

Marginal Structural Models for Risk or Prevalence Ratios for a Point Exposure Using a Disease Risk Score

David B. Richardson*, Alexander P. Keil, Alan C. Kinlaw, and Stephen R. Cole

* Correspondence to Dr. David B. Richardson, Department of Epidemiology, School of Public Health, 2102B McGavran-Greenberg Hall, CB 7435, Chapel Hill, NC 27599 (e-mail: david.richardson@unc.edu).

Initially submitted May 11, 2018; accepted for publication January 25, 2019.

The disease risk score is a summary score that can be used to control for confounding with a potentially large set of covariates. While less widely used than the exposure propensity score, the disease risk score approach might be useful for novel or unusual exposures, when treatment indications or exposure patterns are rapidly changing, or when more is known about the nature of how covariates cause disease than is known about factors influencing propensity for the exposure of interest. Focusing on the simple case of a binary point exposure, we describe a marginal structural model for estimation of risk (or prevalence) ratios. The proposed model incorporates the disease risk score as an offset in a regression model, and it yields an estimate of a standardized risk ratio where the target population is the exposed group. Simulations are used to illustrate the approach, and an empirical example is provided. Confounder control based on the proposed method might be a useful alternative to approaches based on the exposure propensity score, or as a complement to them.

cohort studies; cross-sectional studies; regression analysis; standardization

Abbreviations: IPE, inverse probability of exposure; RR, risk ratio; SMR, standardized mortality ratio.

One approach to control for confounding of an exposure-outcome association by measured covariates commences by summarizing the relationship between these covariates and the exposure or outcome variable of interest, deriving a 1-dimensional variable. A widely used example of this approach is the exposure propensity score, which describes a person's probability of exposure (called "treatment") given their observed covariate pattern (1, 2). The exposure propensity score can be used for confounder control by matching, stratification, or regression model adjustment for the score, and it can be used to derive a weight that can be applied in a weighted regression (3–5).

An alternative summary variable, less widely used by epidemiologists (6), is the disease risk score, which describes a person's risk of disease given their observed covariate pattern (7). Glynn et al. (7) and Arbogast and Ray (8) provide useful overviews of the history and use of disease risk scores. The approach commences by developing a model to predict the outcome as a function of known confounders, which becomes the basis for the disease risk score. Hansen (9) has shown that stratification on a disease risk score can yield consistent effect estimates

with a meaningful counterfactual interpretation, and Stürmer et al. (10) have illustrated that under conditions typically encountered in empirical data analyses, models that control for confounding by regression adjustment for a disease risk score yield similar estimates and confidence intervals to those obtained by matching on an exposure propensity score. Use of the disease risk score for confounder control could be appealing in settings in which estimation of the exposure propensity score is problematic, such as when the exposure of interest is novel or very rare (8, 11) or when the pattern of exposure (or treatment indication) is rapidly evolving (7). In practice, the disease risk score has proved to be a useful tool, most notably in comparative effectiveness research (6, 7).

In this work, we focused on cohort analyses in which we wished to estimate the risk or prevalence ratio for a contrast defined by a point exposure. Building upon the disease risk score approach, we describe a novel way to estimate the parameters of a marginal structural model using a disease risk score. The approach could be a useful tool for flexible regression modeling in some commonly encountered research settings.

METHODS

Consider a study of n individuals in which Y denotes a binary outcome variable. Let X denote a binary exposure variable of primary interest and $\mathbf{Z} = \{Z_1, \dots, Z_k\}$ denote the k covariates that are potential confounders of the associations between X and Y . These study data might be organized in a data set with 1 record per person.

