American Journal of Epidemiology, 2025, 194, 2249–2260
https://doi.org/10.1093/aje/kwae410
Advance access publication date October 23, 2024
Original Contribution

Evaluating natural course performance in parametric g-formula: review of current practice and illustration based on the United Autoworkers–General Motors cohort

Wenxin Lu*[,1], Sally Picciotto[1], Sadie Costello[1], Hilary Colbeth[2], Ellen Eisen[1]

[1]Division of Environmental Health Sciences, University of California Berkeley, 2121 Berkeley Way, Berkeley, CA 94704, United States
[2]Division of Epidemiology, University of California Berkeley, 2121 Berkeley Way, Berkeley, CA 94704, United States

*Correspondence author: 2121 Berkeley Way, Berkeley, CA 94704 (wluac@berkeley.edu)

Abstract

The parametric g-formula is a causal inference method that appropriately adjusts for time-varying confounding affected by prior exposure. Like all parametric methods, it assumes correct model specification, usually assessed by comparing the observed outcome with the simulated outcome under no intervention (natural course). However, it is unclear how to evaluate natural course performance and whether other variables should also be considered. We reviewed current practices for evaluating model misspecification in applications of the parametric g-formula. To illustrate the pitfalls of current practices, we then applied the parametric g-formula to examine cardiovascular disease mortality in relation to occupational exposure in the United Autoworkers–General Motors cohort (UAW-GM), comparing 20 parametric model sets and qualitatively assessing natural course performance for all time-varying variables over follow-up. We found that current practices of evaluating model misspecification are often insufficient, increasing risk of bias and statistical cherry-picking. Based on our motivational analyses of the UAW-GM cohort, good natural course performance of the outcome does not guarantee good simulations of other covariates; poor predictions of exposures and covariates may still exist. We recommend reporting natural course performance for all time-varying variables at all time points. Objective criteria for evaluating model misspecification in the parametric g-formula need to be developed.

Key words: g-formula; causal inference; g-computation; healthy worker effect; natural course; parametric g-formula.

Background

The parametric g-formula, a form of g-computation, is an increasingly popular statistical method that tackles several common challenges in epidemiologic practices. Under certain conditions and assumptions, it has the potential to permit causal conclusions from observational studies with complex longitudinal data structures.[1-4] It appropriately accounts for time-varying confounder-mediators, where traditional modeling methods fail.[5-7] It can be used to conduct mediation analysis in longitudinal settings.[8-10] It offers flexibility in exposure manipulation and thus can be used to estimate the effects of hypothetical interventions and policies,[11,12] the combined effects of multiple interventions,[13-15] and dynamic treatment strategies.[16,17] In recent years, the parametric g-formula has been widely applied in various fields of epidemiology, including occupational, social, environmental, nutritional, chronic disease, and infectious disease epidemiology.[15,18-28]

The statistical procedures for applying the parametric g-formula in longitudinal settings have been clearly described by Taubman et al.[29] The method simulates and compares counterfactual scenarios using data, including baseline and time-varying risk factors (e.g., exposures, treatments, and covariates that affect the outcomes of interest), time-varying risks (disease incidence or other outcomes of interest), and time-varying competing risks (alternative outcomes such as death that prevent the occurrence of the outcome of interest). Briefly, the implementation takes 4 basic steps:

(1) Modeling: Time-varying risks, risk factors, and competing risks at each time point are modeled as functions of prior histories of risk factors.
(2) Simulation: A hypothetical cohort is simulated using baseline factors and coefficients from previously described models. Data at time t are predicted using data up to time t – 1 and used to simulate data at time t + 1. This simulation with no hypothetical intervention applied is often referred to as the "natural course" scenario.
(3) Intervention: Repeat the abovementioned simulation with an additional step. After data at time t are predicted, manipulate the exposure(s) at time t according to the intervention of interest. The manipulated values are used to predict subsequent data.
(4) Estimation: The population risk under each simulated scenario is calculated as the average of subject-specific risks. Risk under a hypothetical intervention is compared with

Received: June 11, 2024. **Accepted:** October 9, 2024

the risk under the natural course or under an alternative intervention to calculate risk differences or risk ratios.

One major limitation of the parametric g-formula is that, like all parametric methods, the validity of risk estimates relies on the correct specification of all parametric models. This assumption of non-misspecification cannot be tested empirically. In complex data structures with multiple time-varying variables, the correct specification of all parametric forms in the models for all variables can be implausible.

Although the non-misspecification assumption cannot be verified, it can be evaluated by comparing the simulated natural course scenario with the observed data, using nonparametric Kaplan-Meier estimates based on the observed data in the presence of censoring.[30,31] If the parametric models capture the true underlying causal mechanisms, the simulated natural course scenario should match the observed scenario. In other words, the close resemblance between the natural course and the observed scenarios (referred to as good "natural course performance" for simplicity in the rest of this article) is a necessary but insufficient condition of the non-misspecification assumption.

The natural course performance has a multidimensional nature, especially in longitudinal settings. First, there are often many time-varying variables, including exposures, covariates, risks, and competing risks, for which natural course performance can be evaluated. Second, for each variable, there are multiple time points at which natural course performance can be evaluated. Third, for each variable at each time point, the density distribution across individuals or units can also be compared. No study has discussed at which dimension(s) the natural course performance can be sufficiently evaluated. Even for the first dimension, there has been no qualitative or quantitative criterion established for a "good" or "acceptable" natural course performance. The evaluation of natural course performance is not compulsory in published applications of the parametric g-formula but is considered a "good practice."[31] Whether and how to assess the potential violation of the key non-misspecification assumption, on which the validity of effect estimates relies, has been largely left to the discretion of the authors.

Here, we first review the existing literature on parametric g-formula applications in observational longitudinal settings and summarize the current practices of parametric model form selection and natural course performance evaluation. After that, we use the United Autoworkers–General Motors cohort (UAW-GM) study as a motivating example to explore the limitations of these common practices as support for the non-misspecification assumption of the g-formula.

Review of current practice

Methods

We reviewed current practices in applying the parametric g-formula in epidemiologic studies. Using the Sysrev online platform,[32] we retrieved PubMed-indexed articles published in and before October 2023 that included "parametric g-formula" or "parametric g-computation" in the title or abstract. We further restricted to articles that met all the following criteria:

(1) Epidemiologic applications of the parametric g-formula. Review articles, commentaries, statistical tutorials, and method comparisons were excluded.
(2) Observational studies. Studies that utilized data collected from randomized controlled trials were excluded.
(3) Longitudinal data structure.

We reviewed and summarized the following aspects of the selected studies:

(A) The type of exposure and outcome variables being studied. If multiple exposures or outcomes were examined in 1 study, the exposure or outcome class of the majority was recorded.
(B) The purposes of applying the parametric g-formula method, as stated in the methods section while describing the method and/or in the discussion section while discussing the advantages of the method.
(C) Whether and how the authors argued for the general non-misspecification of parametric forms in their g-formula models. If the authors compared the simulated natural course scenario and the observed scenario, we further reported what variables were compared and whether the comparison was made at the end of follow-up only, at all time points, or as a summary mean or median across all time points.

Results

Figure 1 summarizes the roadmap for the literature search. We retrieved 192 articles with "parametric g-formula" or "parametric g-computation" keywords in the abstract or title, among which 114 were epidemiologic applications of the parametric g-formula. After restricting to observational studies with a longitudinal data structure, 86 articles were selected for further review.[11-15,17-29,33-100]

The 86 articles studied the following exposure classes: behavioral exposures such as diet, smoking, and alcohol consumption (27%); clinical exposures such as treatment types and strategies (26%); occupational exposures, including workplace exposures, unionization, and retirement (21%); social exposures such as education, urbanization, and poverty (9%); environmental exposures such as air pollution and heavy metal (7%); and other exposures (10%). The classes of outcomes studied were all-cause mortality (33%), cardiovascular outcomes (16%), psychological outcomes (10%), cancer incidence or mortality (8%), neurological outcomes (7%), clinical outcomes (5%), nonmalignant chronic diseases (5%), and other outcomes (16%). Four main purposes for applying the parametric g-formula were reported: addressing time-varying confounding (85%), applying hypothetical interventions (76%), supporting causal conclusions (42%), and conducting causal mediation analyses (10%).

Good natural course performance is the most common evidence offered to support the non-misspecification assumption, but the amount of supporting information presented greatly varies. Table S1 summarizes the information reported for natural course performance in the 86 eligible studies. A total of 63 (73%) studies reported some comparison between the natural course and the observed scenario, and all of them reported comparisons for the outcome variable; fewer reported comparisons for the exposure variables (42%) or time-varying covariates (41%). Forty-four studies reported comparisons at all time points for the outcome variables. Others compared simulated variables at the last time point only or compared a single mean or median across all time points. Only 25 (29%) of studies reported natural course performance for all variables at all time points.

