Testing And Estimating Model-Adjusted Effect-Measure Modification Using Marginal Structural Models And Complex Survey Data

By: Babette A. Brumback, **Erin D. Bouldin**, Hao W. Zheng, Michael B. Cannell, and Elena M. Andresen

Abstract

Recently, it has been shown how to estimate model-adjusted risks, risk differences, and risk ratios from complex survey data based on risk averaging and SUDAAN (Research Triangle Institute, Research Triangle Park, North Carolina). The authors present an alternative approach based on marginal structural models (MSMs) and SAS (SAS Institute, Inc., Cary, North Carolina). The authors estimate the parameters of the MSM using inverse weights that are the product of 2 terms. The first term is a survey weight that adjusts the sample to represent the unstandardized population. The second term is an inverse-probability-of-exposure weight that standardizes the population in order to adjust for confounding; it must be estimated using the survey weights. The authors show how to use the MSM parameter estimates and contrasts to test and estimate effect-measure modification; SAS code is provided. They also explain how to program the previous risk-averaging approach in SAS. The 2 methods are applied and compared using data from the 2007 Florida Behavioral Risk Factor Surveillance System Survey to assess effect modification by age of the difference in risk of cost barriers to health care between persons with disability and persons without disability.

Babette A. Brumback, **Erin D. Bouldin**, Hao W. Zheng, Michael B. Cannell, Elena M. Andresen, Testing and Estimating Model-Adjusted Effect-Measure Modification Using Marginal Structural Models and Complex Survey Data, American Journal of Epidemiology, Volume 172, Issue 9, 1 November 2010, Pages 1085–1091, https://doi.org/10.1093/aje/kwq244. Publisher version of record available at: https://academic.oup.com/aje/article/172/9/1085/148053?

Testing and Estimating Model-Adjusted Effect-Measure Modification Using Marginal Structural Models and Complex Survey Data

Babette A. Brumback*, Erin D. Bouldin, Hao W. Zheng, Michael B. Cannell, and Elena M. Andresen

* Correspondence to Dr. Babette A. Brumback, Department of Epidemiology and Biostatistics, College of Public Health and Health Professions, University of Florida, P.O. Box 100231, Gainesville, FL 32610-0231 (e-mail: brumback@ufl.edu).

Recently, it has been shown how to estimate model-adjusted risks, risk differences, and risk ratios from complex survey data based on risk averaging and SUDAAN (Research Triangle Institute, Research Triangle Park, North Carolina). The authors present an alternative approach based on marginal structural models (MSMs) and SAS (SAS Institute, Inc., Cary, North Carolina). The authors estimate the parameters of the MSM using inverse weights that are the product of 2 terms. The first term is a survey weight that adjusts the sample to represent the unstandardized population. The second term is an inverse-probability-of-exposure weight that standardizes the population in order to adjust for confounding; it must be estimated using the survey weights. The authors show how to use the MSM parameter estimates and contrasts to test and estimate effect-measure modification; SAS code is provided. They also explain how to program the previous risk-averaging approach in SAS. The 2 methods are applied and compared using data from the 2007 Florida Behavioral Risk Factor Surveillance System Survey to assess effect modification by age of the difference in risk of cost barriers to health care between persons with disability and persons without disability.

health surveys; heterogeneity; interaction; logistic regression; models, statistical; probability weighting; standardization; survey analysis

Recently, our colleagues at the Florida Office on Disability and Health conducted research on persons with disability (PWD) using a population-based sample (1). They aimed to document that the difference in risk of cost barriers to health care between PWD and persons without disability was greater for younger persons than for older persons. They also wanted to adjust the risks for confounders, such as race/ethnicity, income, education, and gender. Their goal was to demonstrate for policy-makers that interventions targeting younger PWD are much needed and cost-effective. To argue for cost-effectiveness, it is helpful to use the risk difference. For example, even if the relative risk were constant across age groups (as it was in a preliminary analysis), a greater risk difference in younger persons might imply (if a causal interpretation of results were valid) that intervening to help a group of younger PWD could benefit more people overall than an intervention for an equal-sized group of older PWD.

