Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2013 February ; 22(2): 275–285. doi:
10.1158/1055-9965.EPI-12-1050.

A Comparison of the Polytomous Logistic Regression and Joint Cox Proportional Hazards Models for Evaluating Multiple Disease Subtypes in Prospective Cohort Studies

Xiaonan Xue¹, Mimi Y. Kim¹, Mia M. Gaudet², Yikyung Park³, Moonseong Heo¹, Albert R. Hollenbeck⁴, Howard D. Strickler¹, and Marc J. Gunter⁵

¹Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, NY 10461

²Epidemiology Research Program, American Cancer Society, Atlanta, GA 30303

³Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20852

⁴AARP, Washington D.C

⁵Department of Epidemiology and Biostatistics, School of Public Health, Imperial College, London, W2, 1PG, United Kingdom

Abstract

Background—Polytomous logistic regression models are commonly used in case-control studies of cancer to directly compare the risks associated with an exposure variable across multiple cancer subtypes. However, the validity, accuracy and efficiency of this approach for prospective cohort studies have not been formally evaluated.

Methods—We investigated the performance of the polytomous logistic regression model and compared it to an alternative approach based on a joint Cox proportional hazards model using simulation studies. We then applied both methods to a prospective cohort study to assess whether the association of breast cancer with body size differs according to estrogen and progesterone receptor-defined subtypes.

Results—Our simulations showed that the polytomous logistic regression model but not the joint Cox regression model yielded biased results in comparing exposure and disease subtype associations when the baseline hazards for different disease subtypes are non-proportional. For this reason, an analysis of a real data set was based on the joint Cox proportional hazards model and showed that body size has a significantly greater association with estrogen and progesterone positive breast cancer than with other subtypes.

Conclusions—Because of the limitations of the polytomous logistic regression model for the comparison of exposure-disease associations across disease subtypes, the joint Cox proportional hazards model is recommended over the polytomous logistic regression model in prospective cohort studies.

Corresponding author: Xiaonan Xue, Division of Biostatistics, Department of Epidemiology & Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, xiaonan.xue@einstein.yu.edu, Tel: 718-430-2431, Fax: 718-430-8780.

Potential conflicts of interest: The authors do not have a commercial or other association that might pose a conflict of interest.

Impact—The paper will promote the use of the joint Cox model in a prospective cohort study. Examples of SAS and S-plus programming codes are provided to facilitate use by non-statisticians.

Keywords

time to event data; proportional hazards functions; robust variance; competing risk

Introduction

There is growing recognition of the importance of distinguishing histologic or molecular subtypes in etiologic studies of cancer. As a result, comparing exposure effects across multiple subtypes of a given cancer is often of primary interest. For example, Schroeder et al (1) examined whether the association of cigarette smoking with bladder cancer differed according to the presence of p53 mutations in the tumor specimen. The conventional statistical approach for analyzing case-control studies with more than one outcome subtype is to fit a polytomous logistic regression model (2) (for simplicity, referred to as polytomous model). In the polytomous model, the associations of genetic or environmental factors with all disease subtypes are estimated simultaneously and inferences comparing associations across subtypes can be performed (3–7).

Polytomous models have also been employed in prospective cohort studies to compare the associations between exposure variables and occurrence of multiple disease subtypes (8–13). In the Iowa Women's Health Cohort Study (8), for example, polytomous models were applied to compare the effects of specific risk factors on the development of different types of breast cancer defined according to the presence or absence of estrogen (ER+/-) and progesterone receptor (PR +/-) expression status. However, the performance of the polytomous model, i.e., the validity, accuracy and efficiency of the model, for prospective cohort studies has not been adequately examined in the epidemiologic and statistical literature. A major drawback of applying the polytomous model in prospective cohort studies is that participants are classified only according to whether or not the specific subtype of disease occurred - data on the actual time to development of the different subtypes are not used. An additional limitation of the polytomous method relates to the nature of the control group. For example, to evaluate the association of a specific exposure with ER+ breast cancer, the control group includes subjects who did not develop any type of breast cancer during the follow-up period, excluding those who developed ER- breast cancer. Therefore, to approximate the odds ratio (OR) associated with the risk of ER+ breast cancer, it is necessary to assume that ER- breast cancer is rare (event rate < 10%), and vice versa. However, when the event is not rare, polytomous models could lead to severely biased estimates of exposure-disease associations.

As an alternative in a prospective cohort study setting, we propose to use a joint Cox proportional hazards model (for simplicity, referred to as joint Cox model) to examine associations of risk factors with multiple disease subtypes. The model accommodates time-to-event data, different baseline hazard functions of each disease subtype, and direct comparison of associations with outcome subtypes. This model was originally proposed for the analysis of competing risk events (14–16). Chatterjee, et al (17) extended it to model a disease with a large number of possible subtypes defined by multiple disease characteristics and to allow for the possibility that some cases cannot be categorized into subtypes because of missing characteristics. A two-stage Cox model and an estimating equation approach were proposed to estimate exposure-disease subtype associations. However, polytomous models remain the most commonly used method for the analysis of multiple disease subtypes because of lack of recognition of potential problems associated with the model and

limited familiarity with the formulation and implementation of the two-stage Cox model. In this paper, we formally evaluate the performance of the polytomous model and compare it with the Cox model using both mathematical and simulation-based approaches. For the Cox model, we consider a special case of Chatterjee's model for a limited number of disease subtypes and propose a joint Cox model which can be directly estimated using existing software.

We then apply the methods to a real data set from the National Institute of Health-AARP Diet and Health Study to examine heterogeneity in the association of body size, defined by body mass index (BMI), with breast cancer subtypes characterized by the tumor's ER/PR status (18). In a previous analysis using the polytomous model BMI was shown to have a significant association only with ER+/PR+ breast cancer but not with the other subtypes. However, statistical power was limited because of a relatively small number of cases, particularly for subtypes other than ER+/PR+; in addition, the validity of the polytomous model was not examined. This paper presents an extension of the original analysis with 5 more years of follow-up.