The time trends in reporting natural course performance for outcome variables are illustrated in Figure 2. Since the first eligible study was published in 2009 and the second in 2012, we observed an increasing number of parametric g-formula applications in observational studies with a longitudinal data structure. However, as the method has become more popular, less

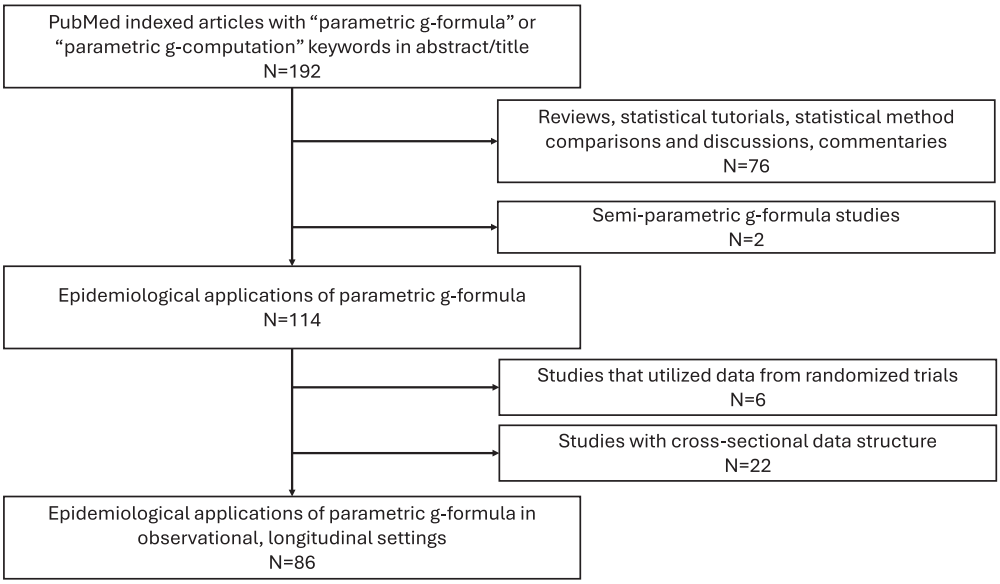

Figure 1. Roadmap of a literature search and inclusion for epidemiologic applications of the parametric g-formula in observational and longitudinal settings. Note: All studies were initially retrieved on the same day in October 2023.

evidence in support of the non-misspecification assumption has been reported. All studies published in or before 2017 reported some information on the natural course performance of the outcome variable. Since 2018, the proportion of studies that reported little to no information on natural course performance has been increasing.

Other methods have been proposed for model selection or evaluation of model misspecification. Sadinski et al[85] compared the parametric g-formula with the inverse probability of treatment weighting and found similar risk estimates. They also selected flexible g-formula models based on the Akaike information criterion (AIC).[85] Inoue et al[71] used machine learning algorithms to find the best predictive model, applied 10-fold cross-validation to minimize the risk of overfitting, and measured the prediction performance with the area under the receiver operating characteristic curve (AUC) and the confusion matrix. Mathisen et al[60] evaluated model predictivity using concordance statistics and assessed model generalizability using internal-external cross-validation. Ibsen et al[46] conducted a sensitivity analysis using different model specifications and found stable results across model sets attempted.

Overall, there has been no consensus on adequate model selection and methods for misspecification risk evaluation among published parametric g-formula applications. Most published applications present only natural course performance for the outcome as evidence against model misspecification.

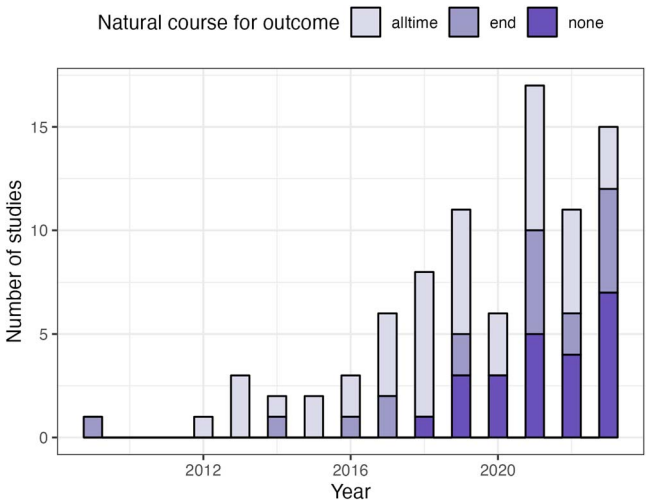

Figure 2. Natural course performance for outcome variables reported in studies that applied the parametric g-formula in observational and longitudinal settings, summarized across calendar years. Note: *All time*: The distributions of the outcome variable in the observed scenario and the natural course scenario were compared at every time point. *End*: The distributions of the outcome variable in the observed scenario and the natural course scenario were compared only at the last time point. *None*: No comparison between the observed and the natural course scenario was reported for the outcome variable.

Motivating example

The UAW-GM cohort

The UAW-GM cohort has been described extensively.[101-105] Briefly, the original cohort included 46 306 hourly workers hired between 1938 and 1982 who worked for at least 3 years at 3 automobile manufacturing plants in Michigan. We restricted this motivating example to those who worked in the Detroit plant. For each worker, follow-up started at the fourth year of employment and ended at death, age 94, or December 31, 1994, whichever occurred first. The final analytic data set included 9288 workers, contributing to 259 671 person-years. Participant characteristics were obtained from company job records and are summarized in Table 1. Baseline characteristics include sex, race, year of entering the cohort, and total metalworking fluid (MWF) exposure during the first 3 years of employment. Time-varying covariates include age, calendar year, and employment status (active vs inactive).

The exposure of interest is total MWF exposure. The quantitative exposure assessment has been described in previous publications.[105-107] Briefly, a time-varying job exposure matrix was constructed based on hundreds of personal and area samples collected in the 1980s across jobs and departments, as well as historical industrial hygiene records. For each person-year, annual time-weighted average daily occupational exposures to straight, soluble, and synthetic MWFs were estimated based on the job exposure matrix and detailed job records. Annual total MWF

Table 1. Summary of study population characteristics and metalworking fluid (MWF) exposures in the United Autoworkers–General Motors cohort (UAW-GM).

	Study population ($n = 9288$)		
	n	**%**	**Median (IQR)**
Race/ethnicity			
Black	3353	36.1	
White	4451	47.9	
Unknown	1284	13.8	
Sex			
Female	530	5.7	
Male	8758	94.3	
Year hired			1960 (1950-1969)
Age hired (years)			26 (21-34)
Duration of employment (years)			16 (5-23)
Duration of follow-up			30 (22-38)
Annual MWF exposure during employment (mg/m^3)			0.4 (0.2-1.3)
Lifetime cumulative MWF exposure (mg/m^3-yr)			11.6 (5.2-25.7)

Note: We restricted to hourly workers hired between 1938 and 1982, who worked for at least 3 years in the Detroit plant, were at least 16 years old when hired, and were missing less than half of the employment history.

exposure (mg/m^3) was calculated as the sum of exposures to the 3 MWF classes. Cumulative total MWF exposure at each person-year was calculated as rolling sums of the annual total MWF exposures.

Vital statuses were determined by linking with the Social Security Administration, death certificates, plant records, state mortality files, and the National Death Index. Causes of death were determined from state vital records, death certificates, and the National Death Index Plus.

Methods

We applied the parametric g-formula to assess the relationship between occupational MWF exposure and cardiovascular disease (CVD) mortality using the SAS GFORMULA macro.[108] The macro applies Monte Carlo simulation and bootstrap sampling. It supports 3 types of outcomes, 4 types of interventions, 7 types of dependent variable transformations, and various independent variable specifications with flexible knots. We explored 20 model sets with different parametric specifications and assessed the natural course performances. We simulated CVD mortality and competing risk (deaths due to other causes) as outcomes, annual and cumulative MWF as exposures, follow-up year, calendar year, age, and leaving work as time-varying covariates. A simplified directed acyclic graph illustrating the causal pathways in this analysis is presented in Figure S2. All parametric forms in all model sets are summarized in Appendix S3, and the sample code for 1 model set is provided in Appendix S4. For each time-varying variable in each model, we plotted the annual average in the natural course scenario and the observed scenario over follow-up years. We visually assessed natural course performance for all time-varying variables by comparing how closely the natural course curve resembles the observed curve and classified natural course performance into 4 categories: (1) good, (2) slight deviation, (3) moderate deviation, and (4) severe deviation. Qualitative criteria for this visual categorization and examples of natural course performance in each category are described in Figure S5. We then categorized the 20 model sets into 4 natural course performance tiers (1-4, from best to worst) based on the overall natural course performances of all modeled time-varying variables. Note that if the natural course plot for a particular covariate X shows deviation from the observed averages over time, this does not necessarily imply that the model for X is misspecified. If X depends very strongly on another covariate L, and the model for L is misspecified, then the natural course performance of X will suffer even if X is modeled correctly. Thus, the assessment of natural course performance applies to the entire set of models and is based on the entire set of plots for all covariates. Apart from the qualitative visual assessment of natural course performance, we also calculated the sum of AICs across the prediction models for all simulated variables and evaluated whether it can be used as a quantitative measure to select model sets with satisfactory natural course performance.