Suppose that the investigator wants to compare the occurrence of Y between groups defined by X , controlling for confounding by \mathbf{Z} . Specifically, assume that the investigator wants to estimate the exposure effect by comparing the exposed group mean (risk of the outcome) with the expected mean of counterfactual outcomes in a group with the same \mathbf{Z} distribution as the exposed (2). This comparison can be summarized as a risk ratio, $E(Y|X = 1)/E(Y_0|\mathbf{Z} = \mathbf{z}, X = 1)$, with a target population of the exposed group, where the potential outcome under the absence of exposure is denoted Y_0 . Note that we are interested in an estimate that is standardized to the covariate distribution in the exposed population,

$$\sum_{\mathbf{Z}} \frac{\Pr(Y|X = 1)}{\Pr(Y|\mathbf{Z} = \mathbf{z}, X = 0)} \Pr(\mathbf{Z} = \mathbf{z}|X = 1),$$

similar to other familiar standardized estimators such as the standardized mortality ratio (SMR) (12, 13). Such analyses estimate the effect of the exposure among the exposed, which often corresponds to the group targeted for interventions, particularly for novel treatments that might be concentrated in a unique subset of the population. This standardized ratio is unconfounded by \mathbf{Z} . If there is effect measure modification by \mathbf{Z} , this standardized ratio offers a useful marginal estimate of the effect of exposure in a population that has the covariate distribution observed among the exposed study subjects; in the absence of effect measure modification, the effect of the exposure among the exposed is equal to the effect of the exposure in the total population.

Disease risk score

The disease risk score, $F(\mathbf{Z})$, is a function of a vector of covariates, \mathbf{Z} , that confers conditional independence between the potential outcome under the absence of exposure, Y_0 , and \mathbf{Z} (i.e., $Y_0 \perp \mathbf{Z}|F(\mathbf{Z})$). Under our proposed approach, $F(\mathbf{Z})$ is estimated by a regression model fitted to the empirical data for the unexposed group; assuming a binary outcome variable, the disease risk score can be estimated as the expected value of Y_0 given \mathbf{Z} . Some authors have suggested that this function could be estimated using data for the entire cohort (i.e., exposed and unexposed) by including statistical terms for the product of exposure and covariates (9). Here, we propose to just use data for the unexposed group to limit potential for model misspecification (e.g., the need to model exposure effect modification). While rare outcomes often motivate case-control designs, rare exposures often motivate cohort studies, and a rare exposure typically implies a relatively large group of unexposed cohort members among whom it might be possible to reliably model the disease risk score. For the case of a single binary regressor variable, $\mathbf{Z} = z$, the disease risk score could be estimated by fitting a logistic model to the

data for the unexposed, $\log\left(\frac{\Pr(Y=1)}{\Pr(Y=0)}|\mathbf{Z}, X = 0\right) = \alpha_0 + \alpha_1\mathbf{Z}$, and then setting $F(z_i) = \exp(\alpha_0 + \alpha_1 z_i)/(1 + \exp(\alpha_0 + \alpha_1 z_i))$ for all individuals i . Therefore, given observed study data Y , X , and \mathbf{Z} , we can estimate disease risk score, $F(\mathbf{Z})$, for each study member.

A marginal structural model based on the disease risk score

Suppose that the data for this study of n individuals are sorted in descending order by X , such that the first m records represent people for whom $X = 1$ and the remaining $n - m$ records represent people for whom $X = 0$. Using these data, we can readily calculate $E(Y|X = 1)$ as $\frac{1}{m} \sum_{i=1}^m y_i$, and, under counterfactual consistency and the causal identification conditions given by Robins et al. (5), we can calculate $E(Y_0|\mathbf{Z} = \mathbf{z}, X = 1)$ as $\frac{1}{m} \sum_{i=1}^m F(z_i)$, and these quantities can be used to derive the desired risk ratio, $E(Y|X = 1)/E(Y_0|\mathbf{Z} = \mathbf{z}, X = 1)$.