Materials and Methods

Consider a disease with only two subtypes, denoted as event 1 and event 2, which are competing risks (e.g., ER+ and ER- breast cancer). The hazard functions for time to event 1 and 2, T_1 and T_2 , are $\lambda_1(t)$ and $\lambda_2(t)$ respectively. Individuals who develop one event are considered censored at the time of the other event. Assume that the hazard for each type of event is affected by a set of common risk factors, X , according to the following Cox proportional hazards model:

$$\lambda_k(t) = \lambda_{0k}(t) \exp(x\beta_k) \quad (1)$$

where $\lambda_{0k}(t)$ is the baseline hazard function and β_k denotes the log hazard ratio associated with a unit increase in X (e.g., 1 kg/m² increase in BMI) for the development of the k^{th} event type ($k=1, 2$). By fitting a separate Cox model for each event, one can estimate β_1 and β_2 but not evaluate if $\beta_1 = \beta_2$. The polytomous logistic regression model transforms the time-to-event data to a categorical outcome in order to examine $\beta_1 = \beta_2$.

Polytomous Logistic Regression Model

Let D denote a polytomous outcome where $D=0$ if the subject remains free of disease during follow up; $D=k$ if event k occurs for $k=1, 2$. The polytomous model, also known as the multinomial logistic regression model, can be expressed as:

$$\log \frac{P(D=k|X)}{P(D=0|X)} = \alpha_k + \gamma_k X. \quad (2)$$

The ratio $P(D=k|X)/P(D=0|X)$ in model (2) is referred to as the "odds-like" term for disease type $D=k$ in comparison to the reference category of no disease ($D=0$). Note that this odds-like term for event k is not the same as the odds of event k , $P(D=k|X)/\{1 - P(D=k|X)\}$, since the odds denominator is the sum of the probability of no event and the probability of all events other than k . It follows from (2) that $e^{\gamma_k} \equiv OR_k^*$, the ratio of the odds-like terms for event k associated with a unit increase in X , which is again different from the usual OR.

The polytomous model tests whether $\gamma_1 = \gamma_2$ or equivalently $OR_1^* = OR_2^*$. It is commonly assumed that the coefficients obtained using the polytomous models are similar to the

corresponding coefficients in the Cox proportional hazards model, i.e., $\gamma_k \approx \beta_k$. However, as we have shown in Appendix I, this is only true when both events are rare and their baseline hazards functions are proportional. For example, $\gamma_1 > \beta_1$ when $\beta_1 > 1$ and the event of interest is not rare. Nonetheless, a test of $\gamma_1 = \gamma_2$ can provide a valid test for $\beta_1 = \beta_2$ if the baseline hazards are proportional. Appendix I shows that as long as the two events have proportional baseline hazards functions, $\gamma_1 - \gamma_2 = \beta_1 - \beta_2$ is true regardless of whether the event rates are rare or not. Furthermore, if the subjects who are lost to follow-up without developing any events are included in the non-event category, the polytomous model continues to provide a valid test for $\beta_1 = \beta_2$ as long as the censoring is uninformative of future events.

If the baseline hazard functions of the different events are not proportional over time, $\gamma_1 - \gamma_2 = \beta_1 - \beta_2$ does not necessarily hold. It is possible that, for example, the baseline hazards function increases with age for one event but decreases with age for the other event. In this case, an alternative approach for comparing the effects of X on the risks of the different events is needed.

Chatterjee extended the competing risk framework to a two-stage Cox model to model disease subtypes, in which the subtypes are defined by multiple disease characteristics some of which may be missing for some of the cases (17,19). However, for the reasons cited above, polytomous models are still the most commonly used methods to compare exposure effects across multiple disease subtypes. Here we consider a special case of Chatterjee's model in which the disease is categorized into a limited number of subtypes to have sufficient number of cases in each category and no further substructure among the disease subtypes is assumed. We propose a joint Cox proportional hazards modeling approach with an estimating procedure that differs slightly from Chatterjee's (17) so that standard procedures in commonly used statistical software can be directly applied to evaluate the exposure and disease associations for multiple disease subtypes in a prospective cohort setting.

Joint Cox Proportional Hazards Model

In a joint Cox regression model, the time-to-event 1 and event 2 are simultaneously modeled. Specifically, each subject contributes two potential event times, one for each event. Suppose the follow-up time for a subject is τ (Appendix II shows the definition of the two events). Marginally, the hazards function for event k is defined the same as the separate Cox model for the event (equation(1)), however, the parameters of interest β_1 and β_2 are estimated simultaneously by stratifying on event type and using a robust variance estimator to account for the correlation between the survival times of the two events (20). Specifically, equation (1) is reparameterized as follows:

$$\lambda_k(t) = \lambda_{0k}(t) \exp(\beta^* X + \theta I_{k=2} X) \quad (3)$$

where $I_{k=2}$ equals to 1 if the event time is for event 2 and 0 otherwise so that the hazard ratios (HR) associated with X for the development of events 1 and 2 are e^{β^*} and $e^{\beta^* + \theta}$, respectively. Note that equation (3) is equivalent to equation(1) so that $\beta_1 = \beta^*$ and $\beta_2 = \beta^* + \theta$. Thus, an estimate of θ provides an estimate of $\beta_1 - \beta_2$ and a test of $\beta_1 = \beta_2$ is equivalent to a test of $\theta = 0$, which can be performed with a Wald test.

Simulation study

We initiated a simulation study to: (1) evaluate the polytomous model estimates of the odds-like ratio as an estimate of OR and HR; (2) evaluate the performance of the polytomous model for comparing the exposure-disease associations between two competing events; (3)

evaluate the performance of the joint Cox regression model for estimating and comparing the exposure-disease associations for both events; and (4) compare the performances of the joint Cox model and the polytomous model when the baseline hazards functions for the two events are non-proportional.