Finally, we applied a hypothetical intervention that sets a maximum annual MWF exposure threshold at 0.5 mg/m^3 (the National Institute for Occupational Safety and Health recommended exposure limit) and recorded the estimated intervention effects under these 20 model sets. This threshold intervention is more dependent on the simulated natural values of the exposure, compared to deterministic interventions that set all exposures to certain values. Risk ratios (RRs) of CVD mortality at 60, 70, and 80 years old comparing the intervention scenario to the natural course scenario were recorded. We hypothesize that model sets producing better natural course performances capture the underlying causal relationships between time-varying variables better, produce point estimates closer to the true causal effects, and also produce point estimates closer to each other.

Results

The natural course performance for all time-varying variables varied considerably across the 20 parametric model sets evaluated, although most (17 of 20) had good natural course performance for the outcome variable (Table S6). Figure 3 illustrates the natural course performance for a tier 1 (best-performing) model set (model set 2): The time-varying distributions of all variables behave similarly under the natural course scenario and the observed scenario. Figure 4 illustrates the natural course performance for a tier 4 (worst-performing) model set (model set 20). Although the outcome variable was simulated close to the observed scenario, there are apparent deviations between the natural course and the observed scenarios for other variables.

This poor natural course performance for model set 20 indicates that those parametric models do not reflect the true causal

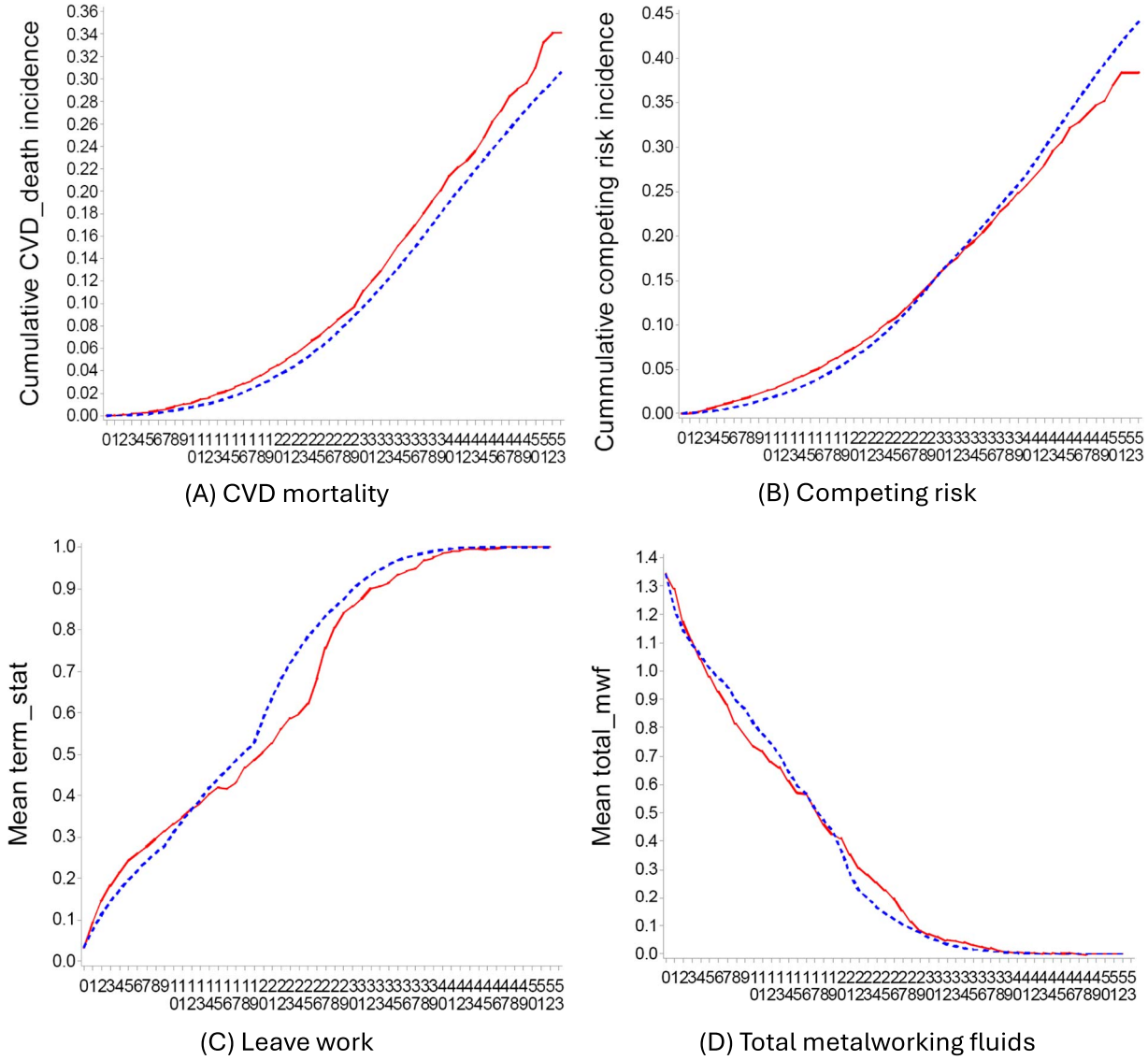

Figure 3. Natural course (dashed line) vs observed (solid line) scenarios for time-varying variables over follow-up years in one of the best-performing parametric g-formula model sets (#2). Note: The x-axis indicates follow-up years. The y-axis indicates the average level of modeled time-varying variables.

relationships among the time-varying variables and implies that the non-misspecification assumption is violated. This failure at the modeling stage, which is the first and fundamental step of the parametric g-formula, causes potentially poor simulations of the counterfactual scenarios of interest at the simulation and intervention stages, resulting in biased effect estimates at the estimation stage. This potential bias is more apparent if the intervention of interest is dependent on the simulated natural value of exposure. For example, MWF exposure is severely underpredicted in model set 20. A hypothetical intervention that sets a maximum annual MWF exposure threshold of 0.5 mg/m^3 will be intervening on much fewer person-years than would have been affected under the observed scenario; this would yield biased estimates for the causal effect of interest.

We calculated the sum of AICs across all modeled time-varying variables and found that a lower sum of AICs at the modeling stage is not a sensitive indication of better natural course performance at the simulation stage (Table S6). Although the average sum of AICs is lower among the tier 1 (best-performing) models compared with tier 2 and tier 3 models, some of the tier 4 models produced a lower sum of AICs than some tier 1 models. For example, if we select a parametric model set with the lowest sum of AICs across all modeled variables, we will end up with model set 5, which has the worst overall natural course performance among the 20 model sets attempted (Figure S7).

Figure 5 summarizes the estimated RRs for the hypothetical intervention of setting a maximum annual MWF exposure limit of 0.5 mg/m^3 at different ages by natural course performance tiers. At the selected ages of 60, 70, and 80 years, the tier 1 model sets produced similar RRs with a narrow range (eg, 0.96-0.98 at age 60). By contrast, all model sets yielded a much larger range of RRs (0.92-1.04), indicating potential model misspecification bias in both directions for models with unsatisfactory natural course performances.