An estimate of the standardized risk ratio can be obtained by fitting a generalized linear model with log link and Poisson distribution by including the log of the disease risk score as an offset,

$$\log(\Pr(Y = 1|X = x)) = \beta_0 + \beta_1 x + \log(F(\mathbf{Z} = \mathbf{z})),$$

where e^{β_1} is an estimate of the desired standardized ratio measure, while β_0 converges to 0 when $F(\mathbf{Z} = \mathbf{z})$ is estimated by a regression model fitted to the empirical data for the unexposed group. This model could be equivalently framed as a weighted Poisson regression in which the disease risk score is included as a weight in a log-linear regression, if one transforms the dependent variable, Y , for each subject to be $Y/F(\mathbf{Z})$. Estimation of robust confidence intervals (14) is recommended given the 2-stage regression (first estimation of disease risk score and second fitting the marginal structural model). In Web Appendix 1 (available at <https://academic.oup.com/aje>) we provide illustrative SAS (SAS Institute, Inc., Cary, North Carolina) code to obtain risk ratios as well as associated robust confidence intervals.

Simulation example

We used simulated data under a cohort study design to demonstrate the implementation of the proposed approach. Data were simulated for 1,000 cohort studies with 10,000 people in each cohort. In each simulation, we generated 10 covariates, denoted Z_1 – Z_{10} . Among these, Z_1 – Z_4 were confounders associated with both exposure and outcome, Z_5 – Z_7 were exposure predictors, and Z_8 – Z_{10} were outcome predictors. $Z_1, Z_3, Z_5, Z_6, Z_8, Z_9$ were random binary variables, and the others were continuous variables assigned as the absolute value of a standard normal random variables with zero mean and unit variance. The relationships between variables follow the structure described by Lee et al. (15), with correlations induced between several of the variables (Web Figure 1). We generated a random binary exposure, X , with exposure prevalence of approximately 10%; we encoded dependence of X on covariates by specifying that X took a value of 1 with probability

$1/(1 + \exp(-(-0.1-1 \times Z_1-0.5 \times Z_2-0.5 \times Z_3-0.5 \times Z_4-0.5 \times Z_5-0.5 \times Z_6-0.5 \times Z_7)))$.

We generated a random binary outcome Y in which we encoded dependence of the outcome on X and covariates. In simulation scenario 1, Y took a value of 1 with probability $\exp(-1 + 1 \times X-0.5 \times Z_1-0.1 \times Z_2-0.5 \times Z_3-0.1 \times Z_4-0.5 \times Z_8-0.5 \times Z_9-0.1 \times Z_{10})$. In simulation scenario 2, Y took a value of 1 with probability $\exp(-1 + 1 \times X-0.5 \times Z_1-0.1 \times Z_2-0.5 \times Z_3-0.1 \times Z_4-0.5 \times Z_8-0.5 \times Z_9-0.1 \times Z_{10} + 0.5 \times X \times Z_1)$. Note that in our first simulation scenario there are no products of X and covariates, while in the second simulation scenario there is a product of X and Z_1 when generating the data. Therefore, in the first simulation the risk ratio is homogeneous across levels of covariates, while in the second it is not. In the first simulation scenario, the average risk ratio among the exposed equals the average risk ratio in the total study population, while in the second scenario it does not.

We estimated the disease risk score by fitting a regression model to predict Y as a function of Z_1 – Z_{10} among the unexposed ($X = 0$), including each covariate as a main effect in the regression model (and not including any terms for products of covariates). We used the method described in this work (and SAS (SAS Institute, Inc.) code in Web Appendix 1) to obtain a marginal estimate of the risk ratio where the target population is the exposed group by fitting a general linear regression model for Y as a function of X with log link, Poisson distribution, including the natural log of the disease risk score as an offset. For comparison we also estimated the effect of X on Y using marginal structural log binomial models (Web Appendix 2), with robust variance and using weights derived from a model for exposure propensity (16). A logistic model was fitted to each simulated cohort to predict X as a function of Z_1, Z_2, \dots, Z_{10} , including each covariate as a main effect in the regression model (and not including any terms for products of covariates). The predicted probability of exposure from the fitted model is the propensity score. For comparability with results obtained using our proposed disease risk score–based approach, we constructed SMR weights to estimate the risk ratio where the target population is the exposed group (4, 17); exposed cohort members are given a weight of 1, while unexposed cohort members are given weights that are defined as the ratio of the estimated propensity score to 1 minus the estimated propensity score. We also constructed stabilized inverse-probability-of-exposure (IPE) weights to estimate the risk ratio where the target population is the total study population; exposed cohort members are given a weight defined as the ratio of the marginal probability of exposure to the estimated propensity score, while unexposed cohort members are given a weight defined as the ratio of 1 minus the marginal probability of exposure to 1 minus the estimated propensity score.

