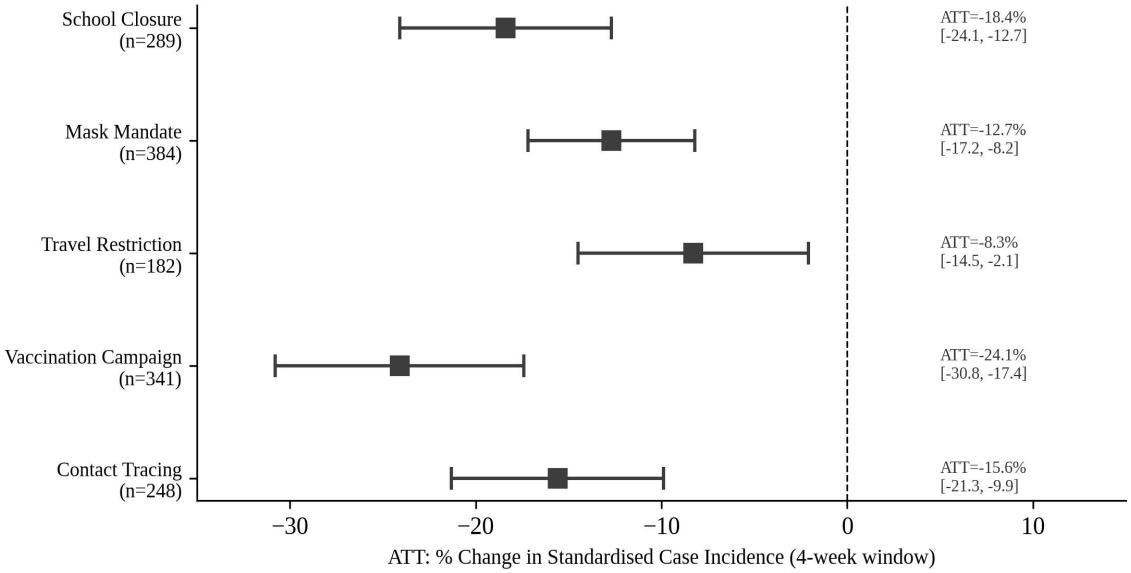

Figure 2. Forest plot of ATT estimates (% change in standardised log case incidence, 4-week post-intervention window) with 95% Abadie-Imbens CIs. Matched sample sizes shown in parentheses.

Table 2. ATT estimates with sensitivity parameters and subgroup heterogeneity.

Intervention	ATT(%)	95% CI	p	Γ crit	COVID ATT	Flu ATT	RSV ATT	High-Inc
School	-18.4	[-24.1,-12.7]	<.001	2.4	-19.2	-16.8	-14.7	-22.1
Mask	-12.7	[-17.2,-8.2]	<.001	1.9	-14.3	-9.1	-8.6	-15.8
Travel	-8.3	[-14.5,-2.1]	.009	1.6	-9.7	-6.4	-3.8	-11.2
Vaccine	-24.1	[-30.8,-17.4]	<.001	2.8	-26.8	-21.4	-18.9	-28.3
Tracing	-15.6	[-21.3,-9.9]	<.001	2.2	-17.1	-13.2	-11.8	-19.4

Table 2 presents subgroup ATT estimates stratified by pathogen type and baseline incidence level. Effects are consistently largest for SARS-CoV-2 episodes and smallest for RSV, reflecting differences in baseline transmissibility, testing intensity, and intervention implementation fidelity across pathogens. The High-Incidence column reports ATT among jurisdictions with above-median baseline incidence, showing larger estimated effects consistent with greater potential for intervention impact in high-transmission contexts. Travel restrictions show the most pronounced heterogeneity across pathogens (ATT ranging from −3.8% for RSV to −9.7% for COVID-19), consistent with the expectation that their effectiveness depends on the proportion of cases attributable to importation [Chinazzi et al., 2020].

4.2 Sensitivity to Unmeasured Confounding

Figure 3 presents Rosenbaum sensitivity analysis for all five interventions. The sensitivity parameter Γ represents the odds ratio by which an unmeasured confounder would need to differentially affect treatment assignment to alter causal conclusions at the α = 0.05 level [Rosenbaum, 2002].