The simulation study was designed as follows:

1. Sample size: $N=2000$; for simplicity, we considered X to be a single binary covariate with $P(X=1)=0.50$. For example, X can be a biomarker that was originally continuous but dichotomized at its median level;
2. Time to event: Times to events 1 and 2 were generated from exponential distributions based on equation (1), while assuming various magnitudes of baseline event rate and strengths of exposure-disease associations (Table 1). Events 1 and 2 have hazards functions that are proportional to each other in this case. To evaluate the effect of non-proportional baseline hazards, time to the second event was generated from a Weibull distribution with an increasing hazard function.
3. Censoring due to competing events and follow-up length of 10 years was generated; censoring due to loss to follow-up was also considered by generating a random censoring time from an independent exponential distribution (See Appendix II for detail).

Eight scenarios were generated (Table 1): scenarios(a)–(d) considered the cases when times to event 1 and event 2 have proportional hazards functions while scenarios(e)–(h) considered the cases when the hazards function for times to event 1 and 2 are not proportional to each other. For scenarios with the proportional hazards functions, (a) and (b) considered rare events while (c) and (d) considered non-rare events; (a) and (c) considered X has an effect on event 1 but not on event 2 while (b) and (d) considered X has the same effects on both events. The set with non-proportional hazards functions were designed similarly, i.e., (e) & (f) for rare events and (g) & (h) for non-rare events and (e) & (g) considered differential effects and (f) & (h) considered same effects for events 1 and 2.

The simulation procedure was repeated 500 times. The performance of each model was summarized according to relative bias, coverage probability of the confidence interval (CI) for estimating OR_2 , e^{β_2} , $e^{\beta_1-\beta_2}$ and empirical power or empirical type I error for testing equality of exposure-disease associations (for simplicity, we omit results for event 1). A coverage probability close to 95% indicates that the variance estimator is consistent, provided that the parameter estimator is unbiased. The empirical power indicates how well the model is able to detect a true exposure effect or differential exposure effects and the closeness of the empirical type I error to the target level of 5% is a measure of the validity of the statistical test.

Comparison of the Polytomous and Joint Cox Model for estimating the association of BMI with different breast cancer subtypes using data from the National Institutes of Health (NIH)-AARP Diet and Health Study

We used the polytomous model and the joint Cox model to formally compare the association of BMI with breast cancer by ER and PR status using data from the NIH-AARP Diet and Health Study. The NIH-AARP study was established in 1995 when a baseline screening questionnaire that queried information on medical history and lifestyle characteristics was returned by 566,398 AARP (formerly known as the American Association of Retired Persons) members. Participants were aged 50–71 and resided in 1 of 6 US States (California, Florida, Louisiana, New Jersey, North Carolina and Pennsylvania) or two metropolitan areas (Atlanta, Georgia, and Detroit, Michigan), selected because they were known to have high-quality cancer registries and a large AARP membership (18).

A total of 54,629 women remained in the analytic cohort after exclusion of pre-menopausal women, women with previous history of cancer as well as current users of hormone therapy and women residing in Florida or Pennsylvania because these two states have substantial proportion of participants with missing ER/PR data. In this cohort, 1,492 cases were diagnosed with breast cancer with known ER/PR status obtained from cancer registry records, in which 246 were classified as ER-/PR- breast cancer cases (16.5%), 231 ER+/PR- cases (15.5%), 18 ER-/PR+ cases (0.01%) and 997 ER+/PR+ cases (66.8%). Body mass index (BMI; kg/m²) was derived from self-reported height and weight on the baseline survey. In statistical models, BMI was expressed as categorical variables (<25.0 kg/m²; 25.0–29.9 kg/m²; 30.0 kg/m²). Covariates included in the multivariable models were consistent with those reported in the prior analysis by Ahn et al (18), namely age, race or ethnic, family history of breast cancer, level of education, age at menarche, age at menopause, age at first birth, parity, smoking status, physical activity, fat intake, alcohol consumption, oophorectomy, and height. The HRs for the association of BMI with the four breast cancer subtypes were estimated simultaneously and the difference in HRs were assessed. The analysis was conducted using SAS 9.2 (SAS Institute, Cary NC).

Results

Simulation Study

When hazards functions for events 1 & 2 are proportional and both outcomes are rare (scenarios(a)–(b)), the odds-like ratio estimated in the polytomous model is unbiased for the true OR as well as the true HR as indicated by the small relative bias (<2%) (Table 2). The variance of the estimator is also unbiased since the coverage probability of CIs is close to 95%. When the events are more common, the odds-like ratio obtained from the polytomous model can be a biased estimate for the true OR and HR. The level of the bias increases with incidence rates and the magnitude of the exposure-disease association: when both event rates are above 10%, the relative bias for the odds-like ratio as an estimate for HR_2 is 4.6% when $HR_2=1$ (scenario(c)) and 6.3% when $HR_2=1.33$ (scenario (d)). Furthermore, the odds-like ratio can be a much poorer estimate of the HR than of the OR: for example in scenario(d), the relative bias is 2.2% for OR_2 and 6.3% for HR_2 ; the estimates are worse with non-proportional hazards functions: in scenario(h), the relative bias is 2.6% for OR_2 and 12.3% for HR_2 and the CI coverage probability is 94.4% for OR_2 and 79.4% for HR_2 .

