Brief Report

Effects of Hypothetical Interventions on Ischemic Stroke Using Parametric G-Formula

Yaser Mokhayeri, PhD; Seyed Saeed Hashemi-Nazari, PhD; Soheila Khodakarim, PhD; Saeid Safiri, PhD; Nasrin Mansournia, MD; Mohammad Ali Mansournia, MD, MPH, PhD; Jay S. Kaufman, PhD; Ashley I. Naimi, PhD

Background and Purpose—Standard analytic approaches (eg, logistic regression) fail to adequately control for time-dependent confounding and, therefore, may yield biased estimates of the total effect of the exposure on the outcome. In the present study, we estimate the effect of body mass index, intentional physical activity, HDL (high-density lipoprotein) cholesterol, LDL (low-density lipoprotein) cholesterol, hypertension, and cigarette smoking on the 11-year risk of ischemic stroke by sex using the parametric g-formula to control time-dependent confounders.

Methods—Using data from the MESA (Multi-Ethnic Study of Atherosclerosis), we followed 6809 men and women aged 45 to 84 years. We estimated the risk of stroke under 6 hypothetical interventions: maintaining body mass index <25 kg/m^2, maintaining normotension (systolic blood pressure <140 and diastolic <90 mm Hg), quitting smoking, maintaining HDL >1.55 mmol/L, maintaining LDL <3.11 mmol/L, and exercising at least 210 minutes per week. The effects of joint hypothetical interventions were also simulated.

Results—In men, the 11-year risk of ischemic stroke would be reduced by 85% (95% CI, 66–96) for all 6 hypothetical interventions. In women, this same effect was estimated as 55% (95% CI, 6–82).

Conclusions—The hypothetical interventions explored in our study resulted in risk reduction in both men and women. **(*Stroke*. 2019;50:00-00. DOI: 10.1161/STROKEAHA.119.025749.)**

Key Words: atherosclerosis ■ blood pressure ■ body mass index ■ cigarette smoking ■ hypertension

Numerous observational studies have investigated the effects of healthy lifestyle and other risk factors on stroke incidence.[1,2] However, none of these studies appropriately adjusted for time-varying confounders affected by prior exposure.[3–5] In fact, standard analytic approaches (eg, logistic regression) fail to adequately control for such confounders and, therefore, may yield biased estimates of the total effect of the exposure on the outcome. Depending on the confounder, the induced bias could over or underestimate the impact of the interventions. To overcome this problem, using the parametric g-formula to appropriately adjust time-varying confounders, we estimate the 11-year cumulative risk of ischemic stroke by sex under various hypothetical interventions.[6] The parametric g-formula enables researchers to quantify the effects of generalized interventions, that is, joint, time-varying, dynamic, and stochastic interventions.[6]

Methods

Because of the sensitive nature of the data collected for this study, requests to access the data set may be sent to National Heart, Lung, and Blood Institute. We used data from the MESA (Multi-Ethnic Study of Atherosclerosis). The participants were 6809 men and women followed for more than 11 years and 5 visits. An outcome was defined by rapid onset of a documented focal neurological deficit lasting 24 hours or until death, or if <24 hours, if there was a clinically relevant lesion on brain imaging[7] (Appendix I in the online-only Data Supplement). Written informed consent was provided by all participants. For this study, we do not need any ethical approval because the data was acquired from National Heart, Lung, and Blood Institute—Research Materials Distribution Agreement V02 1d201208.

We assessed 6 hypothetical treatments and their combinations as joint hypothetical interventions on risk of ischemic stroke by sex based on previous publication[8]: maintain body mass index <25 kg/m^2, maintain LDL (low-density lipoprotein) <3.11 mmol/L, maintain HDL (high-density lipoprotein) >1.55 mmol/L, exercise intentionally at least 210 min a week, avoid smoking cigarettes and maintain normotension (systolic blood pressure <140 and diastolic <90 mm Hg). Additionally, another individual hypothetical intervention on body mass index was specified; all obese or overweight persons (body mass index >25 kg/m^2) lose 10% of weight per visit.

Time-varying covariates measured at all visits including triglyceride, diagnosed diabetes mellitus, diagnosed coronary

Received January 17, 2019; final revision received July 16, 2019; accepted July 18, 2019.