Discussion

We applied the parametric g-formula to analyze the relationship between MWF exposure and CVD mortality in the UAW-GM cohort and qualitatively evaluated the natural course

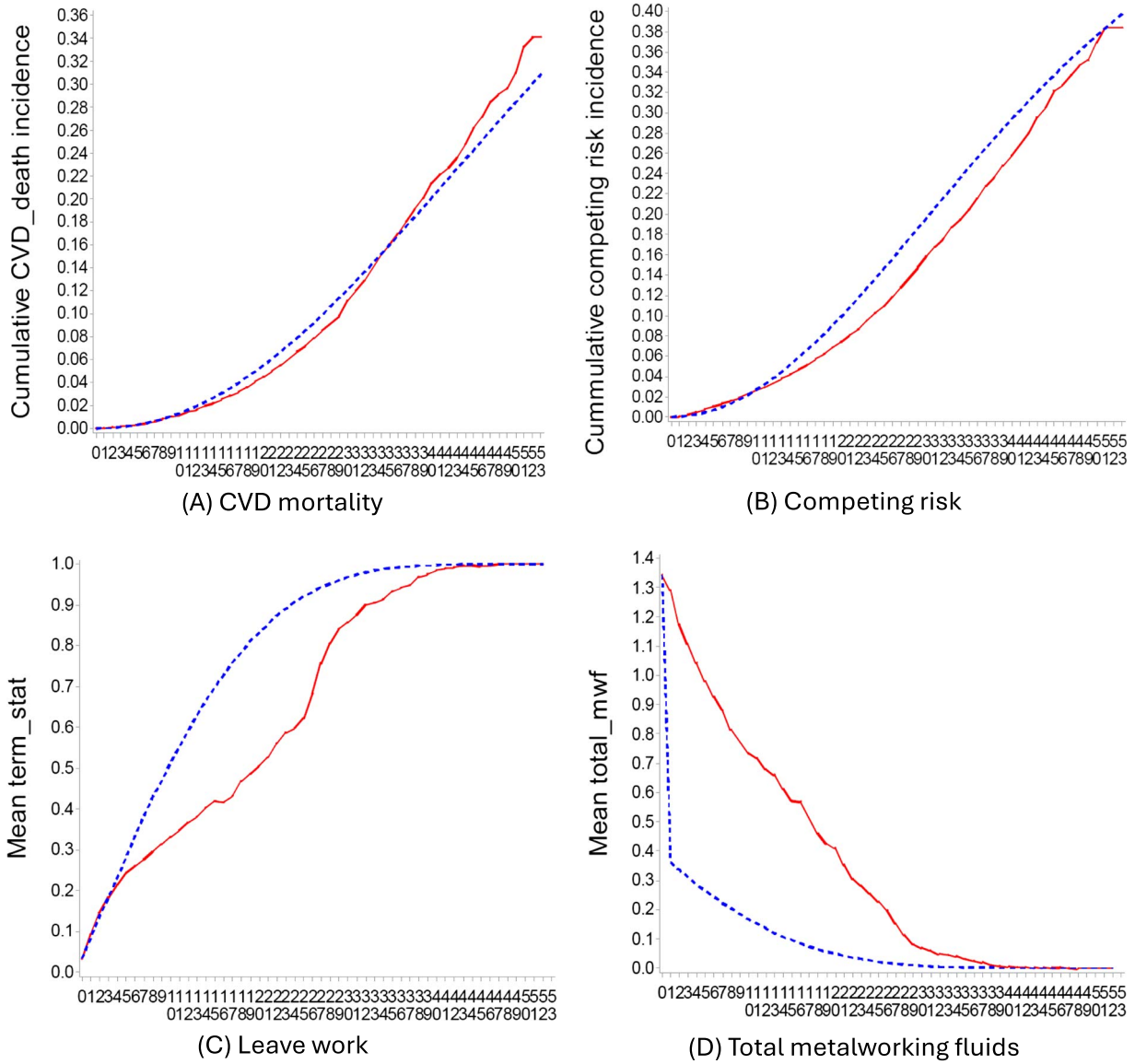

Figure 4. Natural course (dashed line) vs observed (solid line) scenarios for time-varying variables over follow-up years in one of the worst-performing parametric g-formula model sets (#20). Note: The x-axis indicates follow-up years. The y-axis indicates the average level of modeled time-varying variables.

performance under different parametric specifications. More detailed exposure-response analyses on occupational MWF exposures and CVD outcomes in the full UAW-GM cohort have been published elsewhere.[109-112] We found that even if the variables being modeled and the sequence in which they are modeled remain the same, the performance of natural course simulation can vary substantially under different model specifications. A poor natural course performance (Figure 4) indicates that the parametric models specified do not capture the true causal relationship among time-varying variables, that the non-misspecification assumption of the parametric g-formula is violated, and that the effect estimates of the hypothetical interventions of interest are likely to be invalid (an unlikely exception is described in Appendix S8). We also found that model sets with better natural course performances produce effect estimates that are similar to one another (Figure 5), consistent with our hypothesis that good natural course performance indicates better capture of the true causal relationship and results with less risk of bias.

We concur with the previous recommendations by Rudolph et al[31]: Models should be chosen based on the best natural course fit as long as all necessary variables are considered. We further recommend that applications of the parametric g-formula include reports of natural course performance for all modeled time-varying variables at all follow-up time points to minimize the risk of model misspecification. Reporting natural course performance only for the outcome variable is insufficient: Even if the simulated outcome variable under the natural course closely resembles the observed data, the time-varying exposure variables and covariates may still be poorly simulated (Figure 4). Reporting natural course performance for all time-varying variables only at the end of the follow-up or at a single time point is also insufficient; in the poorly performing natural course simulations, the average value of the time-varying variables under the natural course and the observed scenario may still converge at the last time point (Figure 4). The evaluation of natural course performance at all time points is especially important if the research question involves comparing outcome risks at time points other than the

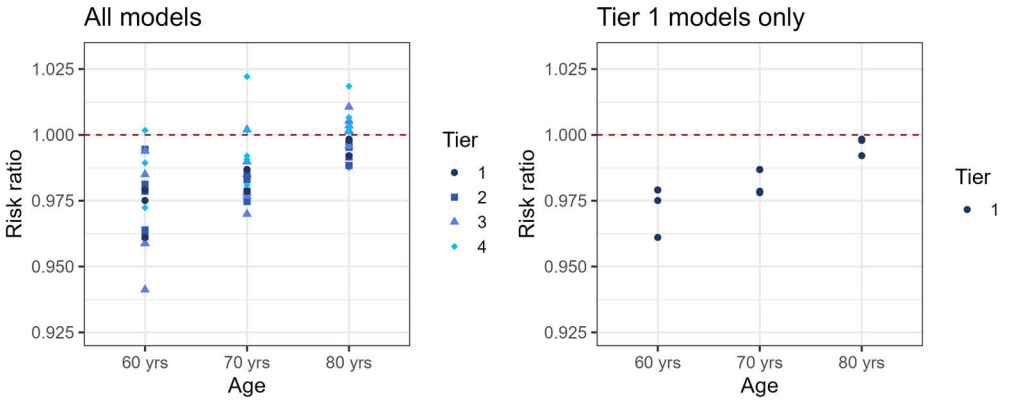

Figure 5. Estimated CVD mortality risk ratios of a hypothetical metalworking fluid exposure intervention, summarized by tiers of natural course performance of parametric 20 model sets. Note: The hypothetical intervention sets a maximum annual total metalworking fluid exposure of 0.5 mg/m^3 over all follow-up years. The risk ratios summarize the reduced cardiovascular mortality risks at different ages under the hypothetical intervention, compared to no intervention (natural course).

end of follow-up. In addition, the population size will shrink over time in closed cohorts with long follow-ups. In the UAW-GM study, the average observed risk and competing risk curves became less smooth after 45 years of follow-up (Figure 3) because the observed population size shrank from 9288 to 1080 after 45 years and to 162 after 50 years of follow-up. The observed distributions of the variables have much smaller sample sizes and larger standard errors toward the end of the follow-up. Thus, the natural course performance at the last time point should not be considered more important than other time points.

Among published applications of the parametric g-formula in observational longitudinal settings, reporting natural course performance is the most common way to demonstrate model validity and non-misspecification. However, the amount of information presented varies across studies, without a common standard. If we consider satisfactory natural course performance for all time-varying variables at all time points as a necessary (but not sufficient) condition for model non-misspecification, then 71% of studies reviewed did not report enough information to evaluate this condition. This underreporting of natural course performance could increase the risk of biased estimates, as well as statistical cherry-picking. In addition, few studies reported a priori selection plans for model parametric forms. If neither the natural course performance nor the model selection processes are reported, authors could potentially explore many different model sets with different parametric forms and report only the one with the largest effect estimate, even though the natural course simulation of the selected model set performs poorly, and the non-misspecification assumption appears violated. For example, the model set that yielded the largest reduction of CVD risk at 60 years old was model set 1 (RR = 0.92), although it is a tier 3 model that had moderate deviations between the natural course and the observed scenarios for MWF exposure (Figure S9). Model set 1 should either be rejected or, if no better model set is identified, be presented with strong caveats about its validity.

An important limitation of our graphical assessment of natural course performance is that it is qualitative, subjective, and bounded by what we observed from the 20 model sets attempted. No quantitative criterion for good natural course performance has been proposed. We considered whether the sum of AICs could be used as a quantitative assessment of the non-misspecification assumption. However, in the UAW-GM example, a lower sum of AICs was not a sensitive indicator of better overall natural course performance. Another shortcoming of AIC in the parametric g-formula is that AICs are not directly comparable between models for different dependent variables.