In each analysis, we summarized results from 1,000 simulated cohorts by computing the mean log risk ratio ($\log(\text{RR})$), the estimated standard deviation of the 1,000 $\log(\text{RR})$ s, and the average estimated robust standard error of the $\log(\text{RR})$ from the GENMOD procedure in SAS (SAS Institute, Inc.), available in versions 9.0 and thereafter. Additional simulations illustrating scenarios varying simulations conditions are included in Web Appendix 3.

Empirical example

We used data from the 2012 North Carolina Live Birth Certificate to illustrate this approach. We selected an example involving a large cohort with an exposure and outcome that were common enough to allow us to reliably estimate models for the disease risk score and the exposure propensity score for the purposes of comparison. The outcome of interest, preterm birth, was defined as a live birth during the 17-week interval beginning with the 21st week of gestation and ending when 37 weeks of gestation are completed. The exposure of primary interest was “smoke,” a dichotomous variable indicating whether the mother reported smoking during the pregnancy. The study population consisted of 64,616 singleton North Carolina live births for which the birth and the entire 17-week risk period for preterm birth (the 21st week of gestation through the 37th week of gestation) occurred in 2012. Covariates, selected from among variables that are available on the North Carolina birth certificate and were considered potential confounders, included age (a continuous variable for maternal age in years), race (a 4-level variable for maternal race/ethnicity: non-Hispanic white, Hispanic white, African-American, other), and “pnc” (a dichotomous variable for prenatal care during the first 20 weeks of gestation).

We estimated disease risk scores by fitting a logistic regression model for preterm as the dependent variable, restricted only to the unexposed population (i.e., individuals with $\text{smoke} = 0$); explanatory variables were age , age^2 , race , and pnc , and product terms were included of the form $\text{age} \times \text{pnc}$, $\text{age}^2 \times \text{pnc}$, $\text{age} \times \text{race}$, and $\text{age}^2 \times \text{race}$. Linear and quadratic terms for age were included to model nonlinearities in the association between maternal age and the log risk of preterm. Higher-order polynomial functions of age led to negligible improvement in model fit. We fitted a log-linear Poisson regression model with the log disease risk score as offset. For comparison, we also fitted models weighted as a function of the exposure propensity score, using SMR weights, and stabilized IPE weights. To calculate denominators for SMR weights in the unexposed and IPE weights for the entire population, we estimated exposure propensity scores for smoke as the dependent variable, with explanatory variables being the same covariate vector used to estimate disease risk scores.

RESULTS

Simulations

In the first simulation scenario, the effect of exposure was homogeneous across levels of covariates. Web Table 1 reports the distributions of Z_1 – Z_{10} between subgroups defined by X in the simulated cohorts. Table 1 reports mean $\log(\text{RR})$ estimates and standard errors based on simulations with 10,000 people in each cohort. The crude estimate ($\log(\text{RR}) = 1.12$) differs from the value specified under the simulation setup (i.e., $\log(\text{RR}) = 1.00$). The mean of estimates obtained using the proposed approach incorporating the log disease risk score as a regression model offset was 1.00, conforming to the true value specified under simulation scenario 1. The robust 95% confidence intervals for the $\log(\text{RR})$ estimates obtained using the proposed approach appeared to have reasonable coverage,

Table 1. Mean Estimated Log Risk Ratios, Empirical Standard Error, and Average Estimated Standard Error for 1,000 Cohorts With 10,000 Observations Each in a Simulation Study

Simulation	log(RR)	ESE	ASE ^a
Scenario 1			
Crude	1.12	0.05	0.05
Disease score (proposed method)	1.00	0.05	0.05
Exposure score (SMR-weighted)	1.00	0.06	0.05
Exposure score (IPE-weighted)	1.00	0.07	0.07
Scenario 2			
Crude	1.32	0.05	0.05
Disease score (proposed method)	1.12	0.05	0.04
Exposure score (SMR-weighted)	1.12	0.05	0.05
Exposure score (IPE-weighted)	1.22	0.06	0.06

Abbreviations: ASE, average estimated standard error; ESE, empirical standard error; IPE, inverse probability of exposure; RR, risk ratio; SMR, standardized mortality ratio.

