Published in final edited form as:

Ophthalmic Epidemiol. 2009 ; 16(2): 98–106. doi:10.1080/09286580902737532.

Evaluation of Risk Factors for Cataract Types in a Competing Risks Framework

Robert J. Glynn, ScD, Bernard Rosner, PhD, and William G. Christen, ScD

Division of Preventive Medicine and Channing Laboratory, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA

Abstract

Purpose—In prospective studies of incident cataract, a person may have developed two or all three of cortical, nuclear sclerotic and posterior subcapsular cataract at the time when clinically significant cataract is apparent. One can compare the impact of risk factors on different types of cataract with methods of competing risk survival analysis that account for tied events.

Methods—We studied risk factors for incident cataract among 20,599 participants in the Physicians' Health Study; 1,923 of whom developed cataract that reduced visual acuity over a median follow-up of 13.0 years. Among incident cases, 45% had two types of cataract and an additional 12% had all three. Proportional hazards models, adjusted for tied failures through use of generalized estimating equations, provided comparisons across different cataract types. Stratified analyses by type with age as the time scale were used to accommodate the strong but different relationships of age with occurrence of different cataract types.

Results—Analyses found different relationships of diabetes and body mass index, but similar relationships of smoking, family history of coronary disease and multivitamin use, across different cataract types. A simplified model provided common estimates for relative risks associated with those characteristics with similar effects across cataract types.

Conclusions—Whereas polytomous logistic regression has proven useful for comparisons of risks across cataract types in retrospective and cross-sectional studies, this approach based on methods of survival analysis has greater flexibility in prospective studies with variable follow-up and frequently tied event times for different cataract types.

Recent data show that cataract remains the leading cause of blindness worldwide, and that the percent of all blindness due to age-related causes such as cataract is increasing, both because of increased longevity as well as advances in preventing blindness from infectious diseases (1). Epidemiologic studies have identified potentially modifiable risk factors for cataract (2–28), but questions remain on how associations of risk factors differ for opacities in different areas of the lens. Comparisons of associations of risk factors with different cataract types can help clarify the etiology of cataract and generate hypotheses on differences in mechanisms for different types. Studies of possible differences in associations of risk factors across types of cataract are challenging because people often develop multiple types of cataract, and these outcomes are not independent.

Previous studies of such comparisons have used a case-control design and included many subjects with multiple cataract types (2,3). Analyses used polytomous logistic regression with control subjects as the referent group, with separate case groups including cases with exactly one type of cataract, and with another group including those cases with multiple types (29,30). Two limitations of this approach are that the substantial numbers of people with multiple types of cataract do not contribute to estimation of the associations of a risk factor with each specific type, and that interpretation of measures of association with the

outcome of multiple types is ambiguous. Alternatively, some published studies have considered separate analyses of risk factors for each type (4–11). However, formal comparisons of effects of risk factors across types have not been considered, perhaps because many subjects have multiple cataract types and estimates of effects on different types are correlated..

Methods for competing risk and multivariate survival analysis provide attractive alternative approaches when information on dates of diagnosis is available (31–34). Similar methods can be used to analyze data collected under two slightly different study designs: one design independently assesses the time of development of each cataract type in each eye or person; while an alternative design identifies the date of diagnosis of clinically significant cataract in a person and then ascertains the types present at that time. We illustrate analyses under the second study design and consider modifications for the first design in our discussion. In the second design, subjects with multiple types of cataract present when clinically significant cataract is diagnosed can contribute information to estimation of associations with each type. One can use methods of generalized estimating equations to account for the correlations among these failures of different types. We illustrate this approach with a comparison of the associations of risk factors with the development of cataract types among over 20,000 US male physicians followed for a median of 13 years in the Physicians' Health Study.

Materials and Methods

Study population

The Physicians' Health Study was a randomized, double-blind, placebo-controlled trial of aspirin and beta carotene in the primary prevention of cardiovascular disease and cancer. The trial randomized 22,071 US male physicians, aged 40–84 years in 1982, who had no history of cancer, myocardial infarction, stroke or transient cerebral ischemia. The aspirin component of the trial was terminated in 1988, but the beta-carotene component continued to its scheduled completion in December 1995 (35,36). The trial protocol was reviewed and approved by the Institutional Review Board of the Brigham and Women's Hospital, and all participants provided written informed consent.