Quantitative measurements of natural course performance or total model misspecification at the simulation stage need to be developed and should (1) have magnitudes that are comparable across variables, (2) account for the different number of observations at each time point (e.g., shrinkage of population size over long follow-ups), and (3) consider applying cross-validation or other methods to avoid overfitting. The model selection methods proposed by Inoue et al[71] and Mathisen et al[60] that incorporated AUC or concordance statistics with machine learning and cross-validation could be a good starting point for developing quantitative measures of natural course performance. For example, AUC may be considered because of its finite range from 0 to 1, simple interpretation, and generalizability, which meet the criterion (1) described above. Cross-validation is preferable to prevent overfitting during modeling, which can lead to poor simulation performance. Before such quantitative measures become available, researchers could also assess the risk of bias due to potential model misspecification by comparing the effect estimates from different parametric model sets, as demonstrated by Ibsen et al.[46]

Reliance on the non-misspecification assumption is an important limitation of the parametric g-formula and parametric methods in general. However, parametric methods have advantages over their nonparametric counterparts, such as higher statistical power and providing informative estimates with higher generalizability.[113-115] For example, Yu et al[116] found that compared with the parametric g-formula, nonparametric targeted minimum loss-based estimation was more sensitive to the inclusion of different covariates. Alternatively, the iterated conditional expectations (ICE) g-formula is another version of the parametric g-formula, which requires only the outcome models to be correctly specified and relaxes the non-misspecification assumption.[117,118] However, the ICE g-formula is less flexible, does not simulate the natural values of variables, and is therefore less applicable when evaluating interventions that depend on the natural value of the exposure.

By pointing out the potential variations of the natural course performance under different parametric specifications, we advocate for more standardized and transparent evaluation practices for the non-misspecification assumption in future applications of the parametric g-formula. Methods for objective, quantitative evaluation of model misspecification in the parametric g-formula

need to be developed to promote more widespread use of this powerful analytic approach to causal inference.

Acknowledgments

A prior version of this work was presented as a poster presentation titled "Quality Assessment of the Natural Course Simulation of the Parametric G-formula: An Illustration Using the Diesel Exhaust in Miners Study (DEMS)" at the 56th Annual Meeting of the Society for Epidemiologic Research in Portland, OR, United States, in June 2023.

Supplementary material

Supplementary material is available at the *American Journal of Epidemiology* online.

Funding

W.L. was supported by the Institute for Research on Labor and Employment (IRLE) faculty award at the University of California, Berkeley. E.E., S.C., and S.P. were supported by the National Institute for Occupational Safety and Health (NIOSH)/Centers for Disease Control and Prevention (CDC) [grant number R01OH011092]. E.E., S.C., and H.C. are additionally supported by a NIOSH Training Grant [grant number T42OH008429].

Conflict of interest

The authors declare no conflicts of interest.

Data availability

The data presented in this study are available on request from the corresponding author.

References

1. Snowden JM, Rose S, Mortimer KM. Implementation of g-computation on a simulated data set: demonstration of a causal inference technique. *Am J Epidemiol*. 2011;173(7):731-738. https://doi.org/10.1093/aje/kwq472
2. Neugebauer R, van der Laan MJ. G-computation estimation for causal inference with complex longitudinal data. *Comput Stat Data Anal*. 2006;51(3):1676-1697. https://doi.org/10.1016/j.csda.2006.06.016
3. Kawahara T, Shiba K, Tsuchiya A. Application of causal inference methods in the analysis of observational neurosurgical data: g-formula and marginal structural model. *World Neurosurg*. 2022;161:310-315. https://doi.org/10.1016/j.wneu.2021.09.141
4. Igelström E, Craig P, Lewsey J, et al. Causal inference and effect estimation using observational data. *J Epidemiol Community Health*. 2022;76(11):960-966. https://doi.org/10.1136/jech-2022-219267
5. Buckley JP, Keil AP, McGrath LJ, et al. Evolving methods for inference in the presence of healthy worker survivor bias. *Epidemiology*. 2015;26(2):204-212. https://doi.org/10.1097/EDE.0000000000000217
6. Brown DM, Picciotto S, Costello S, et al. The healthy worker survivor effect: target parameters and target populations. *Curr Environ. Health Rep*. 2017;4(3):364-372. https://doi.org/10.1007/s40572-017-0156-x
7. Mansournia MA, Etminan J, Danaei G, et al. Handling time varying confounding in observational research. *BMJ*. 2017;359:j4587. https://doi.org/10.1136/bmj.j4587
8. Lin SH, Young J, Logan R, et al. Parametric mediational g-formula approach to mediation analysis with time-varying exposures, mediators, and confounders. *Epidemiology*. 2017;28(2):266-274. https://doi.org/10.1097/EDE.0000000000000609
9. VanderWeele TJ, Tchetgen Tchetgen EJ. Mediation analysis with time varying exposures and mediators. *J R Stat Soc Series B Stat. Methodol*. 2017;79(3):917-938. https://doi.org/10.1111/rssb.12194
10. Wang A, Arah OA. G-computation demonstration in causal mediation analysis. *Eur J Epidemiol*. 2015;30(10):1119-1127. https://doi.org/10.1007/s10654-015-0100-z
11. Vangen-Lonne AM, Ueda P, Gulayin P, et al. Hypothetical interventions to prevent stroke: an application of the parametric g-formula to a healthy middle-aged population. *Eur J Epidemiol*. 2018;33(6):557-566. https://doi.org/10.1007/s10654-017-0344-x
12. Mokhayeri Y, Hashemi-Nazari SS, Khodakarim S, et al. Effects of hypothetical interventions on ischemic stroke using parametric g-formula. *Stroke*. 2019;50(11):3286-3288. https://doi.org/10.1161/STROKEAHA.119.025749
13. Danaei G, Pan A, Hu FB, et al. Hypothetical midlife interventions in women and risk of type 2 diabetes. *Epidemiology*. 2013;24(1):122-128. https://doi.org/10.1097/EDE.0b013e318276c98a
14. Ren J, Li G, Zhang L, Zhang N, et al. Hypothetical interventions for falls among older adults: an application of the parametric g-formula. *Front Med (Lausanne)*. 2021;8:732136. https://doi.org/10.3389/fmed.2021.732136
15. Laouali N, Benmarhnia T, Oulhote Y. Potential benefits of joint hypothetical interventions on diet, lead, and cadmium on mortality in US adults. *Environ Health*. 2022;21(1):93. https://doi.org/10.1186/s12940-022-00905-4
16. Young JG, Cain LE, Robins JM, et al. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. *Stat Biosci*. 2011;3(1):119-143. https://doi.org/10.1007/s12561-011-9040-7
17. Zhang Y, Young JG, Thamer M, et al. Comparing the effectiveness of dynamic treatment strategies using electronic health records: an application of the parametric g-formula to anemia management strategies. *Health Serv Res*. 2018;53(3):1900-1918. https://doi.org/10.1111/1475-6773.12718
18. Neophytou AM, Costello S, Picciotto S, et al. Diesel exhaust, respirable dust, and ischemic heart disease: an application of the parametric g-formula. *Epidemiology*. 2019;30(2):177-185. https://doi.org/10.1097/EDE.0000000000000954
19. Edwards JK, McGrath LJ, Buckley JP, et al. Occupational radon exposure and lung cancer mortality: estimating intervention effects using the parametric g-formula. *Epidemiology*. 2014;25(6):829-834. https://doi.org/10.1097/EDE.0000000000000164
20. Edmonds A, Breskin A, Cole SR, et al. Poverty, deprivation, and mortality risk among women with HIV in the United States. *Epidemiology*. 2021;32(6):877-885. https://doi.org/10.1097/EDE.0000000000001409
21. Blaikie K, Eisenberg-Guyot J, Andrea SB, et al. Differential employment quality and educational inequities in mental health: a causal mediation analysis. *Epidemiology*. 2023;34(5):747-758. https://doi.org/10.1097/EDE.0000000000001629
22. Letellier N, Gutierrez LA, Duchesne J, et al. Air quality improvement and incident dementia: effects of observed and