From the Department of Epidemiology and Biostatistics, School of Public Health and Nutrition, Lorestan University of Medical Sciences, Khorramabad, Iran (Y.M.); Department of Epidemiology, Safety Promotion and Injury Prevention Research Center (S.S.H.-N.) and Department of Epidemiology, School of Public Health (S.K.), Shahid Beheshti University of Medical Sciences, Tehran, Iran; Aging Research Institute, Tabriz University of Medical Sciences, Iran (S.S.); Department of Community Medicine, School of Medicine, Tabriz University of Medical Sciences, Iran (S.S.); Department of Endocrinology, AJA University of Medical Sciences, Tehran, Iran (N.M.); Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran (M.A.M.); Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University, Montreal, CA (J.S.K.); and Department of Epidemiology, University of Pittsburgh PA (A.I.N.).

The online-only Data Supplement is available with this article at https://www.ahajournals.org/doi/suppl/10.1161/STROKEAHA.119.025749.

Correspondence to Mohammad Ali Mansournia, MD, MPH, PhD, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Pour-Sina St, Tehran, Iran. Email mansournia_ma@yahoo.com

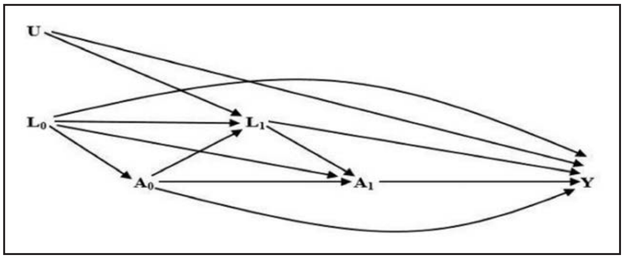

Figure. Causal diagram for the effect of the exposures, for example, body mass index (*A*) on the ischemic stroke (*Y*) or death from ischemic stroke in the MESA (Multi-Ethnic Study of Atherosclerosis). L and U stand for time-varying confounders (such as physical activity) and unmeasured confounders, respectively.

heart disease, hypertension medication, any lipid-lowering medication, and taking aspirin entered in the models as potential confounders. In addition, we included the following potential confounders that were measured only at the baseline: age, race, reported alcohol consumption, anger index, anxiety index, stroke history of parents, stroke history of sibling, homocysteine, fibrinogen, and CRP (C-reactive protein). The Figure is a causal directed acyclic graph to depict the relationship between time-varying exposures and stroke.[9] To simplify this graph; we depict just 2 visits, defined by subscripts of 0 and 1 which correspond to visits 0 and 1 of the study, respectively. *A* stands for time-varying exposures (eg, body mass index); *L* denotes time-varying confounders (eg, diagnosed diabetes mellitus) which are affected by the prior exposures; *Y* corresponds to the outcome (ischemic stroke); and *U* stands for unmeasured confounders. Details on the g-formula and notation are provided in Appendix II in the online-only Data Supplement.

Statistical Analysis

The parametric g-formula[7,10] was applied to estimate 11-year risk of stroke under various hypothetical interventions. The method involves modeling the outcome, each time-varying confounder, and the time-varying exposure at each MESA visit as a function of all variables that occur before the given visit. In this study, we estimated risk difference and population attributable fraction[11] (Appendix III in the online-only Data Supplement).

Results

Baseline characteristics of the 6809 eligible participants are tabulated in Appendix IV in the online-only Data Supplement.

The Table shows results from our g-formula analysis. After hypothetical intervention on all 6 factors, risk of ischemic stroke was lowered by 85% (95% CI, 66–96) in men and by 55% (95% CI, 6–82) in women. Losing weight by 10% during each visit for all obese or overweight persons results in 1% (95% CI, −4 to 5) reduction for men and −1% (95% CI, −5 to 6) reduction for women in the risk of ischemic stroke.

Table. Ischemic Stroke Risk Under Hypothetical Interventions in Men and Women in the Multi-Ethnic Study of Atherosclerosis, United States, 2000–2011