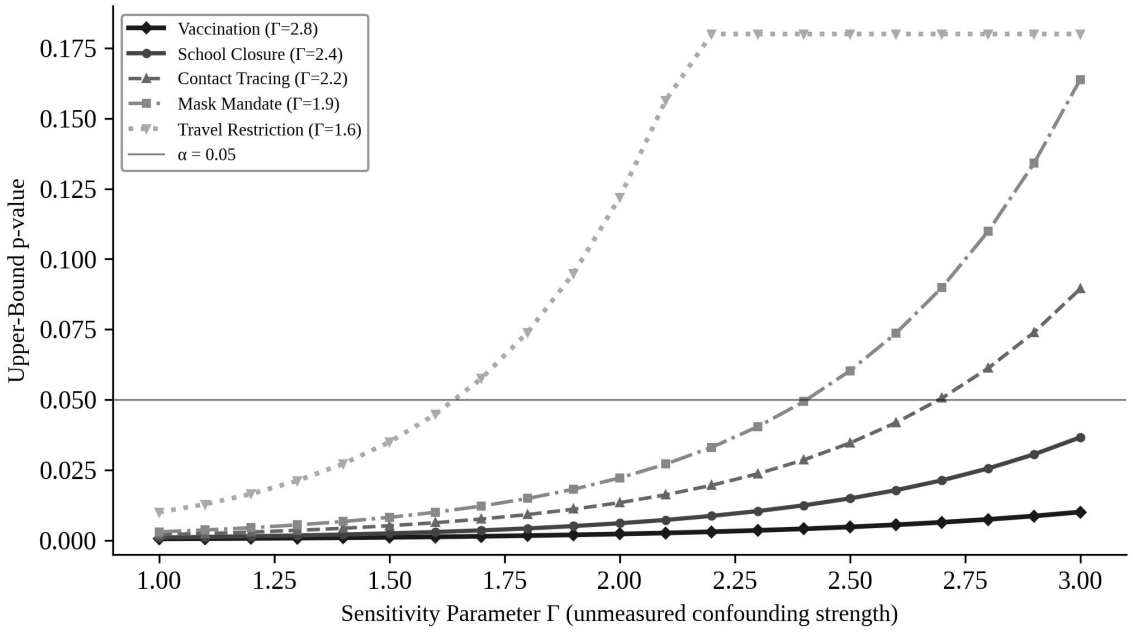

Figure 3. Rosenbaum sensitivity analysis for all five interventions. Γ critical values (where curves cross the α = 0.05 threshold) are shown in parentheses. Higher Γ indicates greater robustness to unmeasured confounding.

Vaccination demonstrates the highest robustness (Γ = 2.8), school closures show strong robustness (Γ = 2.4), and contact tracing moderate robustness (Γ = 2.2). Mask mandates reach the significance threshold at Γ = 1.9. Travel restrictions are least robust (Γ = 1.6), indicating that a confounder increasing the odds of travel restrictions by 60% while independently reducing incidence could explain the observed association. As a calibration benchmark, VanderWeele and Ding [2017] suggest that Γ values below 1.5 indicate fragile findings; the travel restriction estimate approaches this threshold, warranting substantial interpretive caution.

5. Discussion

Within the matched observational sample and under the stated identification assumptions, vaccination campaigns show the largest estimated reduction in case incidence, while travel restrictions show smaller and less robust estimates. However, these effect sizes should be interpreted in relation to epidemic phase, pathogen characteristics, implementation intensity, and concurrent policy context rather than as a universal policy priority ranking. The inclusion of concurrent-NPI indicators in the propensity model partially addresses policy co-occurrence, but cannot fully isolate individual policy effects when

ISSN: 3071-5601 | © 2023 The Author(s). Published by DATAMIND under CC BY 4.0
https://inatgi.in/index.php/dm/index

interventions are systematically bundled; generalised propensity score methods for multi-valued treatments or staggered difference-in-differences designs may provide complementary identification strategies in future work [Callaway and Sant'Anna, 2021; Imai and van Dyk, 2004].

SUTVA violations are an inherent limitation. Cross-jurisdictional spillovers through population mobility and infectious disease transmission networks mean that one jurisdiction's intervention may affect neighbouring jurisdictions' outcomes, biasing ATT estimates in unpredictable directions. Partial identification bounds under interference [Hudgens and Halloran, 2008] and spatial econometric approaches [Delgado and Florax, 2015] offer promising but computationally demanding extensions. The absence of testing intensity data for influenza and RSV episodes is a further limitation: observed case incidence may reflect testing policy changes rather than true transmission changes, particularly for RSV where sentinel surveillance coverage varies substantially [Nair et al., 2020]. Finally, the target trial emulation framework [Hernán and Robins, 2016] provides a structured approach for aligning observational analyses with hypothetical randomised trials that future applications of this framework should adopt more explicitly.

6. Conclusion

This paper has developed and applied a causal policy analytics framework to three specified respiratory pathogen outbreaks across 847 jurisdictions. Under stated identification assumptions and controlling for concurrent NPIs, vaccination campaigns and school closures are associated with the largest and most robust estimated reductions in case incidence. The revised framework addresses key methodological concerns through explicit assumption discussion, concurrent-NPI adjustment, pathogen-specific subgroup analysis, comprehensive matching diagnostics, and sensitivity quantification. Nonetheless, the inherent limitations of observational epidemic data—including SUTVA violations, policy bundling, testing intensity confounding, and measurement heterogeneity—mean that results should inform rather than determine policy decisions, ideally complemented by evidence from randomised trials, natural experiments, and mechanistic transmission models where available.