hypothetical reductions in air pollutant using parametric g-computation. *Alzheimers Dement*. 2022;18(12):2509-2517. https://doi.org/10.1002/alz.12606
23. Chiu YH, Chavarro JE, Dickerman BA, et al. Estimating the effect of nutritional interventions using observational data: the American Heart Association's 2020 dietary goals and mortality. *Am J Clin Nutr*. 2021;114(2):690-703. https://doi.org/10.1093/ajcn/nqab100
24. Lajous M, Willett WC, Robins J, et al. Changes in fish consumption in midlife and the risk of coronary heart disease in men and women. *Am J Epidemiol*. 2013;178(3):382-391. https://doi.org/10.1093/aje/kws478
25. Nianogo RA, Arah OA. Investigating the role of childhood adiposity in the development of adult type 2 diabetes in a 64-year follow-up cohort: an application of the parametric g-formula within an agent-based simulation study. *Epidemiology*. 2019;30(suppl 2):S101-S109. https://doi.org/10.1097/EDE.0000000000001062
26. Garcia-Aymerich J, Varraso R, Danaei G, et al. Incidence of adult-onset asthma after hypothetical interventions on body mass index and physical activity: an application of the parametric g-formula. *Am J Epidemiol*. 2014;179(1):20-26. https://doi.org/10.1093/aje/kwt229
27. Lodi S, Klein M, Rauch A, et al. Sustained virological response after treatment with direct antiviral agents in individuals with HIV and hepatitis C co-infection. *J Int AIDS Soc*. 2022;25(12):e26048. https://doi.org/10.1002/jia2.26048
28. Edwards JK, Cole SR, Westreich D, et al. Age at entry into care, timing of antiretroviral therapy initiation, and 10-year mortality among HIV-seropositive adults in the United States. *Clin Infect Dis*. 2015;61(7):1189-1195. https://doi.org/10.1093/cid/civ463
29. Taubman SL, Robins JM, Mittleman MA, et al. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol*. 2009;38(6):1599-1611. https://doi.org/10.1093/ije/dyp192
30. Chiu Y-H, Wen L, McGrath S, et al. Evaluating model specification when using the parametric g-formula in the presence of censoring. *Am J Epidemiol*. 2023;192(11):1887-1895. https://doi.org/10.1093/aje/kwad143
31. Rudolph JE, Cartus A, Bodnar LM, et al. The role of the natural course in causal analysis. *Am J Epidemiol*. 2022;191(2):341-348. https://doi.org/10.1093/aje/kwab248
32. Bozada T, Borden J, Workman J, et al. Sysrev: a FAIR platform for data curation and systematic evidence review. *Front Artif Intell*. 2021;4:685298. https://doi.org/10.3389/frai.2021.685298
33. Takeuchi M, Ogura M, Inagaki N, et al. Initiating SGLT2 inhibitor therapy to improve renal outcomes for persons with diabetes eligible for an intensified glucose-lowering regimen: hypothetical intervention using parametric g-formula modeling. *BMJ Open Diabetes Res Care*. 2022;10(3):e002636. https://doi.org/10.1136/bmjdrc-2021-002636
34. Laouali N, Benmarhnia T, Lanphear BP, et al. Associations with blood lead and urinary cadmium concentrations in relation to mortality in the US population: a causal survival analysis with g-computation. *Toxics*. 2023;11(2):133. https://doi.org/10.3390/toxics11020133
35. Pierce M, Bai Y, Nevriana A, et al. Prevention of childhood adversities and children's common mental disorders and school grades. *JAMA Netw Open*. 2023;6(10):e2336408. https://doi.org/10.1001/jamanetworkopen.2023.36408
36. Mokhayeri Y, Nazemipour M, Mansournia MA, et al. Does weight mediate the effect of smoking on coronary heart disease? Parametric mediational g-formula analysis. *PLoS One*. 2022;17(1):e0262403. 10.1371/journal.pone.0262403
37. Takeuchi M, Shinkawa K, Yanagita M, et al. Hypothetical intervention of targeted systolic blood pressure control of < 120 mmHg on renal prognosis for persons with stage 3-4 chronic kidney disease: an application of parametric g-formula using health checkup data in Japan. *Clin Exp Nephrol*. 2023;27(6):542-547. https://doi.org/10.1007/s10157-023-02341-1
38. Nilsen L, Sharashova E, Løchen ML, et al. Hypothetical interventions and risk of atrial fibrillation by sex and education: application of the parametric g-formula in the Tromso study. *Eur J Prev Cardiol*. 2023;30(16):1791-1800. https://doi.org/10.1093/eurjpc/zwad240
39. Chen H, Quick M, Kaufman JS, et al. Impact of lowering fine particulate matter from major emission sources on mortality in Canada: a nationwide causal analysis. *Proc Natl Acad Sci U S A*. 2022;119(49):e2209490119. https://doi.org/10.1073/pnas.2209490119
40. Zhang N, Wang C, Li Y, et al. Hypothetical interventions on risk factors for depression among middle-aged and older community-dwellers in China: an application of the parametric g-formula in a longitudinal study. *J Affect Disord*. 2023;327:355-361. https://doi.org/10.1016/j.jad.2023.01.113
41. Cupp MA, Beaudoin FL, Hayes KN, et al. Post-acute care setting after hip fracture hospitalization and subsequent opioid use in older adults. *J Am Med Dir Assoc*. 2023;24(7):971-977.e4. https://doi.org/10.1016/j.jamda.2023.03.012
42. Nisen J, Bijlsma MJ, Martikainen P, et al. The gendered impacts of delayed parenthood: a dynamic analysis of young adulthood. *Adv Life Course Res*. 2022;53:100496. https://doi.org/10.1016/j.alcr.2022.100496
43. Bornhorst C, Pigeot I, De Henauw S, et al. The effects of hypothetical behavioral interventions on the 13-year incidence of overweight/obesity in children and adolescents. *Int J Behav Nutr Phys Act*. 2023;20(1):100. https://doi.org/10.1186/s12966-023-01501-6
44. Chammartin F, Lodi S, Logan R, et al. Risk for non-AIDS-defining and AIDS-defining cancer of early versus delayed initiation of antiretroviral therapy: a multinational prospective cohort study. *Ann Intern Med*. 2021;174(6):768-776. https://doi.org/10.7326/M20-5226
45. Sudharsanan N, Ho JY. Rural-urban differences in adult life expectancy in Indonesia: a parametric g-formula-based decomposition approach. *Epidemiology*. 2020;31(3):393-401. https://doi.org/10.1097/EDE.0000000000001172
46. Ibsen DB, Chiu YH, Gemes K, et al. Hypothetical 22-year intervention with DASH diet lowered risk of heart failure in a general population. *Am J Epidemiol*. 2024;193(1):96-106. https://doi.org/10.1093/aje/kwad181
47. Kim CE, Binder AM, Corvalan C, et al. Time-specific impact of mono-benzyl phthalate (MBzP) and perfluorooctanoic acid (PFOA) on breast density of a Chilean adolescent cohort. *Environ Int*. 2023;181:108241. https://doi.org/10.1016/j.envint.2023.108241
48. Urner M, Jüni P, Rojas-Saunero LP, et al. Limiting dynamic driving pressure in patients requiring mechanical ventilation. *Crit Care Med*. 2023;51(7):861-871. https://doi.org/10.1097/CCM.0000000000005844
49. Guseva Canu I, Gaillen-Guedy A, Antilla A, et al. Lung cancer mortality in the European cohort of titanium dioxide workers: a reanalysis of the exposure-response relationship. *Occup Environ Med*. 2022;79(9):637-640. https://doi.org/10.1136/oemed-2021-108030