Estimating and testing heterogeneity of the crude risk differences using complex survey data are relatively simple to do. For example, one could use SAS PROC SURVEYREG (SAS Institute, Inc., Cary, North Carolina), as we explain below. Adjusting the risk differences for confounding requires a less obvious approach. Model-based standardization (e.g., see chapter 21 in *Modern Epidemiology* (2)) represents a natural approach to this problem. As is described in the textbook (2), the analyst can choose between 1) using an exposure model followed by inverse probability weighting and 2) using an outcome model followed by risk averaging. However, the textbook does not explain how to implement the approaches with complex survey data. Sato and Matsuyama (3) connected the first approach to marginal structural models (4) in the context of a simple random sample. Recently, Bieler et al. (5) explained how to implement the second approach for complex survey data using

SUDAAN (Research Triangle Institute, Research Triangle Park, North Carolina); further details can be found in *Analysis of Health Surveys* (6). Our main purpose in the present article is to explain 1) how to extend the first approach for complex survey data and 2) how to program both approaches in SAS. We consider not only standardization of the risk difference but also standardization of the relative risk and odds ratio.

For both the exposure modeling and the outcome modeling, we use logistic regression, because the logit link preserves the scale of the binary outcome and of the binary exposure. For the outcome model, the outcome is regressed on the exposure, modifier, and confounders, while for the exposure model, the exposure is regressed on the modifier and confounders. For approaches to confounding adjustment based on other, nonlogistic regression models that do not involve complex survey data and do not necessarily involve standardization, see Spiegelman and Hertzmark (7), Zou (8), and Greenland (9), as well as the references therein. When using a population-based sample, the analyst generally needs to account for complex multistage sampling in the estimation of parameters and confidence intervals, as well as in hypothesis tests (see Brogan (10) for examples). We explain how to do this using SAS.

Below we explain the extension to complex survey data of the 2 approaches to model-based standardization of the risks on which the standardization of the risk difference, relative risk, and odds ratio are based. Then we present a simulation study to demonstrate the validity of the approaches. Subsequently, we present the analysis of our motivating example using 2007 Florida Behavioral Risk Factor Surveillance System (BRFSS) (11, 12) survey data. We conclude with a brief discussion.

AN APPROACH BASED ON EXPOSURE MODELING

We first define our target of estimation, the standardized risk, with respect to the population of interest—for example, the state of Florida in 2007. Let Y_i, $i = 1, \ldots, N$ be a binary outcome, namely the presence or absence of a cost barrier to health care. Similarly, let X_i be a binary exposure of interest, here disability status; M_i be the effect modifier, age category; and Z_i be a p-dimensional vector of confounders, such as race/ethnicity, income, education, and gender. In words, our target is an average of the conditional risk of the outcome given exposure, modifier, and confounder, where the average is taken with respect to the population distribution of the confounder conditional on the modifier. In statistical notation, this can be written as $E_{Z|M=m}E(Y|X = x,M = m,Z)$, which is a function of x and m. The inner expectation represents the conditional risk, a function of the nonrandom values x and m as well as the random variable Z, whereas the outer expectation averages that function with respect to the conditional distribution of Z given $M = m$. Note the subtle distinction between our target and the crude estimate of risk $E(Y|X = x,M = m) = E_{Z|M=m,X=x}E(Y|X = x,M = m,Z)$, in which the outer expectation averages with respect to the conditional distribution of Z given both $M = m$ and $X = x$.