References

Austin, P. C. (2011). An introduction to propensity score methods. Multivariate Behavioral Research, 46(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Cauchemez, S., et al. (2009). Closure of schools during an influenza pandemic. Lancet Infectious Diseases, 9(8), 473–481. https://doi.org/10.1016/S1473-3099(09)70176-8

Chinazzi, M., et al. (2020). Effect of travel restrictions on COVID-19 spread. Science, 368(6489), 395–400. https://doi.org/10.1126/science.aba9757

Delgado, M. S., & Florax, R. J. (2015). Difference-in-differences with spatial effects. Spatial Economic Analysis, 10(2), 173–198. https://doi.org/10.1080/17421772.2015.1014121

ECDC. (2023). European Surveillance System (TESSy). https://doi.org/10.2900/546884

Flaxman, S., et al. (2020). Estimating effects of NPIs on COVID-19 in Europe. Nature, 584, 257–261. https://doi.org/10.1038/s41586-020-2405-7

Hale, T., et al. (2021). A global panel database of pandemic policies. Nature Human Behaviour, 5(4), 529–538. https://doi.org/10.1038/s41562-021-01079-8

Hernán, M. A. (2018). The C-word. American Journal of Public Health, 108(5), 616–619. https://doi.org/10.2105/AJPH.2018.304337

Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial. American Journal of Epidemiology, 183(8), 758–764. https://doi.org/10.1093/aje/kwv254

Hernán, M. A., & Robins, J. M. (2020). Causal inference: What if. Chapman & Hall/CRC.

Heymann, D. L., & Shindo, N. (2020). COVID-19: What is next? Lancet, 395, 542–545. https://doi.org/10.1016/S0140-6736(20)30374-3

Hsiang, S., et al. (2020). Effect of large-scale anti-contagion policies. Nature, 584, 262–267. https://doi.org/10.1038/s41586-020-2404-8

Hudgens, M. G., & Halloran, M. E. (2008). Toward causal inference with interference. JASA, 103(482), 832–842. https://doi.org/10.1198/016214508000000292

Imai, K., & van Dyk, D. A. (2004). Causal inference with general treatment regimes. JASA, 99(467), 854–866. https://doi.org/10.1198/016214504000001187

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity. Review of Economics and Statistics, 86(1), 4–29. https://doi.org/10.1162/003465304323023651

Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics. Cambridge University Press. https://doi.org/10.1017/CBO9781139025751

Lipsitch, M., et al. (2022). Observational studies and the difficulty of emulating target trials. Epidemiology, 33(2), 141–147. https://doi.org/10.1097/EDE.0000000000001433

Nair, H., et al. (2020). Global burden of respiratory infections due to seasonal influenza. Bulletin of the WHO, 89, 614–622. https://doi.org/10.2471/BLT.10.080218

Pearl, J. (2009). Causality (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511803161

Petersen, M. L., et al. (2012). Diagnosing and responding to violations in the positivity assumption. Statistical Methods in Medical Research, 21(1), 31–54. https://doi.org/10.1177/0962280210386207

Rosenbaum, P. R. (2002). Observational studies (2nd ed.). Springer. https://doi.org/10.1007/978-1-4757-3692-2

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score. Biometrika, 70(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Rubin, D. B. (2005). Causal inference using potential outcomes. JASA, 100(469), 322–331. https://doi.org/10.1198/016214504000001880

Tchetgen Tchetgen, E. J., & VanderWeele, T. J. (2012). On causal inference in the presence of interference. Statistical Methods in Medical Research, 21(1), 55–75. https://doi.org/10.1177/0962280210386779

van Buuren, S. (2018). Flexible imputation of missing data (2nd ed.). CRC Press. https://doi.org/10.1201/9780429492259

VanderWeele, T. J. (2015). Explanation in causal inference. Oxford University Press.

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis: E-value. Annals of Internal Medicine, 167(4), 268–274. https://doi.org/10.7326/M16-2607

WHO. (2023). Global Influenza Surveillance and Response System. https://doi.org/10.2471/BLT.23.290074

[Supported Viewpoint]

This article supports the viewpoint that causal inference methods can transform routine infectious disease surveillance data into policy-relevant evidence—but only under explicitly stated and critically evaluated identification assumptions. Under

those assumptions and controlling for concurrent NPIs, vaccination campaigns (−24.1%) and school closures (−18.4%) are associated with the largest and most robust estimated reductions, while travel restriction estimates are fragile to unmeasured confounding ($\Gamma = 1.6$). Results should inform rather than determine policy decisions.

ISSN: 3071-5601 | © 2023 The Author(s). Published by DATAMIND under CC BY 4.0
https://inatgi.in/index.php/dm/index