In contrast, when the hazards functions for events 1 & 2 are proportional (scenarios(a)–(d)), the estimate of the ratio of the two HRs, estimated using the ratio between the two odds-like ratios obtained from the polytomous model, is accurate regardless of the disease rate: the relative bias is within 3%, similar to that found in the joint Cox model (see below). This is consistent with the results from the mathematical approach (Appendix I). The variance estimates from the simulation studies are also accurate so that CI coverage probability and type I error rate are adequate. Furthermore, statistical power using this approach is comparable to that of the joint Cox model. Therefore, the use of the polytomous model to compare exposure-disease subtype associations is valid as long as the hazard functions for the disease subtypes are proportional. When the two events do not have proportional hazards functions, the polytomous model results in a larger bias (relative bias 4–9% for scenarios(e)–(h)) in estimating the ratio of HRs and the bias is higher when the disease rate is higher: the relative bias is over 10 times higher than that for the joint Cox model: 8.8% vs 0.6% in scenario(h), although the power and type I error rate for testing $\beta_1 = \beta_2$ are comparable to what was obtained for the joint Cox model.

Table 3 shows that the joint Cox model provides an unbiased point and interval estimates of the true HR for each event, regardless of the event rates and the proportionality of the

hazards functions - the relative bias is small and the coverage probability is always close to 95%. The same is true for the estimation of the ratio of the HRs.

As a sensitivity analysis, we considered cases when one event is rare and the other is not. The same conclusion remains for both models as if both events are common; a less common exposure, i.e., $P(X=1) < 0.5$ as well as a continuous X were also considered and the results remained similar (not shown).

Association of BMI with Breast Cancer subtypes in the NIH-AARP Diet and Health Study

We first applied the polytomous model and found that compared to leaner women ($BMI < 25.0 \text{ kg/m}^2$), those who were overweight ($25.0\text{--}29.9 \text{ kg/m}^2$) or obese ($\geq 30.0 \text{ kg/m}^2$) had an increased risk of ER+/PR+ breast tumors (OR = 1.59 and 2.10, respectively) with a statistically significant linear trend (p for trend < 0.0001 , Table 4). In contrast, overweight (OR = 1.42) but not obesity was associated with a higher risk of ER-/PR- breast tumors without a significant linear trend ($p = 0.28$). BMI was not associated with risk of ER+/PR- breast cancer. Since all the breast cancer subtypes are rare, the odds-like ratio provides a reasonable estimate of the OR and the HR. We then applied the joint Cox model that also demonstrated that both overweight (HR = 1.55, 95% CI: 1.30, 1.85) and obese (HR = 1.76, 95% CI: 1.47, 2.10) are significantly associated with ER+/PR+ cancer and a significant linear trend over BMI groups only among ER+/PR+ cancers ($p < 0.0001$) but not for other types (ER-/PR-, p for trend = 0.69 for ER-/PR- and = 0.60 ER-/PR+) (Table 4). Among the ER-/PR- breast tumors, overweight (HR = 1.43) but not obesity was associated with a higher risk of this subtype as in the polytomous model. Although both the joint Cox model and the polytomous model allow for the other covariates in the model to have the same effects across the different disease subtypes, implementation of such polytomous models is not as straightforward as the joint Cox model using standard procedures in SAS. Therefore in this analysis the joint Cox model but not the polytomous model assumed other covariates to have the same effects across the disease subtypes. Consequently, despite the small number of cases, the joint Cox model was able to yield estimates of HRs for BMI associated with ER-/PR+ breast cancer, albeit with wide CI, while the polytomous model failed to converge and produce estimates of ORs. However, one should be cautious about the interpretation of the result for the ER-/PR+ category because of its limited size.

We next formally evaluated whether the magnitude of the association of BMI with disease differed across subtypes. Since the validity of the polytomous model for this analysis depends on the proportional hazards assumption, we examined if the four subtypes have proportional baseline hazards. Graphical approaches as well as statistical tests can be used to assess proportionality. However, the statistical tests are influenced by sample size and cannot provide insights on how the data deviate from proportionality. Therefore, we used a graphical approach based the scaled Schoenfeld residuals (21): non-proportionality of baseline hazards would be suggested by a systematic decreasing or increasing trend in a lowess smoothed curve over schoenfeld residuals. Figure 1 indicates a decreasing trend between ER-/PR- and ER+/PR+ and Figure 2 suggests an increasing trend between ER-/PR- and ER+/PR+. Because the baseline hazards function of ER+/PR+ appeared to be non-proportional to those of the ER-/PR- and ER+/PR-, the joint Cox model was used to estimate the ratio of the HRs: the hazard ratio associated with obesity for ER-/PR- is significantly lower than that for ER+/PR+ (relative HR = 0.59, 95% CI: 0.40, 0.88); the relative HR associated with each increase of categories of BMI is 0.78 (0.65, 0.94), i.e., the linear trend for ER+/PR+ is 22% stronger than that for ER-/PR-.

4. Discussion

Polytomous models are currently the most commonly used method to compare exposure-disease associations across multiple subtypes of disease; however, the model has several potential limitations when applied to prospective epidemiological studies. We have provided evidence that supports our recommendation that the joint Cox regression model should be used for time-to-event data. This approach has been under-utilized for the comparison of exposure effects on different subtypes of disease in epidemiologic studies because lack of awareness of potential problems associated with the polytomous models and lack of familiarity of the data augmentation and implementation of the joint Cox model using standard statistical software. We demonstrated through mathematical and extensive simulation studies that both methods perform well in comparing exposure and disease subtype associations under the proportionality assumption. However, when the proportionality in hazards functions across disease subtypes is violated, polytomous models can yield biased estimate of differential effects, particularly when at least one event rate exceeds 10%. Joint Cox regression models, in contrast, are a valid approach regardless of disease rates and whether or not proportionality across subtypes satisfies. Since the proportional hazards assumption can often be of question, the joint Cox model provides a more robust solution for examining heterogeneity in exposure-cancer subtype associations. Joint Cox regression models are also easy to implement with commonly used statistical software (see Appendix III for codes in SAS and Splus or R). We note that when adjusting for other covariates or potential confounders in a polytomous model, standard procedures in SAS typically assume that the covariates have differential effects on different subtypes of disease while MultinomRob (22) in R and mlogit in STATA allow some or all of the covariates to be constrained to have the same effect across disease subtypes. In contrast, a joint Cox model can be easily implemented using standard procedures in commonly used statistical software, regardless of whether the effects of the covariates are assumed to be the same or different across disease subtypes.