The approach based on exposure modeling uses the identity $E_{Z|M=m}E(Y|X = x,M = m,Z) = E^{W(X,M,Z)}(Y|X = x,M = m)$, where the latter expectation refers to a weighted population average of Y within the stratum defined by $X = x$ and $M = m$, using weights $W(X,M,Z) = P(X|M)/P(X|M,Z)$. Because one cannot divide by zero in constructing the weights, the exposure modeling approach can only be used when $P(X|M,Z)$ is bounded away from zero. Under causal modeling assumptions, the weighted population average is equivalent to a saturated marginal structural model $E(Y_x|M = m)$, where Y_x denotes the potential outcome to $X = x$. Specifically, one must assume that the collection of Y_x for all persons and values of x exist and are well-defined (e.g., that the potential outcome of one person does not depend on the exposure realized by other persons), and furthermore that Z is a sufficient set of confounders for estimating the parameters of the marginal structural model (i.e., the potential outcomes are independent of X given Z and M. (See Robins et al. (4), VanderWeele (13), and Chiba et al. (14) for more details.) For researchers who only want to standardize estimates rather than estimate causal effects, the link to marginal structural models is theoretically interesting but practically unimportant.

The major difficulty in estimating $E^{W(X,M,Z)}(Y|X = x,M = m)$ is due to the need to model $W(X,M,Z) = P(X|M)/P(X|M,Z)$, which in turn involves 2 models for the probability of exposure. The numerator model is conditional on the modifier only, whereas the denominator model is conditional on the modifier and the confounder. The weight $W(X,M,Z)$ is obviously directly related to the inverse-probability weight $1/P(X|M,Z)$; the latter weight is sufficient for estimation of the marginal structural model expectation $E(Y_x|M = m)$, because only data from persons who are homogeneous in X and M (and hence for whom $P(X|M)$ is constant) will be averaged.

For high-dimensional Z, as in our example with 4 categories for race/ethnicity (Z_1), 5 categories for income (Z_2), 4 categories for education (Z_3), and 2 categories for gender (Z_4), one generally needs to use an unsaturated model for $P(X|M,Z)$. We use a logistic regression model of the form $\text{logit}(P(X = 1|M,Z)) = M\alpha_0 + Z_1\alpha_1 + Z_2\alpha_2 + Z_3\alpha_3 + Z_4\alpha_4$, in which the covariates M and $Z_1, \ldots, Z_4$ are horizontal vectors and the parameters $\alpha_0, \ldots, \alpha_4$ are vertical vectors. It is important to note that this is a model for the population and not for the sample. Because of the complex survey design, the sample is subject to selection bias. Hence, to estimate $\alpha_0, \ldots, \alpha_4$, we must use the survey weights. We do this using SAS PROC SURVEYLOGISTIC. We can also compute $P(X|M)$ using an analogous regression, and then finally we compute the weight $W(X,M,Z)$. Note, however, that $1/P(X|M,Z)$ would suffice.

The weights $W(X,M,Z)$ effectively standardize the population to adjust for confounding by Z and M. If we had a simple random sample from the population, these weights would be sufficient to conduct model-based standardization. However, we have a complex multistage sample. Hence, after first weighting the sample to represent the population, we must then reweight the weighted sample to standardize our representation of the population. Let S_i denote the survey weight for individual i, and let W_i denote $W(X_i,M_i,Z_i)$ computed for individual i. Then the combined weight that adjusts for selection bias and confounding is simply $S_i \times W_i$.

We can now estimate our target $E_{Z|M=m}E(Y|X = x, M = m,Z)$ via a procedure that estimates proportions using survey data. For example, one could use SAS PROC SURVEYFREQ or SAS PROC SURVEYREG with the weight equal to $S_i \times W_i$ and using the primary stratification and primary clustering variables in the strata and cluster statement. We are also interested in risk differences, relative risks, odds ratios, and tests of homogeneity. For estimating and testing homogeneity of the risk differences, we can use PROC SURVEYREG with the 2007 Florida BRFSS data, as shown in Appendix 1.