^a For crude models, ASE is the average standard error estimated across 1,000 simulated cohorts. For weighted regression models, ASE is the average robust standard error.

encompassing the value specified under the simulation setup in 95.5% of simulations. For comparison, we fitted a marginal structural log-binomial model using weights derived from an exposure propensity score model; the mean of the estimates obtained using SMR-weighted regression models also conformed to the true value specified under the simulation ($\log(\text{RR}) = 1.00$). Similar results also were obtained when we fitted a marginal structural log-binomial model using stabilized IPE weights (average $\log(\text{RR}) = 1.00$), reflecting the fact that the estimate of the standardized risk ratio obtained when the target population of the study is the exposed group (which is the estimate derived under the proposed approach and is derived by SMR-type weighting using exposure propensity score–based weights) equals the estimate of the standardized risk ratio when the target population is the total population (which is the estimate derived under inverse probability of exposure weighting) when there is no modification of the effect measure across levels of covariates. The mean stabilized IPE weight was 1.00, the minimum weight was 0.24, and the 5th, 25th, 50th, 75th, and 95th percentiles of weights were 0.77, 0.94, 0.97, 1.03, and 1.21.

In simulation scenario 2, data were generated such that the risk ratio was heterogeneous across strata of Z_1 . The estimated standardized $\log(\text{RR})$ from the regression model with the log disease risk score as offset was 1.12. Similarly, when we fitted a marginal structural model using weights derived from an exposure propensity score model and SMR-type weighting, the average estimated standardized log risk ratio was 1.12. In contrast, fitting a model based on stabilized IPE weighting yielded an estimate of 1.22, which estimates the average treatment effect among the total population, which we expect to differ in simulation scenario 2 from an estimate of the standardized risk ratio when the target population is the exposed group, as obtained from the proposed model incorporating a disease risk score as offset.

Additional simulations were conducted to illustrate performance of the proposed method under conditions of smaller cohort size, lower prevalence of exposure, and nonlinearity/nonadditivity of effects (Web Tables 2–4). Under the first simulation scenario the effect of exposure was homogeneous across levels of covariates. Given a smaller cohort size (1,000 observations per cohort), we found that the mean of estimates obtained using the proposed approach was 1.00, as was the mean of the estimates obtained using SMR-weighted regression models, conforming to the true value specified under the simulation scenario. Similar results also were obtained when we fitted a marginal structural log-binomial model using stabilized IPE weights, although the mean of the estimates was slightly less than 1.0. In simulations where the exposure prevalence was 5% or 2.5%, the mean of estimates obtained using the proposed disease risk score approach, and obtained using SMR-type weights, were 0.98, and 0.97, respectively (Web Tables 3 and 4). When fitting a marginal structural log-binomial model using IPE weights, the average $\log(\text{RR})$ was 0.97 and 0.92; across the simulations, the average standard errors tended to be larger under models based on IPE weighting than under the proposed disease risk score model.

Web Appendix 3 also reports simulation scenarios where there was nonlinearity and/or nonadditivity in the effects of covariates while the fitted disease risk score and exposure propensity score models included only linear and additive terms for covariates. Overall, the 3 methods performed well in simulation scenarios where the exposure prevalence was 10% or 5% (Web Tables 2 and 3). The mean of estimates obtained using the proposed disease risk score approach, and those obtained using SMR-type weights, ranged from 0.97 to 1.00 under the scenarios considered, and the average $\log(\text{RR})$ estimates obtained using a marginal structural log-binomial model with IPE weights ranged from 0.96 to 1.00 under these scenarios. When the exposure prevalence was 2.5% (Web Table 4), the mean of estimates obtained using the proposed disease risk score approach, and those obtained using SMR-type weights, ranged from 0.94 to 0.98 under the scenarios considered. Under a marginal structural log-binomial model using IPE weights, the average $\log(\text{RR})$ estimates ranged from 0.90 to 0.92 under the scenarios considered; across the simulations, the average standard errors tended to be larger under models based on SMR-type and IPE weights than under the proposed disease risk score model.