Other alternative approaches to the polytomous models have also been used. For example, Prentice, et al. used a stratified Cox regression model to compare the effects of a dietary intervention on multiple subtypes of breast cancer in which independence across competing event types was assumed (23–24). However, a woman with multiple risk factors might be more likely to develop any subtype of breast cancer; therefore, risk of these two subtypes could be positively correlated. Consequently, the estimated coefficients for the same covariate could be correlated as well. Failure to take into account these correlations can over-estimate the variation of the difference in coefficients, resulting in a loss of power for detecting heterogeneity in disease associations.

When a rare disease such as cancer is considered, a more efficient prospective study design such as a nested case-control study or a case-cohort study is often used. The extension of the joint Cox model to these study designs for the analysis of disease with multiple subtypes would be of great interest for future research.

Although not our main focus, this study also served to further show that BMI has a significant positive association with ER+/PR+ breast cancer and the strength of the association for ER+/PR+ breast tumors is significantly greater than that for other breast cancer subtypes. The current study represents an update of the prior data from the NIH-AARP study and includes four times more breast cancer cases. Because the rare disease assumption applied in this study and the departure from proportionality of the baseline hazards was not substantial, results from the polytomous models did not differ materially from that of the joint Cox model.

In conclusion, to our knowledge, this paper is the first to formally evaluate and compare the polytomous models and the joint Cox models within a prospective setting. We hope that the result of this study encourage greater use of the joint Cox model for the analysis of cancer subtypes in prospective cohort studies.

Acknowledgments

The authors thank the referees for their excellent comments which helped greatly improve the paper.

References

1. Schroeder JC, Conway K, Li Y, Mistry K, Bell DA, Taylor JA. P53 mutations in bladder cancer: evidence for exogenous versus endogenous risk factors. *Cancer Research*. 2003; 63:7530–7538. [PubMed: 14612556]
2. Hosmer, DW.; Lemeshow, S. *Applied Logistic Regression*. 2. New York: John Wiley & Sons, INC; 2000. p. 273
3. Rosenblatt K, Daling JR, Chen C, Sherman KJ, Schwartz S. Marijuana use and risk of oral squamous cell carcinoma. *Cancer Research*. 2004; 64:4049–4054. [PubMed: 15173020]
4. Kelemen LE, Cerhan JR, Lim U, Davis S, Cozen W, Schenk M, et al. Vegetables, fruit and antioxidant-related nutrients and risk of non-Hodgkin lymphoma: a National Cancer Institute-Surveillance, Epidemiology, and End Results population-based case-control study. *Am J Clin Nutr*. 2006; 83:1401–10. [PubMed: 16762953]
5. Agalliu I, Gern R, Leanza S, Burk RD. Associations of high-grade prostate cancer with BRCA1 and BRCA2 founder mutations. *Clin Cancer Res*. 2009; 15:1112–1120. [PubMed: 19188187]
6. Blackmore KM, Lesosky M, Barnett H, Raboud JM, Vieth R, Knight JA. Vitamin D from dietary intake and sunlight exposure and the risk of hormone-receptor-defined breast cancer. *Am J Epidemiol*. 2008; 168:915–924. [PubMed: 18756015]
7. Morris AP, Lindgren CM, Zeggini E, Timpson NJ, Frayling TM, Hattersley AT, et al. A powerful approach to sub-phenotype analysis in population-based genetic association studies. *Genet Epidemiol*. 2010; 34(4):335–343. [PubMed: 20039379]
8. Potter J, Cerhan J, Sellers TA, McGovern PG, Drinkard C, Kushi LR, et al. Progesterone and estrogen receptors and mammary neoplasia in the Iowa Women's Health Study: How many kinds of breast cancer are there? *Cancer Epidemiol Biomarkers & Prev*. 1995; 4:319–326.
9. Gapstur SM, Morrow M, Sellers TA. Hormone replacement therapy and risk of breast cancer with a favorable histology: results of the Iowa Women's health study. *JAMA*. 1999; 281:2091–2097. [PubMed: 10367819]
10. Ziv E, Tice J, Smith-Bindman R, Shepherd J, Cummings S, Kerlikowske K. Mammographic density and estrogen receptor status of breast cancer. *Cancer Epidemiol Biomarkers & Prev*. 2004; 13:2090–2095.
11. MacInnis RJ, English DR, Gertig DM, Hopper JL, Giles GG. Body size and composition and prostate cancer risk. *Cancer Epidemiol, Biomarkers & Prev*. 2003; 12:1417–1421.
12. Cui Y, Shikany JM, Liu S, Shagufra Y, Rohan T. Selected antioxidants and risk of hormone receptor-defined invasive breast cancers among postmenopausal women in the Women's Health Initiative Observational Study. *Am J Clin Nutr*. 2008; 87:1009–18. [PubMed: 18400726]
13. Kenfield SA, Wei EK, Stampfer MJ, Rosner RA, Colditz GA. Comparison of aspects of smoking among the four histological types of lung cancer. *Tobacco Control*. 2008; 17:198–204. [PubMed: 18390646]
14. Lunn M, McNeil D. Applying Cox regression to competing risk. *Biometrics*. 1995; 51:524–532. [PubMed: 7662841]
15. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. *Statist Med*. 2007; 26:2389–2430.
16. Glynn RJ, Rosner B. Comparison of risk factors for the competing risks of coronary heart disease, stroke and venous thromboembolism. *Am J Epidemiol*. 2005; 162:975–982. [PubMed: 16207808]