Exposure and outcome measures

The Physicians' Health Study assessed risk factors and identified potential outcomes through questionnaires sent to participants before randomization and annually during follow-up. The risk factors for cataract we considered were those measured at baseline that were previously found to be related to the overall risk of cataract in this population (24–28). Specifically, men reported their date of birth, weight in pounds and height in inches (from which we calculated body mass index in kg/m^2), their current regular use of multi-vitamin supplements, current and former cigarette smoking and number of cigarettes smoked in current smokers, a history of diabetes, and a history of myocardial infarction in either parent before 60 years of age.

After randomization, annual follow-up questionnaires assessed adherence to treatments and study outcomes, including cataract. Through the use of three mailings supplemented by telephone contact of nonrespondents, response rates at each survey were greater than 98 percent. At the end of the trial, 99.2 percent of participants provided information on morbidity. When a participant reported an incident cataract, a written consent identifying the treating ophthalmologist or optometrist was requested. Eye doctors were contacted by mail and asked to supply information by questionnaire about the presence of lens opacities, visual acuity loss, cataract extraction, other ocular abnormalities that could explain visual acuity loss, cataract type (nuclear sclerotic, cortical, and/or posterior subcapsular), and cause

(including age-related, traumatic, congenital, inflammatory, or surgery- or steroid-induced). Alternatively, the doctor could send photocopies of relevant medical records. Reports from eye doctors or medical records were obtained from over 92 percent of participants reporting cataract. As for all other outcomes in this study, if records for confirmation were unavailable, the participant was assumed to be event-free as long as he remained alive until the end of the trial. Figure 1 illustrates the process of cataract evaluation in the study.

The study outcome was the incidence of cataract and the types of cataract that contributed to its diagnosis. Cataract was defined as a self-report confirmed by medical record review to be initially diagnosed after randomization, age-related in origin (congenital cataracts and those due to trauma, steroids, intraocular inflammation, or surgery were excluded), with best-corrected visual acuity of 20/30 or worse and with no alternate ocular disease to explain the visual acuity loss. Follow-up time of participants with confirmed cataract ended on the date of diagnosis. Type of cataract was determined from records after the definition of incident cataract was met, and men could have two or all three types of incident cataract present at the time of diagnosis. Among 1,977 men in this analysis with confirmed cataract, 54 (2.7%) had missing information on cataract type. These men were censored at the time of confirmed cataract.

The approach to evaluation of cataract suggested a competing risks framework with respect to cataract types. Each participant had a single follow-up time, either when he was diagnosed with clinically significant cataract or at the end of his follow-up without any significant cataract. For those developing cataract, interest focused on the types of cataract that contributed to the visual loss. Types of cataract were not independently evaluated unless clinically significant cataract was noted. An alternative design would have separately evaluated the time to development of each cataract type in each eye of a study subject, giving rise to a multivariate survival study. In this design, eyes would be censored at the time of cataract extraction. In the discussion, we describe the similarities in analytic strategies applicable to this alternative design, compared to the competing risk approach for the current design.

Statistical analysis

Individuals, rather than eyes were the unit of analysis because eyes were not examined independently. Participants were classified according to the status of the worse eye as defined by an earlier date of diagnosis or cataract surgery. If the two eyes had the same date of diagnosis, the eye with the worse visual acuity was designated the worse eye. If one eye was excluded because of visual loss attributed to other ocular abnormality or a cataract that was not age-related, the fellow eye was used for classification.

We excluded the 1,103 men who reported a cataract at baseline, plus the 369 men with missing values on one of the baseline covariates of interest, leaving a cohort of 20,599 men for this analysis. For each participant, we calculated the time in years until the first of a confirmed diagnosis of cataract, the end of the study or death. At the time of cataract diagnosis, a man could have one, two or all three types of cataract present as contributing causes.