	11-Year Risk, %		Risk Difference, %		PAF, %	
Intervention	Men	Women	Men	Women	Men	Women
(0) No intervention	3.13	2.15	0	0	0	0
(1) Maintain BMI <25 kg/m²	2.75	1.81	−0.38 (−1.41 to 0.42)	−0.34 (−0.96 to 0.23)	12.11 (−13.39 to 47.09)	15.75 (−10.92 to 44.63)
(2) No hypertension*	1.36	1.50	−1.77 (−3.00 to −0.56)	−0.65 (−1.64 to 0.36)	56.65 (21.96 to 80.62)	30.31 (−17.86 to 66.96)
(3) No smoking	2.65	2.40	−0.48 (−1.57 to 0.73)	0.24 (−0.28 to 0.76)	15.43 (−24.88 to 47.12)	−11.33 (−33.79 to 11.85)
(4) Exercise at least 210 min/wk	2.74	2.10	−0.39 (−1.07 to 0.06)	−0.06 (−0.63 to 0.38)	12.56 (−1.85 to 30.59)	2.60 (−18.74 to 26.10)
(5) Desirable HDL†	1.68	1.52	−1.45 (−2.84 to −0.3)	−0.63 (−1.17 to 0.10)	46.32 (11.09 to 73.38)	29.28 (−3.82 to 49.51)
(6) Desirable LDL‡	3.01	2.07	−0.12 (−0.51 to 0.12)	−0.09 (−0.44 to 0.34)	3.91 (−4.37 to 14.79)	3.98 (−16.25 to 19.01)
(7) Joint interventions (2, 5, 6)	0.68	1.03	−2.45 (−4.00 to −1.21)	−1.12 (−2.03 to 0.03)	78.20 (51.87 to 93.00)	52.28 (−1.88 to 82.23)
(8) Joint interventions (1, 2, 5, 6)	0.61	0.91	−2.53 (−4.12 to −1.4)	−1.24 (−2.13 to −0.31)	80.64 (58.44 to 93.74)	57.71 (16.07 to 81.59)
(9) Joint interventions (3, 4)	2.36	2.35	−0.77 (−1.79 to 0.34)	0.20 (−0.47 to 0.85)	24.50 (−10.72 to 51.38)	−9.35 (−37.84 to 19.26)
(10) Joint interventions (1, 4)	2.47	1.77	−0.66 (−1.77 to 0.15)	−0.38 (−1.13 to 0.36)	21.05 (−4.63 to 56.06)	17.53 (−19.50 to 51.60)
(11) Joint interventions (1,3,4)	2.10	2.02	−1.03 (−2.31 to 0.40)	−0.13 (−1.01 to 0.69)	32.79 (−13.01 to 65.65)	6.07 (−34.29 to 47.95)
(12) Joint interventions (1–4)	0.94	1.39	−2.19 (−3.69 to −1.13)	−0.77 (−1.83 to 0.40)	69.92 (43.59 to 89.43)	35.60 (−17.78 to 72.86)
(13) All interventions (1–6)	0.46	0.98	−2.67 (−4.33 to −1.56)	−1.17 (−2.08 to −0.13)	85.26 (65.80 to 95.55)	54.60 (6.13 to 82.35)
(14) Lose 10% of BMI if BMI ≥25 kg/m²	3.09	2.17	−0.04 (−0.22 to 0.15)	0.02 (−0.14 to 0.11)	1.42 (−4.13 to 5.53)	−0.95 (−5.16 to 5.99)
(15) Joint interventions (4, 14)	2.74	2.08	−0.39 (−1.07 to 0.03)	−0.07 (−0.63 to 0.32)	12.47 (−1.23 to 32.54)	3.23 (−15.37 to 27.16)
(16) Joint interventions (3, 4, 14)	2.32	2.29	−0.81 (−1.79 to 0.37)	0.14 (−0.47 to 0.81)	26.01 (−13.59 to 52.35)	−6.32 (−39.87 to 20.53)
(17) Joint interventions (2, 3, 4, 14)	1.02	1.60	−2.11 (−3.36 to −1.01)	−0.55 (−1.71 to 0.68)	67.33 (36.15 to 86.24)	25.50 (−35.91 to 68.85)
(18) Joint interventions (2, 5, 6, 14)	0.68	1.01	−2.45 (−3.97 to −1.24)	−1.14 (−2.02 to −0.01)	78.26 (52.58 to 92.92)	53.00 (0.72 to 82.10)
(19) All interventions (2–6 and 14)	0.51	1.09	−2.62 (−4.26 to −1.37)	−1.06 (−2.03 to 0.44)	83.74 (55.80 to 94.52)	49.39 (−21.02 to 81.98)

BMI indicates body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein; and PAF, population attributable fraction.
*Systolic blood pressure lower than 140 and diastolic lower than 90 mm Hg.
†Min 1.55 mmol/L.
‡Max 3.11 mmol/L.