Because PROC SURVEYREG uses Taylor linearization (i.e., the delta method) to estimate sampling variability, one can use it with binary outcomes to obtain correct asymptotic confidence intervals and tests of hypotheses about risk differences when the model statement is correctly specified (or saturated, as in our usage) and when the weights are known. When the weights are estimated, as in our example, the confidence intervals and hypothesis tests will be asymptotically conservative (e.g., 95% confidence intervals will have coverage slightly greater than 95%), as described by Robins et al. (4). In our example, we estimated the weights (''combinedweight'') as shown in Appendix 2.

The test of heterogeneity is standard output of SAS PROC SURVEYREG; it is the ''test of model effects'' for the age_3level*disability term, which has 2 df. Because the denominator degrees of freedom are so large for our example, due to the large sample size of the 2007 Florida BRFSS, this F test is effectively a chi-squared 2-df test.

For the odds ratio, an analogous analysis can be done using SAS PROC SURVEYLOGISTIC. For the relative risk, unfortunately, SAS does not yet have a survey procedure for log-linear models. One could use SAS PROC GENMOD instead, ignoring the survey stratification when estimating sampling variability; this is a conservative approach in that the variability will be overestimated.

Alternatively, one could use PROC SURVEYREG or PROC SURVEYLOGISTIC in combination with PROC IML to compute the relative risks and confidence intervals and to conduct a test of homogeneity based on a parametric bootstrap simulation, as in the article by Greenland (15). Briefly, one computes the parameter estimates and asymptotic covariance matrix using a saturated model, such as the one above in PROC SURVEYREG. One uses the parameter estimates to construct the log relative risks (for our example, call these θ_1, θ_2, and θ_3 for the 3 age groups). Then, one uses the asymptotic covariance matrix to simulate random error terms; these are added to the original parameter estimates to form simulated bootstrap parameter estimates. For each bootstrap parameter estimate, one computes the log relative risks. One then computes the bootstrap sample covariance matrix for a suitable set of contrasts of the log relative risks (for our example, we would have 2 contrasts: one for the middle age range versus the younger age range and one for the older age range versus the younger age range). Let the contrasts be c_1 and c_2 in our example (e.g., $\theta_2 - \theta_1$ and $\theta_3 - \theta_1$), and let the covariance matrix be V. Then an approximately chi-squared test statistic with degrees of freedom equal to the number of linearly independent contrasts (2 in our example) is given by, for example, $(c_1\ c_2)V^{-1}(c_1\ c_2)^T$. Confidence intervals for the log relative risks or differences of log relative risks are based on a normal approximation and the appropriate components of V. SAS code is available from the authors upon request.

A nonparametric bootstrap is also possible, resampling primary sampling units within primary strata. Using this approach, one would reestimate the weights for each bootstrap sample and thus could obtain confidence intervals and hypothesis tests that are correct when the weights are estimated rather than known. This approach is also possible in SAS, but the programming would be more difficult because one could not rely on the standard SAS macro for the nonparametric bootstrap (which does not have an option to resample within strata).

AN APPROACH BASED ON OUTCOME MODELING

Recall that the target of estimation is $E_{Z|M=m}E(Y|X = x,M = m,Z)$. Let D be a vector of dummy variables representing the interaction of X and M, and let Z_1, Z_2, Z_3, and Z_4 similarly represent the confounders race/ethnicity, income, education, and age. The approach based on outcome modeling uses a logistic regression for $E(Y|X,M,Z)$; our model is of the form $\text{logit}(P(Y = 1|X,M,Z)) = X\beta + M\gamma + D\lambda + Z_1\eta_1 + Z_2\eta_2 + Z_3\eta_3 + Z_4\eta_4$. We estimate the parameters of this model using PROC SURVEYLOGISTIC together with the survey weights and stratum and clustering variables. For individual i in the subset of the population with $M_i = m$, we plug these parameter estimates into the expression $R_i(x) = \text{logit}^{-1}(x\beta + m\gamma + d\lambda + Z_{i1}\eta_1 + Z_{i2}\eta_2 + Z_{i3}\eta_3 + Z_{i4}\eta_4)$, where d is the value of D for a person with $M = m$ and $X = x$. The target is then estimated as