17. Chatterjee N, Sinha S, Diver R, Feigelson H. Analysis of cohort studies with multivariate, partially observed, disease classification data. *Biometrika*. 2010; 97:683–698. [PubMed: 22822252]
18. Ahn J, Schatzkin A, Lacey JV, Albanes D, Ballard-Barbash R, Adams KF, et al. Adiposity, adult weight change and postmenopausal breast cancer risk. *Arch Intern Med*. 2007; 167(19):2091–2102. [PubMed: 17954804]
19. Chatterjee N. A two stage regression model for epidemiological studies with multivariate disease classification data. *JASA*. 2004; 99(465):127–138.
20. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *JASA*. 1989; 84:1065–1073.
21. Schoenfeld D. Partial residuals for the Proportional hazards regression model. *Biometrika*. 1982; 69:239–241.
22. Mebane, WR., Jr; Sekhon, JS. 2012 Sep 29. Version 1.8-6. <http://cran.r-project.org/web/packages/multinomRob/index.html>
23. Prentice RL, Caan B, Chlebowski RT, Patterson R, Kuller LH, Ockene JK, et al. Low-fat dietary pattern and risk of invasive breast cancer: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA*. 2006; 295(6):629–42. [PubMed: 16467232]
24. Kalbfleisch, JD.; Prentice, RL. *The Statistical Analysis of Failure Time Data*. 2. New York, NY: John Wiley & Sons Inc; 2002. p. 255

APPENDIX I

Proof of the validity of a polytomous model to compare exposure-disease associations.

Without loss of generality, consider a disease with only two subtypes, denoted as event 1 and event 2, which are competing risks for one another. The hazards functions for times to event 1 and 2, T_1 and T_2 , are $\lambda_1(t)$ and $\lambda_2(t)$ respectively. Assume that the hazard for each type of event is affected by a set of common risk factors, X , according to the following Cox proportional hazards model:

$$\lambda_k(t) = \lambda_{0k}(t) \exp(x\beta_k) \quad (1)$$

where $\lambda_{0k}(t)$ is the baseline hazard function for the k^{th} event type and β_k denotes the log hazard ratio associated with a unit increase in X for the development of the k^{th} event type ($k=1, 2$). Let D denote a polytomous outcome where $D=0$ if the subject remains free of disease during follow up; $D=k$ if event k occurs for $k=1, 2$. The polytomous logistic regression model, also known as the multinomial logistic regression model, can be expressed as:

$$\log \frac{P(D=k|X)}{P(D=0|X)} = \alpha_k + \gamma_k X. \quad (2)$$

Note that

$$P(D=1|X=x) = P(T_1 < \tau, T_1 < T_2 | X=x)$$

and

$$P(D=2|X=x) = P(T_2 < \tau, T_2 < T_1 | X=x)$$

and

$$P(D=0|X=x)=P(T_1>\tau, T_2>\tau|X=x).$$

Assume the baseline hazard functions for event 1 and event 2 are proportional to each other, i.e., $\lambda_{02}(t)=\delta\lambda_{01}(t)$. Then we have

$$P(D=1|X=x)=\frac{e^{\beta_1 x}}{e^{\beta_1 x}+\delta e^{\beta_2 x}}(1-e^{-(e^{\beta_1 x}+\delta e^{\beta_2 x})\Lambda_{01}(\tau)})$$

and

$$P(D=2|X=x)=\frac{\delta e^{\beta_2 x}}{e^{\beta_1 x}+\delta e^{\beta_2 x}}(1-e^{-(e^{\beta_1 x}+\delta e^{\beta_2 x})\Lambda_{01}(\tau)})$$

where $\Lambda_{01}(\tau)$ and $\Lambda_{02}(\tau)$ are the cumulative hazards functions by time τ for event 1 and 2, respectively, $\Lambda_{02}(\tau)=\delta\Lambda_{01}(\tau)$ and

$$P(D=0|X=x)=e^{-(e^{\beta_1 x}+\delta e^{\beta_2 x})\Lambda_{01}(\tau)}.$$

$$\text{Thus } e^{\gamma_1}=\frac{(1+\delta)e^{\beta_1}}{e^{\beta_1}+\delta e^{\beta_2}} * \frac{e^{(e^{\beta_1}+\delta e^{\beta_2})\Lambda_{01}(\tau)}-1}{e^{(1+\delta)\Lambda_{01}(\tau)}-1}$$

And $e^{\gamma_2}=\frac{(1+\delta)e^{\beta_2}}{e^{\beta_1}+\delta e^{\beta_2}} * \frac{e^{(e^{\beta_1}+\delta e^{\beta_2})\Lambda_{01}(\tau)}-1}{e^{(1+\delta)\Lambda_{01}(\tau)}-1}$. Therefore, $\gamma_1 = \beta_1$ and $\gamma_2 = \beta_2$, however, $e^{\gamma_1-\gamma_2}=e^{\beta_1-\beta_2}$. That is, a test of $\gamma_1 = \gamma_2$ is equivalent to a test of $\beta_1=\beta_2$.

When both event rates are rare, $\Lambda_{01}(\tau)$ and $\Lambda_{02}(\tau)$ are small, so that $e^{\gamma_1} \approx e^{\beta_1}$ and $e^{\gamma_2} \approx e^{\beta_2}$. Thus only under the rare event assumption we can use γ_1 to estimate β_1 and γ_2 to estimate β_2 .