50. Jain P, Suemoto CK, Rexrode K, et al. Hypothetical lifestyle strategies in middle-aged women and the long-term risk of stroke. *Stroke*. 2020;51(5):1381-1387. https://doi.org/10.1161/STROKEAHA.119.026761
51. Mathisen J, Nguyen TL, Jensen JH, et al. Impact of hypothetical improvements in the psychosocial work environment on sickness absence rates: a simulation study. *Eur J Public Health*. 2022;32(5):716-722. https://doi.org/10.1093/eurpub/ckac109
52. O'Donnell A, Pham N, Battisti L, et al. Estimating the causal effect of treatment with direct-acting antivirals on kidney function among individuals with hepatitis C virus infection. *PLoS One*. 2022;17(5):e0268478. https://doi.org/10.1371/journal.pone.0268478
53. Eisenberg-Guyot J, Mooney SJ, Barrington WE, et al. Union burying ground: mortality, mortality inequities, and sinking labor union membership in the United States. *Epidemiology*. 2021;32(5):721-730. https://doi.org/10.1097/EDE.0000000000001386
54. Williamson EJ, Polak J, Simpson JA, et al. Sustained adherence to a Mediterranean diet and physical activity on all-cause mortality in the Melbourne collaborative cohort study: application of the g-formula. *BMC Public Health*. 2019;19(1):1733. https://doi.org/10.1186/s12889-019-7919-2
55. Zhang N, Chen F, Wang C, et al. Incidence of cognitive impairment after hypothetical interventions on depression, nighttime sleep duration, and leisure activity engagement among older Chinese adults: an application of the parametric g-formula. *Front Public Health*. 2023;11:1088833. https://doi.org/10.3389/fpubh.2023.1088833
56. Rojas-Saunero LP, Hilal S, Murray EJ, et al. Hypothetical blood-pressure-lowering interventions and risk of stroke and dementia. *Eur J Epidemiol*. 2021;36(1):69-79. https://doi.org/10.1007/s10654-020-00694-5
57. Bornhorst C, Reinders T, Rathmann W, et al. Avoiding time-related biases: a feasibility study on antidiabetic drugs and pancreatic cancer applying the parametric g-formula to a large German healthcare database. *Clin Epidemiol*. 2021;13:1027-1038. https://doi.org/10.2147/CLEP.S328342
58. Karaboyas A, Morgenstern H, Fleischer NL, et al. Replicating randomized trial results with observational data using the parametric g-formula: an application to intravenous iron treatment in hemodialysis patients. *Clin Epidemiol*. 2020;12:1249-1260. https://doi.org/10.2147/CLEP.S283321
59. Li Y, Wang Q, Liang S, et al. Effect of switching antiretroviral treatment regimen in patients with drug-resistant HIV-1 infection: retrospective observational cohort study. *JMIR Public Health Surveill*. 2022;8(6):e33429. https://doi.org/10.2196/33429
60. Mathisen J, Nguyen TL, Jense JH, et al. Reducing employee turnover in hospitals: estimating the effects of hypothetical improvements in the psychosocial work environment. *Scand J Work Environ Health*. 2021;47(6):456-465. https://doi.org/10.5271/sjweh.3969
61. Danaei G, Robins JM, Young JG, et al. Weight loss and coronary heart disease: sensitivity analysis for unmeasured confounding by undiagnosed disease. *Epidemiology*. 2016;27(2):302-310. https://doi.org/10.1097/EDE.0000000000000428
62. Yang YI, Hodge AM, Dugue PA, et al. Mortality effects of hypothetical interventions on physical activity and TV viewing. *Med Sci Sports Exerc*. 2021;53(2):316-323. https://doi.org/10.1249/MSS.0000000000002479
63. Matilla-Santander N, Matthews AA, Gunn V, et al. Causal effect of shifting from precarious to standard employment on all-cause mortality in Sweden: an emulation of a target trial. *J Epidemiol Community Health*. 2023;77(11):736-743. https://doi.org/10.1136/jech-2023-220734
64. Pitkanen J, Bijlsma MJ, Remes H, et al. The effect of low childhood income on self-harm in young adulthood: mediation by adolescent mental health, behavioural factors and school performance. *SSM Popul Health*. 2021;13:100756. https://doi.org/10.1016/j.ssmph.2021.100756
65. Zhou Z, Cai L, Fu J, et al. Hypothetical interventions on risk factors for cognitive impairment among Chinese older adults: an application of the parametric g-formula. *Int J Environ Res Public Health*. 2020;17(3):1021. https://doi.org/10.3390/ijerph17031021
66. Lodi S, Phillips A, Logan R, et al. Comparative effectiveness of immediate antiretroviral therapy versus CD4-based initiation in HIV-positive individuals in high-income countries: observational cohort study. *Lancet HIV*. 2015;2(8):e335-e343. https://doi.org/10.1016/S2352-3018(15)00108-3
67. Wilsgaard T, Vangen-Lønne AM, Mathiesen E, et al. Hypothetical interventions and risk of myocardial infarction in a general population: application of the parametric g-formula in a longitudinal cohort study-the Tromso study. *BMJ Open*. 2020;10(5):e035584. https://doi.org/10.1136/bmjopen-2019-035584
68. Mohammadi N, Alimohammadian M, Feizesani A, et al. The marginal causal effect of opium consumption on the upper gastrointestinal cancer death using parametric g-formula: an analysis of 49,946 cases in the Golestan cohort study, Iran. *PLoS One*. 2021;16(1):e0246004. https://doi.org/10.1371/journal.pone.0246004
69. Hale JM, Bijlsma MJ, Lorenti A. Does postponing retirement affect cognitive function? A counterfactual experiment to disentangle life course risk factors. *SSM Popul. Health*. 2021;15:100855. https://doi.org/10.1016/j.ssmph.2021.100855
70. van den Heuvel L, Evers LJW, Meinders MJ, et al. Estimating the effect of early treatment initiation in Parkinson's disease using observational data. *Mov Disord*. 2021;36(2):407-414. https://doi.org/10.1002/mds.28339
71. Inoue K, Nianogo R, Telesca D, et al. Low HbA1c levels and all-cause or cardiovascular mortality among people without diabetes: the US National Health and nutrition examination survey 1999-2015. *Int J Epidemiol*. 2021;50(4):1373-1383. https://doi.org/10.1093/ije/dyaa263
72. Chen C, Chen H, van Donkelaar A, et al. Using parametric g-computation to estimate the effect of long-term exposure to air pollution on mortality risk and simulate the benefits of hypothetical policies: the Canadian community health survey cohort (2005 to 2015). *Environ Health Perspect*. 2023;131(3):37010. https://doi.org/10.1289/EHP11095
73. Keil AP, Richardson DB. Reassessing the link between airborne arsenic exposure among Anaconda copper smelter workers and multiple causes of death using the parametric g-formula. *Environ Health Perspect*. 2017;125(4):608-614. https://doi.org/10.1289/EHP438
74. Lesko CR, Todd JV, Cole SR, et al. Mortality under plausible interventions on antiretroviral treatment and depression in HIV-infected women: an application of the parametric g-formula. *Ann Epidemiol*. 2017;27(12):783-789.e2. https://doi.org/10.1016/j.annepidem.2017.08.021
75. Eisenberg-Guyot J, Mooney SJ, Barrington WE, et al. Does the union make us strong? Labor-union membership, self-rated health, and mental illness: a parametric g-formula approach. *Am J Epidemiol*. 2021;190(4):630-641. https://doi.org/10.1093/aje/kwaa221
76. Bijlsma MJ, Tarkiainen L, Myrskylä M, et al. Unemployment and subsequent depression: a mediation analysis using the