$$\frac{\sum_{i \text{ such that } M_i=m} S_i R_i(x)}{\sum_{i \text{ such that } M_i=m} S_i}, \qquad (1)$$

where the survey weights S_i are used to ensure that the standard distribution $Z|M = m$ reflects the overall population rather than the selected population. This is the method programmed in SUDAAN by Bieler et al. (5) and explained in more detail by Korn and Graubard (6). Our target of estimation, contrasts of functions of this target, and confidence intervals and hypothesis tests can be programmed in SAS PROC IML using the same kind of approach as outlined above based on simulation as in the Greenland article (15). Briefly, the parameter estimates and their asymptotic covariance matrix are computed using PROC SURVEYLOGISTIC, as just described; then bootstrap replications are computed as above. For each bootstrap replicate, expression 1 and functions of expression 1 are computed. Confidence intervals and hypothesis tests can be programmed as outlined above. SAS code is available from the authors upon request.

SIMULATION STUDY

To illustrate the validity of the 2 approaches, we conducted a simulation study using SAS. In our simulation study, M is a binary effect modifier, with $P(M = 1) = 0.4$. We consider a single confounder Z distributed as $N(0,1)$.

Table 1. Results of a Simulation Study Comparing 2 Approaches (Exposure Modeling and Outcome Modeling) to Estimating the Standardized Population Risk Difference

Effect Modifier	True Risk Difference	Estimated Risk Difference		95% Confidence Interval Coverage Probability, %	
		Exposure Modeling	Outcome Modeling	Exposure Modeling	Outcome Modeling
$M = 1$	0.441	0.439	0.441	94.0	95.4
$M = 0$	0.232	0.230	0.231	95.0	95.6

The population exposure model is $\text{logit}(P(X = 1|M,Z)) = 0.5 - 1 \times M + 1 \times Z + 0.5 \times M \times Z$, and the population outcome model is $\text{logit}(P(Y = 1|X,M,Z)) = -0.5 + 1 \times X - 0.5 \times M + 1 \times X \times M + 0.5 \times Z$.

To obtain the true values of the parameters of interest, which are the population standardized risk differences at each level of M, we must average the population outcome model across the distribution of Z within M; we used Monte Carlo integration with a simulated population of size 100,000.

A complex survey sample was simulated as follows. A sample of size 3,500 was initially sampled from the population models for M, Z, X, and Y, respectively. If $Y = X$, then the observation was kept with probability 0.5 and was assigned a survey weight $S = 2$. Otherwise, the observation was kept with probability 1 and was assigned a survey weight $S = 1$. We simulated 500 such complex survey samples and applied the 2 approaches to each one. The average size of the complex survey samples was 2,267.

Table 1 presents the results. Both approaches accurately estimate the true risk difference within each level of the effect modifier. Additionally, the 95% confidence intervals have the correct coverage. We suspect that if our simulation were larger than 500, the approaches would be even more comparable; however, because of the bootstrap necessitated by the outcome modeling approach, our simulation was very computer-intensive, and thus we limited the size to 500.

EXAMPLE

Next we apply the 2 model-based standardization approaches to the motivating example, in which the primary target of estimation is the standardized risk of a cost barrier to health care as a function of disability status, within each of 3 age groups. Each standardization adjusts for confounding due to race/ethnicity, income, education, and gender using either the exposure modeling approach or the outcome modeling approach. We use data from the 2007 Florida BRFSS Survey (11, 12). Participants were assessed as having a cost barrier to health care if they reported that they could not visit a doctor in the past year due to cost. Disability was determined according to the definition used by the Centers for Disease Control and Prevention, where a person was classified as having a disability if he or she reported having a limitation or using special equipment (16). Race/ethnicity was categorized into 4 groups: non-Hispanic white; non-Hispanic black; non-Hispanic of another race or multiracial non-Hispanic; and Hispanic of any race. Annual household income was categorized into 5 groups: less than $20,000, $20,000–$24,999, $25,000–$34,999, $35,000–$49,000, and $50,000 or more. Education was categorized into 4 groups: less than high school, high school graduation or equivalent, some college, and college degree or higher. We excluded participants with missing data on any of these variables. Our final sample included 31,590 participants.