Initial analyses evaluated the association of age with the incidence of each of nuclear sclerotic, cortical and posterior subcapsular cataract. We considered five age groups (40–49, 50–59, 60–69, 70–79, and ≥80 years) and calculated incidence rates; subjects were allowed to contribute time to multiple groups as they aged. Interpretation of Kaplan-Meier curves in the setting of competing risks is problematic and tends to produce an overestimate of the cumulative incidence of a cause specific failure (37,38). We used the nonparametric estimate of the cause-specific failure probability at time t , $\hat{F}_k(t)$, defined as follows (37). Let

$0 < t_{(1)} < t_{(2)} < \dots < t_{(r)}$ denote the r distinct, ordered ages at the time of incident, age-related cataract. At each age $t_{(i)}$, let n_i be the number at risk of cataract, let d_i be the number developing cataract, and let d_{ki} be the number developing cataract of type k , where $k=1, 2$, or 3 indexes the three types of cataract. Since men can have more than one incident type of

cataract at time $t_{(i)}$, $\sum_k d_{ki} \geq d_i$. The Kaplan-Meier estimate of the probability of remaining

free of any cataract at time $t_{(j-1)}$ is then $\widehat{S}(t_{(j-1)}) = \prod_{i=1}^{j-1} (1 - \frac{d_i}{n_i})$, and the

cause-specific probability of a cataract of type k by age t is $\widehat{F}_k(t) = \sum_{\{j: t_{(j)} \leq t\}} \frac{d_{kj}}{n_j} \widehat{S}(t_{(j-1)})$. These conditional probabilities are taken as measures of cumulative incidence. Mortality constitutes another important competing risk, so these estimates of type specific cataract by a given age assume that a man survives to that age.

These initial analyses indicated strong but different relationships of age with the occurrence of cataract in different areas of the lens. Hence, multivariate survival analyses used age as the time scale and stratified on type of outcome (39). Specifically, the approach described by Lunn and McNeil that stratifies on event type allows for estimation of the separate associations of each risk factor with the relative hazard of each outcome under a proportional hazard assumption and, with age as the time scale, makes no assumptions about the shape of the relationship of age with the hazard of each event type (31). Through the use of data augmentation, one can implement the approach with standard statistical software packages for the proportional hazards model such as SAS or Stata. The Appendix illustrates the data augmentation and analysis steps in SAS.

The stratified proportional hazards model specifies that the hazard of developing clinically significant cataract at age t that is attributable by the diagnosing eye doctor to cataract type k ($k=1,2,3$) for the i th participant with covariate vector X_i is $\lambda_k(t) \exp(\beta' X_i)$, where $\lambda_k(t)$ is the unspecified, underlying hazard of failure due to type k at age t and β is the vector of regression parameters relating covariates to type-specific hazard. The covariate vector X_i can include some variables with an assumed common relationship across types, as well as interaction terms involving type.

The data augmentation step of the analysis requires that each subject have a separate observation for each of the three cataract types. This requires construction of a data set with three observations for each subject, each with the same follow-up time but with a separate indicator variable for the presence of each cataract type. Specifically, for men who developed clinically significant cataract, a type-specific indicator was set to one in the observation corresponding to that type if the diagnosing eye doctor noted the presence of that type, and set to zero otherwise. Our first multivariate analysis assumed a separate association of each risk factor with each type of cataract. This model included type-specific covariate effects defined to be equal to the value of the covariate in the observation corresponding to that type, and to be equal to zero in observations corresponding to other types. Alternative coding for type by covariate interactions is possible in the stratified Cox model, as described by Therneau and Grambsch (40), but our approach leads to direct interpretation of each covariate. Many men had multiple types of cataract present at the time of diagnosis, so these outcomes were not independent. We applied generalized estimating equations, as described by Lee et al., to account for the correlations among these related outcomes (33). Specifically, parameter estimates were obtained under the assumption of independence and then standard errors were adjusted by consideration of the correlations among the outcomes.

A particular hypothesis of interest was whether a risk factor had the same or a different relationship with different cataract types. We used generalized Wald tests to evaluate this hypothesis for each risk factor, allowing all other risk factors to have different effects for different cataract types (41). In the case of multiple variables measuring a single construct (e.g., former smoking, current light smoking and current heavier smoking), a single test compared the associations of all variables across cataract types. On the basis of the results of these tests, we produced a simplified model that equated those relationships of a risk factor with different outcomes when there was little evidence for differences. We used $P < 0.05$ as a criterion to keep different effects of a given risk factor on different cataract types.