Empirical example

In the North Carolina live birth data, the cumulative incidence of preterm birth is 9.8%. The prevalence of the primary exposure, maternal smoking (*smoke*), is 10.6%. In the observed data, the average maternal age among those who reported smoking (*smoke* = 1) was 2 years less than among those who did not (*smoke* = 0), and 86% of those with *smoke* = 1 received prenatal care while 92% of those with *smoke* = 0 received it. Using the proposed disease risk score approach, the (standardized) risk ratio for the association between smoking and preterm birth was 1.37 (robust 95% confidence interval: 1.28, 1.46) (Table 2). For comparison, we fitted a log-binomial model to the SMR-weighted data for the association between smoking and preterm birth, which yielded an estimate of the standardized risk ratio of 1.37 (robust 95% confidence interval: 1.27, 1.47), essentially

identical to the disease risk score–based marginal structural model estimate of the standardized risk ratio among the exposed. We also fitted a log-binomial model to the IPE-weighted data for the association between smoking and preterm birth, which yielded an estimate of the standardized risk ratio of 1.46 (robust 95% confidence interval: 1.32, 1.60), which is an estimate of the smoking effect in the total population. The mean stabilized IPE weight was 1.00, the minimum weight was 0.30, and the 5th, 25th, 50th, 75th, and 95th percentiles of weights were 0.77, 0.94, 0.98, 1.02, and 1.20.

DISCUSSION

Based on the variety of simulation settings we have considered, we found that the proposed approach yields a marginal estimate of the risk ratio that is equivalent to an estimate obtained in an analysis using an SMR-type weighted log-binomial regression model based on exposure propensity scores—namely a marginal estimate of the risk ratio where the target population is the exposed. We found that an IPE-weighted estimate of the average risk ratio in the total study population was not well estimated, and we noted in simulations that as the exposure prevalence diminished (from 10% to 5% to 2.5%), the IPE-weighted estimate tended to perform less well than the proposed disease risk score–based method. Also, the empirical standard error and average estimated standard error for the IPE method tended to diverge from each other and tended to be larger than the empirical standard error and average estimated standard error obtained for the proposed disease risk score–based method (Web Appendix 3). We caution, however, that conclusions from our limited range of simulation-based results are not conclusive. Moreover, in these simulations we focused solely on results for the disease risk score, SMR-type weighted, and IPE-weighted regression models quantified with robust standard errors, and we did not consider comparability of bootstrap-based measures of precision. In an empirical data analysis we found that an estimate of the association of interest between maternal smoking and preterm birth obtained using the proposed approach (i.e., incorporating the log of the disease risk score as a regression model offset) yielded an equivalent estimate of the standardized risk ratio to that obtained from a weighted log-binomial regression with SMR-type weights derived based on an exposure propensity score model.

In our proposed approach, we fitted a model to estimate the disease risk score using information only for the unexposed,

for whom Y_0 is observed under counterfactual consistency. Scores are then extrapolated using this model to predict the counterfactual disease risk for exposed individuals within the cohort. Alternatively, this model could be fitted in the total population by including the exposure of interest in the regression model (18); disease risk scores are then assigned to each individual after setting treatment status to zero. Arbogast and Ray (19) evaluated the performance of these two strategies and found that if the disease risk score model is correctly specified, then confounding control tended to be better when the score was modeled using the full cohort versus only within the comparator subcohort. Another approach, and the one classically used in SMR analyses, is to estimate the disease risk score in data for an external reference population.