parametric g-formula. *Soc Sci Med*. 2017;194:142-150. https://doi.org/10.1016/j.socscimed.2017.10.011
77. Westreich D, Cole SR, Young JG, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. *Stat Med*. 2012;31(18):2000-2009. https://doi.org/10.1002/sim.5316
78. Nascimento GG, Peres MA, Mittinty MN, et al. Diet-induced overweight and obesity and periodontitis risk: an application of the parametric g-formula in the 1982 Pelotas birth cohort. *Am J Epidemiol*. 2017;185(6):442-451. https://doi.org/10.1093/aje/kww187
79. Agodi A, Barchitta M, Quattrocchi A, et al. Preventable proportion of intubation-associated pneumonia: role of adherence to a care bundle. *PLoS One*. 2017;12(9):e0181170. https://doi.org/10.1371/journal.pone.0181170
80. Cole SR, Richardson DB, Chu H, et al. Analysis of occupational asbestos exposure and lung cancer mortality using the g formula. *Am J Epidemiol*. 2013;177(9):989-996. https://doi.org/10.1093/aje/kws343
81. Neophytou AM, Picciotto S, Costello S, et al. Occupational diesel exposure, duration of employment, and lung cancer: an application of the parametric g-formula. *Epidemiology*. 2016;27(1):21-28. https://doi.org/10.1097/EDE.0000000000000389
82. Jain P, Danaei G, Robins JM, et al. Smoking cessation and long-term weight gain in the Framingham heart study: an application of the parametric g-formula for a continuous outcome. *Eur J Epidemiol*. 2016;31(12):1223-1229. https://doi.org/10.1007/s10654-016-0200-4
83. Lodi S, Costagliola D, Sabin C, et al. Effect of immediate initiation of antiretroviral treatment in HIV-positive individuals aged 50 years or older. *J Acquir Immune Defic Syndr*. 2017;76(3):311-318. https://doi.org/10.1097/QAI.0000000000001498
84. Johnson KW, Glicksberg BS, Hodos RA, et al. Causal inference on electronic health records to assess blood pressure treatment targets: an application of the parametric g formula. *Pac Symp Biocomput*. 2018;23:180-191. https://doi.org/10.1142/9789813235533_0017
85. Sadinski LM, Westreich D, Edmonds A, et al. Hypertension and one-year risk of all-cause mortality among women with treated HIV in the United States. *AIDS*. 2023;37(4):679-688. https://doi.org/10.1097/QAD.0000000000003461
86. Neophytou AM, Costello S, Picciotto S, et al. Accelerated lung function decline in an aluminium manufacturing industry cohort exposed to PM(2.5): an application of the parametric g-formula. *Occup Environ Med*. 2019;76(12):888-894. https://doi.org/10.1136/oemed-2019-105908
87. Schleimer JP, Pear VA, McCort CD, et al. Unemployment and crime in US cities during the coronavirus pandemic. *J Urban Health*. 2022;99(1):82-91. https://doi.org/10.1007/s11524-021-00605-3
88. Dickerman BA, Giovannucci E, Pernar CH, et al. Guideline-based physical activity and survival among US men with non-metastatic prostate cancer. *Am J Epidemiol*. 2019;188(3):579-586. https://doi.org/10.1093/aje/kwy261
89. Neophytou AM, Picciotto S, Brown DM, et al. Estimating counterfactual risk under hypothetical interventions in the presence of competing events: crystalline silica exposure and mortality from 2 causes of death. *Am J Epidemiol*. 2018;187(9):1942-1950. https://doi.org/10.1093/aje/kwy077
90. Breskin A, Westreich D, Hurt CB, et al. The effects of hepatitis C treatment eligibility criteria on all-cause mortality among people with human immunodeficiency virus. *Clin Infect Dis*. 2019;69(9):1613-1620. https://doi.org/10.1093/cid/ciz008
91. Ueda P, Gulayin P, Danaei G. Long-term moderately elevated LDL-cholesterol and blood pressure and risk of coronary heart disease. *PLoS One*. 2018;13(7):e0200017. https://doi.org/10.1371/journal.pone.0200017
92. Shahn Z, Li Y, Sun Z, et al. G-computation and hierarchical models for estimating multiple causal effects from observational disease registries with irregular visits. *AMIA Jt Summits Transl Sci Proc*. 2019;2019:789-798. https://www.ncbi.nlm.nih.gov/pubmed/31259036
93. Lodi S, Günthard HF, Dunn D, et al. Effect of immediate initiation of antiretroviral treatment on the risk of acquired HIV drug resistance. *AIDS*. 2018;32(3):327-335. https://doi.org/10.1097/QAD.0000000000001692
94. Keil AP, Richardson DB, Westreich D, et al. Estimating the impact of changes to occupational standards for silica exposure on lung cancer mortality. *Epidemiology*. 2018;29(5):658-665. https://doi.org/10.1097/EDE.0000000000000867
95. Breskin A, Westreich D, Cole SR, et al. The effects of hepatitis C infection and treatment on all-cause mortality among people living with human immunodeficiency virus. *Clin Infect Dis*. 2019;68(7):1152-1159. https://doi.org/10.1093/cid/ciy588
96. Garcia E, Picciotto S, Neophytou AM, et al. Lung cancer mortality and exposure to synthetic metalworking fluid and biocides: controlling for the healthy worker survivor effect. *Occup Environ Med*. 2018;75(10):730-735. https://doi.org/10.1136/oemed-2017-104812
97. Sudharsanan N. Population-level mortality benefits of improved blood pressure control in Indonesia: a modelling study. *Int J Epidemiol*. 2019;48(3):954-965. https://doi.org/10.1093/ije/dyy232
98. Omorou AY, Achit H, Wieczorek M, et al. Impact of comorbidities and functional impairment on 5-year loss of health utility in patients with lower-limb osteoarthritis in the KHOALA cohort. *Qual Life Res*. 2019;28(11):3047-3054. https://doi.org/10.1007/s11136-019-02243-4
99. Ferguson JM, Costello S, Elser H, et al. Chronic obstructive pulmonary disease mortality: the Diesel Exhaust in Miners Study (DEMS). *Environ Res*. 2020;180:108876. https://doi.org/10.1016/j.envres.2019.108876
100. Bijlsma MJ, Wilson B, Tarkiainen L, et al. The impact of unemployment on antidepressant purchasing: adjusting for unobserved time-constant confounding in the g-formula. *Epidemiology*. 2019;30(3):388-395. https://doi.org/10.1097/EDE.0000000000000985
101. Costello S, Chen K, Picciotto S, et al. Metalworking fluids and cancer mortality in a US autoworker cohort (1941-2015). *Scand J Work Environ Health*. 2020;46(5):525-532. https://doi.org/10.5271/sjweh.3898
102. Eisen EA, Tolbert PE, Monson RR, et al. Mortality studies of machining fluid exposure in the automobile industry I: a standardized mortality ratio analysis. *Am J Ind Med*. 1992;22(6):809-824. https://doi.org/10.1002/ajim.4700220604
103. Eisen EA, Bardin J, Gore R, et al. Exposure-response models based on extended follow-up of a cohort mortality study in the automobile industry. *Scand J Work Environ Health*. 2001;27(4):240-249. https://doi.org/10.5271/sjweh.611
104. Tolbert PE, Eisen EA, Pothier LJ, et al. Mortality studies of machining-fluid exposure in the automobile industry. II. Risks associated with specific fluid types. *Scand J Work Environ Health*. 1992;18(6):351-360. https://doi.org/10.5271/sjweh.1562
105. Hallock MF, Smith TJ, Woskie SR, et al. Estimation of historical exposures to machining fluids in the automotive industry.

Am J Ind Med. 1994;26(5):621-634. https://doi.org/10.1002/ajim.4700260505
106. Woskie SR, Smith TJ, Hallock MF, et al. Size-selective pulmonary dose indices for metal-working fluid aerosols in machining and grinding operations in the automobile manufacturing industry. *Am Ind Hyg Assoc J*. 1994;55(1):20-29. https://doi.org/10.1080/15428119491019221
107. Woskie SR, Virji MA, Hallock M, et al. Summary of the findings from the exposure assessments for metalworking fluid mortality and morbidity studies. *Appl Occup Environ Hyg*. 2003;18(11):855-864. https://doi.org/10.1080/10473220390237377
108. Logan RW, Young JG, Taubman S, et al. *GFORMULA MACRO*. 2015; Available from: https://causalab.sph.harvard.edu/
109. Picciotto S, Peters A, Eisen EA. Hypothetical exposure limits for oil-based metalworking fluids and cardiovascular mortality in a cohort of autoworkers: structural accelerated failure time models in a public health framework. *Am J Epidemiol*. 2015;181(8):563-570. https://doi.org/10.1093/aje/kwu484
110. Picciotto S, Ljungman PL, Eisen EA. Straight metalworking fluids and all-cause and cardiovascular mortality analyzed by using g-estimation of an accelerated failure time model with quantitative exposure: methods and interpretations. *Am J Epidemiol*. 2016;183(7):680-688. https://doi.org/10.1093/aje/kwv232
111. Elser H, Chen KT, Arteaga D, et al. Metalworking fluid exposure and stroke mortality among US autoworkers. *Am J Epidemiol*. 2022;191(6):1040-1049. https://doi.org/10.1093/aje/kwac002
112. Costello S, Picciotto S, Rehkopf DH, et al. Social disparities in heart disease risk and survivor bias among autoworkers: an examination based on survival models and g-estimation. *Occup Environ Med*. 2015;72(2):138-144. https://doi.org/10.1136/oemed-2014-102168
113. Mircioiu C, Atkinson J. A comparison of parametric and non-parametric methods applied to a Likert scale. *Pharmacy (Basel)*. 2017;5(2):26. https://doi.org/10.3390/pharmacy5020026
114. Altman DG, Bland JM. Parametric v non-parametric methods for data analysis. *BMJ*. 2009;338(1):a3167. https://doi.org/10.1136/bmj.a3167
115. Conzuelo Rodriguez G, Bodnar LM, Brooks MM, et al. Performance evaluation of parametric and nonparametric methods when assessing effect measure modification. *Am J Epidemiol*. 2022;191(1):198-207. https://doi.org/10.1093/aje/kwab220
116. Yu Y-H, Bodnar LM, Brooks MM, et al. Comparison of parametric and nonparametric estimators for the association between incident prepregnancy obesity and stillbirth in a population-based cohort study. *Am J Epidemiol*. 2019;188(7):1328-1336. https://doi.org/10.1093/aje/kwz081
117. Wen L, Hernán MA, Robins JM. Multiply robust estiamtors of causal effects for survival outcomes. *Scand Stat Theory Appl*. 2022;49(3):1304-1328. https://doi.org/10.1111/sjos.12561
118. Wen L, Young JG, Robins JM, et al. Parametric g-formula implementations for causal survival analyses. *Biometrics*. 2021;77(2):740-753. https://doi.org/10.1111/biom.13321

American Journal of Epidemiology, 2025, **194**, 2249–2260
https://doi.org/10.1093/aje/kwae410
Original Contribution