In the presence of independent censoring (for example, censoring due to early loss to follow-up), we denote the censoring time by C so that

$$P(D=1|X=x)=P(T_1<\tau, T_1<T_2, T_1<C|X=x)$$

and

$$P(D=2|X=x)=P(T_2<\tau, T_2<T_1, T_2<C|X=x).$$

Then,

$$\begin{aligned}
 P(D=1|X=x) &= \int_0^\tau P(T_2 > u|X=x)P(C > u|X=x)\lambda_1(u)e^{-\Lambda_1(u)}du \\
 &= \int_0^\tau \lambda_1(u)e^{-\Lambda_1(u)-\Lambda_2(u)}P(C > u|X=x)du
 \end{aligned}$$

and

$$\begin{aligned}
 P(D=2|X=x) &= \int_0^\tau P(T_2 > u|X=x)P(C > u|X=x)\lambda_2(u)e^{-\Lambda_2(u)}du \\
 &= \int_0^\tau \lambda_2(u)e^{-\Lambda_1(u)-\Lambda_2(u)}P(C > u|X=x)du \\
 &= \delta e^{(\beta_2-\beta_1)x} \int_0^\tau \lambda_1(u)e^{-\Lambda_1(u)-\Lambda_2(u)}P(C > u|X=x)du
 \end{aligned}$$

so that

$$\begin{aligned}
 e^{\gamma_1-\gamma_2} &= \frac{P(D=1|X=1)/P(D=1|X=0)}{P(D=2|X=1)/P(D=2|X=0)} \\
 &= \frac{\int_0^\tau \lambda_1(u)e^{-\Lambda_1(u)-\Lambda_2(u)}P(C > u|X=1)du / \int_0^\tau \lambda_{01}(u)e^{-\Lambda_{01}(u)-\Lambda_{02}(u)}P(C > u|X=0)du}{\delta e^{(\beta_2-\beta_1)x} \int_0^\tau \lambda_1(u)e^{-\Lambda_1(u)-\Lambda_2(u)}P(C > u|X=1)du / \int_0^\tau \lambda_{01}(u)e^{-\Lambda_{01}(u)-\Lambda_{02}(u)}P(C > u|X=0)du} \\
 &= e^{\beta_1-\beta_2}.
 \end{aligned}$$

Therefore, a polytomous model can still provide a valid estimate of ratio of hazard ratios between event 1 and event 2 in the presence of early loss to follow-up as long as both events have hazards functions which are proportional to each other.

APPENDIX II

Below we describe a method to simulate an independent exponential censoring time.

Without censoring due to early loss of follow-up, the probability of no events by the end of follow-up τ is $P(D=0|x) = e^{-\Lambda_1(\tau)-\Lambda_2(\tau)}$, where $\Lambda_1(\tau)$ and $\Lambda_2(\tau)$ are the cumulative hazards functions up to time τ for event 1 and 2, respectively. With the presence of censoring due to early loss of follow-up, the same probability becomes

$$P(D=0|x, C) = \frac{\Lambda_c(\tau)}{\Lambda_1(\tau) + \Lambda_2(\tau) + \Lambda_c(\tau)} (1 - e^{-\Lambda_1(\tau) - \Lambda_2(\tau) - \Lambda_c(\tau)})$$

where $\Lambda_c(\tau)$ is the cumulative hazards function for an independent censoring time. We limit the increase in probability of no event with early loss of follow-up to be not more than 5% then the same probability without early loss of follow-up, which implies

$$\frac{P(D=0|x, C)}{P(D=0|x)} = \frac{\Lambda_c(\tau)}{\Lambda_1(\tau) + \Lambda_2(\tau) + \Lambda_c(\tau)} (e^{\Lambda_1(\tau) + \Lambda_2(\tau)} - e^{-\Lambda_c(\tau)}) < 1.05.$$

For exponential times to event 1 and 2 and time to censoring, the above equation reduces to

$$\frac{P(D=0|x, C)}{P(D=0|x)} = \frac{\lambda_c}{\lambda_1 + \lambda_2 + \lambda_c} (e^{\lambda_1 \tau + \lambda_2 \tau} - e^{-\lambda_c \tau}) \approx 1.05$$

where we numerically solve for λ_c with a given λ_1 and λ_2 based on the event rates we assumed for each scenario; for exponential time to event 1 and Weibull for time to event 2, the rate is calculated similarly.

APPENDIX III

Data Structure and Program Code

Definition of the Joint Survival Outcome under Three Scenarios

Type of events	Event Type	(T,C)
Developed event 1 at t	1	(t, 1)
	2	(t, 0)
Developed event 2 at t	1	(t, 0)
	2	(t, 1)
No event by the end of follow-up τ	1	(τ , 0)
	2	(τ , 0)

Therefore, the time to event data should be reorganized as follows:

Subject	Time	Status	Event 2	Exposure
1	5	1	0	0
1	5	0	1	0
2	10	0	0	1
2	10	0	1	1
3	6	0	0	0
3	6	1	1	0
4	...			

Note: in the above table, subject 1 who was not exposed developed event 1 at time point 5; subject 2 who was exposed did not develop any event after 10 units of follow-up; subject 3 who was not exposed developed event 2 at time point 6.

SAS code for the Joint Cox Model:

```
proc phregcovs(aggregate);
model time*status(0)=exposure expotype;
strata event2;
```

```
id subject;
```

```
expotype=exposure*event2;
```

```
run;
```

R orSplus code for the joint Cox model:

```
coxph(Surv(time, status)~exposure*strata(event2)+cluster(subject), robust=T)
```

Figure 1.
Residual plot for ER-/PR- breast cancer compared to ER+/PR+ cancer

Figure 2.
Residual plot for ER+/PR- breast cancer compared to ER+/PR+ cancer

Table 1

Parameters chosen for different scenarios considered in the simulation study

Scenario	Event 1		Event 2				$e^{\beta_1-\beta_2}$	
	$P(T_1 < 10)$		OR_1		e^{β_1}		OR_2	
	X=0	X=1	X=0	X=1	X=0	X=1	X=0	X=1
(a)	0.030	0.075	2.62	2.56	0.030	0.030	1.00	2.56
(b)	0.010	0.020	2.02	2.01	0.015	0.030	2.03	2.01
(c)	0.150	0.225	1.65	1.57	0.150	0.150	1.00	1.57
(d)	0.150	0.195	1.37	1.33	0.202	0.260	1.39	1.33
(e)	0.020	0.060	3.13	3.06	0.030	0.030	1.00	3.06
(f)	0.030	0.060	2.06	2.03	0.040	0.080	2.08	2.03
(g)	0.100	0.200	2.25	2.12	0.300	0.300	1.00	2.12
(h)	0.200	0.280	1.56	1.47	0.300	0.408	1.61	1.47

Note: The odds ratio for event k, i.e., OR_k is defined as the ratio of odds of developing event k for the exposed ($X=1$) and the odds of event k for the unexposed ($X=0$), $k=1, 2$; the hazard ratio for event k, i.e., β_k is defined as the hazard ratio for developing event k associated with exposure; scenarios (a)–(d) considered events 1 and 2 have proportional hazards functions and scenarios (e)–(h) considered events 1 and 2 do not have proportional hazards functions.