Prior authors have described confounder control with the disease risk score implemented through matching, stratification, or covariate adjustment (e.g., entering the score by means of indicator terms for categories) (7–9). In the present study, we describe an approach to estimation of a standardized risk ratio (where the target population is the exposed group) by incorporating the log disease risk score as an offset in a regression model. This avoids coarsening (as might occur when matching, stratifying, or adjusting for categories defined by the disease risk score) or the need to specify a parametric form (as occurs when including the disease risk score as a continuous explanatory variable in a regression model). It was recently noted that the epidemiologic literature provides no theory or examples of methods for weighting on the disease risk score (20). The present study fills that gap; we note that the approach could be implemented by weighting on the disease risk score (after appropriately transforming the outcome variable).

The disease risk score has been less widely used by epidemiologists than the exposure propensity score (6). While there might be historical (21, 22) as well as theoretical reasons for this (7–9), one practical limitation of disease risk score–based methods (compared with exposure propensity score–based approaches) is that it is not easy to assess balance of covariates in a disease risk–based analysis, as can be readily done with exposure propensity score–based approaches. For example, with IPE-weighted marginal structural models, one can undertake a simple assessment of covariate balance across exposure groups using the weighted data. Despite such limitations, the proposed disease risk score model might be appealing when one wishes to obtain a standardized (i.e., marginal) estimate of association and the exposure propensity score is difficult to estimate well.

Glynn et al. (7) reviewed study design features that influence the value or feasibility of analyses using a disease risk score relative to using an exposure propensity score. One such setting is when there is substantial historical information available to help inform model development for the disease risk score, while little is known about indications for the treatment or exposure. Development of disease risk score models could be informed by prior literature on the nature of how covariates cause the disease of interest, which is useful if information available for modeling the probability of exposure is sparse. Another setting is when the exposure conditions, or indications for treatment, are rapidly evolving. Rare exposures for small, or moderately sized, cohorts also can pose challenges for estimation of the exposure propensity score (8). In addition, prior simulation studies have illustrated problems of bias

Table 2. Estimates of Association Between Maternal Smoking and Preterm Birth Among 66,172 Live Births in North Carolina, 2012

Model	RR	95% CI ^a
Crude	1.33	1.25, 1.43
Disease score (proposed method)	1.37	1.28, 1.46
Exposure score (SMR-weighted)	1.37	1.27, 1.47
Exposure score (IPE-weighted)	1.46	1.32, 1.60

Abbreviations: CI, confidence interval; IPE, inverse probability of exposure; RR, risk ratio; SMR, standardized mortality ratio.

^a Robust confidence intervals were estimated to account for within-subject correlation induced by weighting.

that can arise in exposure propensity score analyses when the exposure is rare (11). For the proposed disease risk score approach, rather than requiring exposed and unexposed at each level of covariate, the disease risk score model requires only positivity across the balancing disease risk score (often a weaker condition termed “risk positivity,” requiring no values of disease risk at which treatment is received or not received with certainty) (19).

Prior authors have also suggested that the disease risk score might be appealing when there are more than 2 exposures of interest (7) or when the exposure of interest is polytomous or continuous (8). Exposure propensity score modeling can get complex in such situations, while in principle polytomous or continuous exposure variables pose no additional problems for the disease risk score–based models (because estimation of a disease risk score is no more difficult in a cohort study where the exposure of interest is a polytomous variable than it is if the exposure is a dichotomous variable). However, if the exposure of interest is not binary, but rather is a polytomous variable, one needs to reflect carefully on interpretation of the estimates that involve pair-wise comparisons with the reference group, because different exposure groups might be standardized to different *Z* distributions. Consequently, to demonstrate the applicability of disease risk scores for estimating a standardized risk ratio for the average effect of exposure in the exposed, we have focused in the present study on the simple case of a binary exposure.

The proposed model-based standardization using the disease risk score provides a potentially useful tool in cohort analysis for estimation of risk or prevalence ratios. One can obtain comparable estimates to those derived using marginal structural models with exposure propensity score–based weighting, yielding a standardized estimate of the risk ratio where the target population is the exposed.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (David B. Richardson, Alexander P. Keil, Stephen R. Cole); Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Alan C. Kinlaw); Department of Pediatrics, University of North Carolina School of Medicine, Chapel Hill, North Carolina (Alan C. Kinlaw); and Division of Pharmaceutical Outcomes and Policy, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Alan C. Kinlaw).