Table 2 presents estimates of the crude risks and risk differences within age categories, obtained using SAS PROC SURVEYREG and the original survey weights. The risk differences are clearly heterogeneous, with a P value less than 0.001.

Table 3 presents the results from the exposure modeling or marginal structural model approach, applied to estimate the standardized risks and risk differences and to test heterogeneity. The risk differences are heterogeneous ($P < 0.001$). We observe that younger adults have a higher risk difference than older adults, as our colleagues at the Florida Office on Disability and Health expected. The risk differences for the 2 younger age groups (18–29 years and 30–64 years) are estimated as similar; this stands in contrast to the crude risk differences, which are rather different.

Table 4 presents analogous results based on the outcome modeling and risk-averaging approach. Again, the risk differences are heterogeneous ($P < 0.001$). In this set of results, the younger adults again have higher risk differences than the older adults, but now the risk difference for the

Table 2. Testing and Estimating Effect-Measure Modification by Age of the Crude Risk Difference for the Effect of Persons With Disability on Cost Barriers to Health Care, Using Linear Regression for Complex Survey Data, Florida Behavioral Risk Factor Surveillance System Survey, 2007

Age Group, years	Risk for Persons With Disability[a]	95% CI	Risk for Persons Without Disability[a]	95% CI	Risk Difference[a,b]*	95% CI
18–29	0.453	0.332, 0.573	0.219	0.181, 0.256	0.234	0.108, 0.360
30–64	0.313	0.284, 0.343	0.133	0.121, 0.145	0.180	0.149, 0.212
≥65	0.071	0.053, 0.088	0.033	0.024, 0.042	0.038	0.018, 0.058

Abbreviation: CI, confidence interval.

*$P < 0.001$.

[a] Not adjusted for covariates.

[b] Test of effect-measure modification: χ^2 (2 df) = 30.87.

Table 3. Testing and Estimating Effect-Measure Modification by Age of the Adjusted Risk Difference for the Effect of Persons With Disability on Cost Barriers to Health Care, Using Marginal Structural Models for Complex Survey Data, Florida Behavioral Risk Factor Surveillance System Survey, 2007

Age Group, years	Risk for Persons With Disability[a]	95% CI	Risk for Persons Without Disability[a]	95% CI	Risk Difference[a,b,*]	95% CI
18–29	0.332	0.189, 0.475	0.225	0.183, 0.268	0.107	−0.043, 0.256
30–64	0.260	0.224, 0.296	0.142	0.129, 0.155	0.118	0.080, 0.156
≥65	0.063	0.045, 0.081	0.041	0.029, 0.054	0.021	−0.001, 0.044

Abbreviation: CI, confidence interval.
* $P < 0.001$.
[a] Adjusted for race/ethnicity, income, education, and gender using the exposure modeling approach.
[b] Test of effect-measure modification: χ^2 (2 df) = 9.55.

youngest age group is estimated as higher than that for the middle age group, as with the crude risk differences.

DISCUSSION

We have presented 2 approaches to model-based standardization of risks, risk differences, relative risks, and odds ratios using complex survey data and SAS programming. One approach models the exposure as a function of confounders and the effect modifier, and the other approach models the outcome as a function of confounders, the effect modifier, and the exposure. Motivated by our research on disability, we applied the 2 approaches to 2007 Florida BRFSS data in order to assess heterogeneity of the risk differences across age groups. Either analysis enables our colleagues to document that the risk difference is larger for the younger age group (<65 years) than for the older age group. The difference in estimates for the youngest age group (18–25 years) from the 2 approaches leads us naturally to wonder what results a doubly robust procedure (17) would give. Briefly, doubly robust procedures give design-consistent estimators if at least 1 of the exposure model or the outcome model is correct, whereas our exposure modeling approach is only correct if the exposure model is correct, and our outcome modeling approach is only correct if the outcome model is correct. We are currently working on developing a doubly robust procedure for complex survey data.