We considered whether birth cohort confounded relationships of risk factors with occurrence of cataract types through stratification on categories of birth year in the proportional hazards model. Patterns of Schoenfeld residuals with time were considered to identify possible violations of the proportional hazards assumption (42). Additionally, we fitted separate models for the first 6.5 years of follow-up and thereafter, to describe any changes in associations of risk factors with cataracts diagnosed in early versus late follow-up.

Results

Associations of age with cataract types

Among the 20,599 men included in this analysis, older age was associated with current use of multivitamins, a lower prevalence of myocardial infarction in either parent before age 60, a higher prevalence of diabetes, and a greater prevalence of former cigarette smoking (Table 1).

Over the median follow-up of 13.0 years, 1,977 men (9.6 percent of study subjects) developed confirmed, incident age-related cataract. The majority of cataract cases developed more than one type of cataract (Table 2). The percent of cases with more than one type of cataract ranged across four age groups from 54 percent in men originally aged 40–49 years to 59 percent in men originally aged 70–84 years.

Cortical and posterior subcapsular cataract had comparable incidence rates that were less than half the rate of nuclear sclerotic cataract (Table 3). However, age-specific incidence rates of posterior subcapsular cataract in men initially under age 60 years were more than twice the rates of cortical cataract. After age 70, incidence of cortical cataract was much greater than that of posterior subcapsular cataract. This crossing of hazard rates led to an estimated cumulative incidence of cortical cataract that was less than that of posterior subcapsular cataract before age 78, but greater after that age (Figure 2). At age 90, the overall cumulative incidence of cataract was 0.63 whereas cumulative incidence estimates for nuclear sclerotic, cortical and posterior subcapsular cataract were 0.58, 0.27, and 0.21, respectively.

Multivariate analysis

The initial multivariate analysis of risk factors allowed for a possibly different relationship of each determinant with each type of cataract (Table 4). Estimated hazard ratios showed some heterogeneity but no evidence of qualitative differences across sites. Hazards associated with body mass index, family history of premature myocardial infarction, diabetes, and current smoking were generally greater for posterior subcapsular cataract than for other sites. Formal tests of heterogeneity, accounting for the possibility of cataract at multiple sites, found evidence for different relationships of body mass index ($P = 0.002$) and diabetes ($P = 0.003$) with different types of cataract. Specifically, body mass index had little relationship with cortical cataract and diabetes had an apparently weaker relationship with nuclear sclerotic cataract compared to other sites. Differences across types in relationships

with multivitamin use, family history of myocardial infarction, and cigarette smoking may be due to chance.

Based on these findings, we fitted a simplified model that maintained the differences across sites for body mass index and diabetes, but constrained to be equal the relationships of other variables across cataract types (Table 5). In this model, we found an estimated 13 percent reduction in the hazard of each cataract type associated with use of multivitamins (95% CI: 2%–22%), and a 40 to 60 percent increased hazard of each cataract type associated with different levels of current cigarette smoking. The estimated percent increase in hazard associated with diabetes was about three times greater for cortical and posterior subcapsular cataract, compared to the percent increase in hazard for nuclear sclerotic cataract.

A proportional hazards model stratified on both birth cohort and event type found similar estimated effects. Evaluation of the proportional hazards assumption over the follow-up period revealed some evidence for violation of this assumption. In separate analyses of early (first 6.5 years) versus later (after 6.5 years) follow-up, the reduced hazard associated with multivitamin use was somewhat greater in early follow-up (HR 0.79; 95% CI: 0.66–0.95) compared to later follow-up (HR 0.93; 95% CI: 0.80–1.07), and the increased hazard associated with smoking at least a pack per day was also somewhat greater in early follow-up (HR 1.99; 95% CI: 1.49–2.65) compared to later follow-up (HR 1.34; HR: 1.05–1.71). Attenuation of effects over time may reflect increased measurement error due to changes in these behaviors that were not updated during follow-up. Estimated relative hazards associated with diabetes, body mass index and a family history of premature myocardial infarction changed little over time (data not shown).