A.C.K. received funding support from a National Research Service Award Post-Doctoral Traineeship from the Agency for Healthcare Research and Quality sponsored by the Cecil G. Sheps Center for Health Services Research at the University of North Carolina at Chapel Hill (grant 5T32 HS000032-28). A.P.K. received funding support from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant DP2-HD-08-4070).

Conflict of interest: none declared.

REFERENCES

1. Robins JM, Mark SD, Newey WK. Estimating exposure effects by modelling the expectation of exposure conditional on confounders. *Biometrics*. 1992;48(2):479–495.
2. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41–55.
3. Kurth T, Walker AM, Glynn RJ, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *Am J Epidemiol*. 2006;163(3):262–270.
4. Brookhart MA, Wyss R, Layton JB, et al. Propensity score methods for confounding control in nonexperimental research. *Circ Cardiovasc Qual Outcomes*. 2013;6(5):604–611.
5. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550–560.
6. Tadrous M, Gagne JJ, Stürmer T, et al. Disease risk score as a confounder summary method: systematic review and recommendations. *Pharmacoepidemiol Drug Saf*. 2013;22(2):122–129.
7. Glynn RJ, Gagne JJ, Schneeweiss S. Role of disease risk scores in comparative effectiveness research with emerging therapies. *Pharmacoepidemiol Drug Saf*. 2012;21(suppl 2):138–147.
8. Arbogast PG, Ray WA. Use of disease risk scores in pharmacoepidemiologic studies. *Stat Methods Med Res*. 2009;18(1):67–80.
9. Hansen BB. The prognostic analogue of the propensity score. *Biometrika*. 2008;95(2):481–488.
10. Stürmer T, Schneeweiss S, Brookhart MA, et al. Analytic strategies to adjust confounding using exposure propensity scores and disease risk scores: nonsteroidal antiinflammatory drugs and short-term mortality in the elderly. *Am J Epidemiol*. 2005;161(9):891–898.
11. Hajage D, Tubach F, Steg PG, et al. On the use of propensity scores in case of rare exposure. *BMC Med Res Methodol*. 2016;16:38.
12. Breslow NE, Day NE. *Statistical Methods in Cancer Research: The Design and Analysis of Cohort Studies*. Lyon, France: International Agency for Research on Cancer; 1987.
13. Keil AP, Edwards JK, Richardson DB, et al. The parametric g-formula for time-to-event data: intuition and a worked example. *Epidemiology*. 2014;25(6):889–897.
14. Huber PJ. The behavior of maximum likelihood estimates under non-standard conditions. Presented at Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, 1967.
15. Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. *Stat Med*. 2010;29(3):337–346.
16. Richardson DB, Kinlaw AC, MacLehose RF, et al. Standardized binomial models for risk or prevalence ratios and differences. *Int J Epidemiol*. 2015;44(5):1660–1672.
17. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology*. 2003;14(6):680–686.
18. Miettinen OS. Stratification by a multivariate confounder score. *Am J Epidemiol*. 1976;104(6):609–620.
19. Arbogast PG, Ray WA. Performance of disease risk scores, propensity scores, and traditional multivariable outcome regression in the presence of multiple confounders. *Am J Epidemiol*. 2011;174(5):613–620.
20. Wyss R, Glynn RJ, Gagne JJ. A review of disease risk scores and their application in pharmacoepidemiology. *Curr Epidemiol Rep*. 2016;3(4):277–284.

21. Pike MC, Anderson J, Day N. Some insights into Miettinen's multivariate confounder score approach to case-control study analysis. *J Epidemiol Community Health*. 1979;33(1): 104–106.
22. Cook EF, Goldman L. Performance of tests of significance based on stratification by a multivariate confounder score or by a propensity score. *J Clin Epidemiol*. 1989;42(4): 317–324.