Model-based standardization provides an assessment of an unconditional exposure effect on the entire subpopulation within a level of the effect modifier. Ordinary regression approaches return a conditional exposure effect for a subset of persons not just with the same level of the effect modifier but also with identical levels of the confounders. For many applications, as in ours, the unconditional effect is of primary interest; the unconditional risk difference is particularly useful for cost-effectiveness analyses. For the special case of the risk difference, model-based standardization can be viewed as averaging the conditional exposure effect across the distribution of confounders within the modifier.

Our methods have some limitations. For expository purposes, we have excluded persons with missing data for any of the key variables so that we could better focus on the model-based standardization. If data on only a few variables were missing, the combined weights could accommodate a third term representing the inverse probability of a complete observation. Our methods lead to causal interpretations only under the strong assumptions we presented. The first method additionally requires that the exposure model is correct, and the second method requires that the outcome model is correct. The first method leads to confidence intervals and hypothesis tests that are asymptotically conservative because of the estimated weights; a nonparametric bootstrap approach would produce asymptotically exact results. Both methods are limited to large samples.

A subtle issue deserves mention. Our approaches to standardization of the risks within a given age category have

Table 4. Testing and Estimating Effect-Measure Modification by Age of the Adjusted Risk Difference for the Effect of Persons With Disability on Cost Barriers to Health Care, Based on Averaging Risks From Standard Logistic Regression for Complex Survey Data, Florida Behavioral Risk Factor Surveillance System Survey, 2007

Age Group, years	Risk for Persons With Disability[a]	95% CI	Risk for Persons Without Disability[a]	95% CI	Risk Difference[a,b,*]	95% CI
18–29	0.383	0.257, 0.509	0.225	0.183, 0.267	0.158	0.023, 0.293
30–64	0.240	0.212, 0.267	0.142	0.130, 0.155	0.097	0.067, 0.127
≥65	0.072	0.053, 0.091	0.040	0.028, 0.052	0.032	0.010, 0.054

Abbreviation: CI, confidence interval.
* $P < 0.001$.
[a] Adjusted for race/ethnicity, income, education, and gender using the outcome modeling approach.
[b] Test of effect-measure modification: χ^2 (2 df) = 14.38.

averaged risks with respect to the population distribution of confounders within that age category. One might instead seek to standardize all risks to the same distribution of confounders—for example, to the distribution of confounders in the overall population. Doing so would be procedurally equivalent to treating the modifier as a second exposure in the marginal structural model.

ACKNOWLEDGMENTS

Author affiliation: Department of Epidemiology and Biostatistics, College of Public Health and Health Professions, University of Florida, Gainesville, Florida (Babette A. Brumback, Erin D. Bouldin, Hao W. Zheng, Michael B. Cannell, and Elena M. Andresen).

This research was supported by the Centers for Disease Control and Prevention (grant U59 DD000273). In addition, the Suwannee River Area Health Education Center supported the public health internship of M. B. C.

The authors thank Janet Hess at the University of South Florida, Dr. David Wood at the University of Florida in Jacksonville, Dr. Phyllis Sloyer at Children's Medical Services, and Dr. John Reiss at the University of Florida, on behalf of the Health Care Transition Services Task Force for Youth and Young Adults with Disabilities and the Florida Health and Transition Services program, for their collaboration and direction on the disability research. They also thank Youjie Huang and Melissa Murray from the Florida Behavioral Risk Factor Surveillance System office for their helpful support.

Conflict of interest: none declared.