Discussion

Our analyses illustrate the advantages of a competing risks framework as an approach for the evaluation and comparison of relative hazards of risk factors for cataract in different areas of the lens. The approach allows for varying follow-up, accommodates different underlying hazards with age for different cataract types, includes people with multiple types at diagnosis in the evaluation of each type, yields tests for equality of effects across types, and can produce parsimonious models by equating comparable effects. With the use of data augmentation, it can also be readily implemented in available software packages including SAS, Stata and Splus.

An approach used in several previous prospective studies of cataract type is separate analysis of risk factors for each type of cataract. With this alternative approach, one could use bootstrap methods to test for differences in risk factors across types, accounting for possibly multiple types in the same person. Advantages of the approach proposed in the current paper are the ability to estimate common associations for some risk factors across types and relevant confidence intervals, and possibly more efficient tests of differences for some risk factors after equating some associations across types.

We have adopted a competing risks perspective for the analysis of cataract types, primarily because our study design called for evaluation of cataract types at the time when clinically significant cataract was present. Some prospective studies of cataract development separately evaluate the times of occurrence of each type. With this alternative design, one could use multivariate survival methods for joint evaluation of risk factors for each cataract type. A very similar analytic approach could be used with this alternative perspective, with separate observations for each cataract type, and adjustment for possibly multiple types through the use of generalized estimating equations, as described by Wei et al (34).

It is important to appreciate that the choice among alternative study designs has implications not only for risk factor evaluation but also for cumulative incidence estimates such as those in our Figure 2. Specifically, a reviewer has pointed out that cortical cataract alone may not affect visual acuity sufficiently to qualify as clinically significant based on our study definition.

Other approaches for evaluation of competing risks are also available but their applicability to this setting with multiple causes at the time of failure is not clear. For example, Beyersmann et al (43) and Putter et al (44) describe the value of subdistribution hazard functions if interest focuses on the impact of covariates on the cumulative incidence of a specific cause such as posterior subcapsular cataract. Software for implementation of these approaches is available (43,44), but applications to date have required that each failure has a unique cause. The approach used here produces readily interpretable estimates of cause-specific relative hazards.

Our results have some similarities and some differences with findings from previous studies of risk factors for specific types of cataract. Age is the dominant risk factor for cataract, and other studies have noted the sharper increase with age in incidence of nuclear cataract, relative to cortical and posterior subcapsular cataract (7,10,21). By treating age as the time scale in survival analysis, our analytic approach has the flexibility to accommodate the distinct and nonuniform changes with age in the relative hazards of specific types of cataract. By contrast, a logistic regression or proportional hazards analysis with age as a fixed covariate at baseline would not account for this changing relative hazard. A competing risk survival analysis with age as the time scale may be better able to control for the influence of age in evaluation of the relationships of other risk factors with types of cataract.

Current cigarette smoking is a consistently identified risk factor for cataract, with some evidence for heterogeneity in relationships across cataract types (5,12). Our possibly different finding of a common relationship of smoking across cataract types may be due to differences in quantification of smoking and our different definition of endpoints with the requirement of visual deficits. Specifically, some investigators have considered pack-years of cigarettes smoked (5). Also, cortical cataract is generally less likely to reduce visual acuity compared to nuclear sclerotic and posterior subcapsular cataracts. Some previous studies also had smaller sample sizes and lacked an approach for formal evaluation of differences in smoking risks across types.

Several studies support our finding of an apparently different relationship of diabetes with different types of cataract, with a weaker relationship with nuclear cataract. The Beaver Dam (6), Lens Opacities Case-Control Study (3) and the Barbados Eye Study (9) found diabetes to be associated with an increased risk of both posterior subcapsular and cortical opacities, whereas the Age-Related Eye Disease Study (4) found lower rates of nuclear opacities in those with versus without diabetes. Diabetes appears to be associated with both increased incidence as well as higher rates of progression of cortical and posterior subcapsular cataracts (6).

Relationships of body mass index with risk of cataract have varied qualitatively across studies, with reports of positive linear relationships (13,20,26), reduced risk in heavier people (2,3,14), and U-shaped associations (17). Cross-sectional studies may be problematic because cataract is a marker for poor health that may induce weight loss. Also, socioeconomic status may confound relationships of body mass index with development of cataract, as poor nutrition may increase risk of cataract. Prospective studies in well-nourished populations (US male physicians, other male health professionals, and female

nurses, 13,26) have found an increased risk of cataract associated with greater body mass index, with a particularly strong association with posterior subcapsular cataract.