Table 2

Performance of the polytomous model estimates ($OR_2^*=e^{\gamma_2}$ and $OR_1^*=e^{\gamma_1-\gamma_2}$ in model (2)) of OR_2 and e^{β_2} and $e^{\beta_1-\beta_2}$

Scenario	Event 2 ¹				Comparison			
	OR_2		e^{β_2}		$e^{\beta_1-\beta_2}$			
	Bias ²	Cov	Bias	Cov	Power/Type I error	Bias	Cov	Power/Type I error
(a)	1.5	95.6	1.5	95.6	4.4	2.4	94.8	78.8
(b)	0.6	95.8	1.3	95.6	59.6	2.0	96.2	3.8
(c)	4.6	94.2	4.6	94.2	5.8	-0.1	97.8	73.2
(d)	2.2	94.6	6.3	92.2	86.4	-0.0	97.0	3.0
(e)	0.1	96.6	0.1	96.6	3.4	5.3	95.4	87.8
(f)	-0.4	95.2	1.8	96.4	94.2	5.6	95.4	4.6
(g)	2.9	94.0	2.9	94.0	6.0	4.3	96.2	99.6
(h)	2.6	94.4	12.3	79.4	99.8	8.8	96.4	3.6

Note:

¹ Scenarios considered for event 1 are also considered for event 2, that is, both events include the scenarios of positive associations with the exposure and null associations with the exposure. For the same scenario, the results for event 2 are similar to those for event 1. Therefore, for simplicity, in Tables 2–3 we omit the result for event 1 and only report for results on parameters associated with event 2 and parameters associated with the contrast between events 1 and 2;

² Bias is defined as (estimated value-parameter of interest)/parameter of interest * 100; coverage is the percentage of times the confidence interval included the parameter of interest; power is the percentage of times the confidence interval did not include 1 when the alternative is true and type I error rate is the same percentage but when the null is true.

Table 3

Performance of the joint Cox model estimate (model (3)) of β_2 and $\beta_1 - \beta_2$

β_2				$\beta_1 - \beta_2$			
bias	cov	Power/Type I error		bias	cov	Power/Type I error	
(a)	0.9	95.2	4.8	2.4	94.2		80.4
(b)	0.0	95.8	59.2	1.9	96.0		4.0
(c)	-0.4	95.4	4.6	-0.1	95.4		77.6
(d)	-0.6	94.6	82.0	-0.0	94.4		5.6
(e)	0.9	96.6	3.4	3.8	95.5		87.8
(f)	-1.1	95.2	93.8	3.5	95.6		4.4
(g)	-0.1	94.8	5.2	0.7	94.4		99.8
(h)	-0.2	95.4	99.8	0.6	95.4		4.6

Note: Bias, coverage and power/type I error were defined the same as in Table 2.

Table 4

Polytomous model and the Joint Cox Model^{*} on the AARP data to evaluate current BMI with four subtypes of breast cancer defined by the tumor's ER and PR status

Polytomous model	BMI <25.0	BMI 25.0–29.9	Relative OR(CI)	BMI 30.0	Relative OR(CI)	p-value for trend	p-value Difference in trend
ER+/PR+ Case No. OR(CI)	286 1(ref)	345 1.589(1.333,1.894) ⁺	1(ref)	366 2.100(1.753,2.516) ⁺	1(ref)	<0.0001	Ref
ER-/PR- Case No. OR(CI)	84 1(ref)	91 1.424(1.025,1.979) ⁺	0.901 (0.617,1.316)	71 1.192(0.824,1.725)	0.577 (0.394,0.844) ⁺	0.283	0.012
ER-/PR+ Case No. OR(CI)	9 1(ref)	4		5			
ER+/PR- Case No. OR(CI)	97 1(ref)	73 0.885(0.635,1.233)	0.585 (0.385,0.890)	61 0.902(0.626,1.299)	0.433 (0.286,0.654) ⁺	0.531	<0.0001
Joint Cox model	Relative HR(CI)		Relative HR(CI)		Relative HR & CI for trend		
ER+/PR+	1(ref)	1.551(1.304,1.845) ⁺	1(ref)	1.758(1.471,2.101)	1(ref)	<0.0001	Ref
ER-/PR-	1(ref)	1.434(1.034,1.990) ⁺	0.925 (0.639,1.338)	1.043(0.729,1.493)	0.593 (0.399,0.883) ⁺	0.693	0.783 (0.654,0.938) ⁺
ER-/PR+	1(ref)	0.531(0.139,2.021)	0.342 (0.089,1.316)	0.753(0.224,2.531)	0.428 (0.126,1.456)	0.599	0.631 (0.317,1.257)
ER+/PR-	1(ref)	0.861(0.620,1.194)	0.555 (0.384,0.802) ⁺	0.756(0.532,1.075)	0.430 (0.291,0.636) ⁺	0.113	0.659 (0.544,0.798) ⁺

Note

^{*} model adjusted for age, age at menarche, age at first live birth, parity, smoking, educational level, race, family history of breast cancer, fat intake, alcohol consumption, oophorectomy, physical activity and height (in polytomous model, all the adjusting variables have different coefficients across subtypes; in the joint Cox model, those adjusting variables with the similar coefficient across subtypes are set to have the same coefficient);

⁺ indicates p-value<0.05.