REFERENCES

1. Cannell B, Nur U, Bouldin ED, et al. *Access to Health Care and Health Behaviors Among Young Adults With and Without Disability in Florida: A Report From the Florida Behavioral Risk Factor Surveillance System 2007*. Gainesville, FL: Florida Office on Disability and Health, University of Florida; 2010. (http://fodh.phhp.ufl.edu/publications/). (Accessed March 1, 2010).
2. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology.* 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
3. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology.* 2003;14(6):680–686.
4. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550–560.
5. Bieler GS, Brown GG, Williams RL, et al. Estimating model-adjusted risks, risk differences, and risk ratios from complex survey data. *Am J Epidemiol*. 2010;171(5):618–623.
6. Korn E, Graubard B. *Analysis of Health Surveys*. New York, NY: John Wiley & Sons, Inc; 1999.
7. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol*. 2005; 162(3):199–200.
8. Zou GA. modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159(7): 702–706.
9. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. *Am J Epidemiol*. 2004;160(4): 301–305.
10. Brogan D. Software for sample survey data: misuse of standard packages. In: Armitage P, Colton T, eds. *Encyclopedia of Biostatistics*. 2nd ed. Chichester, United Kingdom: John Wiley & Sons Ltd; 2005:5057–5064.
11. Gentry EM, Kalsbeek WD, Hogelin GC, et al. The behavioral risk factor surveys: II. Design, methods, and estimates from combined state data. *Am J Prev Med*. 1985;1(6):9–14.
12. Remington PL, Smith MY, Williamson DF, et al. Design, characteristics, and usefulness of state-based behavioral risk factor surveillance: 1981–1987. *Public Health Rep*. 1988; 103(4):366–375.
13. VanderWeele TJ. On the distinction between interaction and effect modification. *Epidemiology.* 2009;20(6):863–871.
14. Chiba Y, Azuma K, Okumura J. Marginal structural models for estimating effect modification. *Ann Epidemiol*. 2009;19(5): 298–303.
15. Greenland S. Interval estimation by simulation as an alternative to and extension of confidence intervals. *Int J Epidemiol*. 2004;33(6):1389–1397.
16. Centers for Disease Control and Prevention. *Disability and Health State Chartbook, 2006: Profiles of Health for Adults With Disabilities*. Atlanta, GA: Centers for Disease Control and Prevention; 2006. (http://www.cdc.gov/ncbddd/dh/chartbook/Chartbook%20Text.pdf). (Accessed February 27, 2010).
17. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics*. 2005;61(4):962–973.

APPENDIX 1

```
proc surveyreg data=rd.stand;
class age_3level disability;
strata strat;
model costproblem = age_3level disability
age_3level*disability / solution clparm;
title 'Risk of low access to care because
of cost';
title2 'Linear regression of age and dis-
ability on access, adjusted';
weight combinedweight;
estimate '65 and over' disability -1 1
age_3level*disability 0 0 0 0 -1 1;
estimate '30-64' disability -1 1 age_
3level*disability 0 0 -1 1 0 0;
estimate '18-29' disability -1 1 age_
3level*disability -1 1 0 0 0 0;
run;
```

APPENDIX 2

```
proc surveylogistic data=rd.data1;
class age_3level race_4level income_
5level sex educ_4level / param=ref;
model disability (event = 'PWD') = age_
3level race_4level income_5level sex
educ_4level;
```

```
weight surveywt;
output out=rd.propout (keep=id propen-
sity) p=propensity;
run;

proc sort data=rd.data1;
by id;
run;

proc sort data=rd.propout;
by id;
run;

proc surveylogistic data=rd.data1;
class age_3level race_4level income_
5level sex educ_4level / param=ref;
model disability (event = 'PWD') =
age_3level;
weight surveywt;
output out=rd.pout (keep=id puncond)
p=puncond;
run;

proc sort data=rd.pout;
by id;
run;

data rd.stand;
merge rd.data1 rd.propout rd.pout;
by id;
if pwd=1 then iptw = puncond/propensity;
if pwd=0 then iptw = (1-puncond)/
(1-propensity);
combinedweight = iptw*surveywt;
run;
```