Discussion

Hypothetical interventions predicted the absolute risk reduction as 85% in men and 55% in women for ischemic stroke, however, due to the imprecision in the estimate for women, there is only weak evidence for heterogeneity of these reductions (interaction *P* value =0.14). In our analysis, the parametric g-formula was used to account for the time-varying nature of our exposures on incidence of stroke. Additionally, we improve on previous work by conducting subgroup analyses by sex.[12] Different population attributable fraction's between men and women could be ascribed to some extent to higher prevalence of risk factors in men.[13]

We acknowledge the possibility of sparse data bias and low precision of effect estimates (represented by wide confidence intervals).[14] We did not adjust for dietary pattern because for 577 participants there was no measurement of the dietary pattern, which challenges the assumption of no unmeasured confounders. Moreover, physical activity was not assessed in visit 4 of the study. Therefore, we carried forward the values visit 3 to the missed visit. It should be noted that the proportion of missing data was less than 4% for all covariates. G-formula assumptions, limitations, and the graphs as a description of the model fit may be found in Appendix V in the online-only Data Supplement.

Conclusions

Our analyses indicated that in men 85% of cases could have been prevented by compliance with 6 hypothetical interventions. In women, only 54% of cases could have been similarly prevented.

Acknowledgments

This article was prepared using MESA Research Materials obtained from the National Heart, Lung, and Blood Institute; Research Materials Distribution Agreement (RMDA) V02 1d20120806.

Disclosures

None.

References

1. Tikk K, Sookthai D, Monni S, Gross ML, Lichy C, Kloss M, et al. Primary preventive potential for stroke by avoidance of major lifestyle risk factors: the European prospective investigation into cancer and nutrition-heidelberg cohort. *Stroke*. 2014;45:2041–2046. doi: 10.1161/STROKEAHA.114.005025
2. Willey JZ, Moon YP, Paik MC, Boden-Albala B, Sacco RL, Elkind MS. Physical activity and risk of ischemic stroke in the Northern Manhattan study. *Neurology*. 2009;73:1774–1779. doi: 10.1212/WNL.0b013e3181c34b58
3. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ*. 2017;359:j4587. doi: 10.1136/bmj.j4587
4. Mansournia MA, Altman DG. Inverse probability weighting. *BMJ*. 2016;352:i189. doi: 10.1136/bmj.i189
5. Shakiba M, Mansournia MA, Salari A, Soori H, Mansournia N, Kaufman JS. Accounting for time-varying confounding in the relationship between obesity and coronary heart disease: analysis with G-estimation: the ARIC Study. *Am J Epidemiol*. 2018;187:1319–1326. doi: 10.1093/aje/kwx360
6. Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol*. 2009;38:1599–1611. doi: 10.1093/ije/dyp192
7. Aho K, Harmsen P, Hatano S, Marquardsen J, Smirnov VE, Strasser T. Cerebrovascular disease in the community: results of a WHO collaborative study. *Bull World Health Organ*. 1980;58:113–130.
8. Lichtenstein AH, Appel LJ, Brands M, Carnethon M, Daniels S, Franch HA, et al. Diet and lifestyle recommendations revision 2006 a scientific statement from the American heart Association nutrition committee. *Circulation*. 2006;114:82–96.
9. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol*. 2013;42:860–869. doi: 10.1093/ije/dyt083
10. Mansournia MA, Naimi AI, Greenland S. The implications of using lagged and baseline exposure terms in longitudinal causal and regression models. *Am J Epidemiol*. 2019;188:753–759. doi: 10.1093/aje/kwy273
11. Mansournia MA, Altman DG. Population attributable fraction. *BMJ*. 2018;360:k757. doi: 10.1136/bmj.k757
12. Vangen-Lønne AM, Ueda P, Gulayin P, Wilsgaard T, Mathiesen EB, Danaei G. Hypothetical interventions to prevent stroke: an application of the parametric g-formula to a healthy middle-aged population. *Eur J Epidemiol*. 2018;33:557–566. doi: 10.1007/s10654-017-0344-x
13. O'Donnell MJ, Chin SL, Rangarajan S, Xavier D, Liu L, Zhang H, et al; INTERSTROKE Investigators. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. *Lancet*. 2016;388:761–775. doi: 10.1016/S0140-6736(16)30506-2
14. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. *BMJ*. 2016;352:i1981. doi: 10.1136/bmj.i1981