Several observational studies have found reduced rates of cataract associated with use of multivitamins (3,16,18), with some suggestion of possible heterogeneity across cataract types (5). Potential confounding by determinants of the decision to use multivitamins is a concern. The randomized Linxian cataract studies found evidence that vitamin supplements may decrease the incidence of nuclear cataract in a nutritionally-deprived population (22). However, the randomized Age-Related Eye Disease Study (23) found that allocation to a high-dose formulation of vitamin C, vitamin E, and beta carotene relative to placebo was associated with a relative odds of increased opacity grade or cataract surgery of 0.97 (95% confidence interval: 0.84–1.11), with similar associations observed for specific cataract types.

The prospective design of our study precludes potential bias that arises if risk factors change after the development of cataract. Restriction of the study population to well-nourished individuals of higher socioeconomic status, who were free of cancer and cardiovascular disease at baseline restricts the generalizability of results. However, it provides control for potential risk factors for cataract that are difficult to measure precisely. Reliance on self-report to identify cataracts is another limitation. However, the population of physicians has high access to care and enhanced ability to report both risk factors and outcomes reliably.

In summary, our goal was to illustrate an approach to identify similarities and evaluate differences in the associations of risk factors with different types of cataract. Comparisons of associations across types can provide evidence for differences in etiology of different cataract types. Advantages of our approach include the ability to account for multiple simultaneous outcomes in the same person, to follow subjects for varying intervals, to find reduced models based on widely-available techniques of generalized estimating equations, and to estimate meaningful measures of relative hazards. The approach can use familiar diagnostics to identify features such as the crossing hazard for cortical and posterior subcapsular cataract. Violations of proportional hazards can be readily accommodated by stratification. All analyses can be implemented in popular software packages including SAS, S-Plus and Stata.

Acknowledgments

Supported by NIH grant EY 12269 from the National Eye Institute

References

1. World Health Organization. Magnitude and causes of visual impairment. Fact Sheet No 282. November. 2004 Accessed at www.who.int/mediacentre/factsheets/fs282/en
2. Mohan M, Sperduto RD, Angra SK, Milton RC, Mathur RL, Underwood BA, Pandya JN, Chhabra VK, Vajpayee RB. India-US case-control study of age-related cataracts. India-US Case-Control Study Group. Arch Ophthalmol. 1989; 107:670–676. [PubMed: 2818712]
3. Leske MC, Chylack LT Jr, Wu SY. The Lens Opacities Case-Control Study. Risk factors for cataract. Arch Ophthalmol. 1991; 109:244–251. [PubMed: 1993036]
4. Age-Related Eye Disease Study Group. Risk factors associated with age-related nuclear and cortical cataract: a case-control study in the Age-Related Eye Disease Study, AREDS Report No. 5. Ophthalmology. 2001; 108:1400–1408. [PubMed: 11470690]
5. Klein BEK, Klein R, Lee KE, Meuer SM. Socioeconomic and lifestyle factors and the 10-year incidence of age-related cataract. Am J Ophthalmol. 2003; 136:506–512. [PubMed: 12967805]

6. Klein BEK, Klein R, Lee KE. Diabetes, cardiovascular disease, selected cardiovascular risk factors, and the 5-year incidence of age-related cataract and progression of lens opacities: the Beaver Dam Eye Study. *Am J Ophthalmol.* 1998; 126:782–790. [PubMed: 9860001]
7. Klein BEK, Klein R, Lee KE. Incidence of age-related cataract: the Beaver Dam Eye Study. *Arch Ophthalmol.* 1998; 116:219–225. [PubMed: 9488275]
8. Leske MC, Wu SY, Nemesure B, Hennis A. Risk factors for incident nuclear opacities. *Ophthalmology.* 2002; 109:1303–1308. [PubMed: 12093655]
9. Hennis A, Wu S-Y, Nemesure B, Leske MC. Risk factors for incident cortical and posterior subcapsular lens opacities in the Barbados Eye Studies. *Arch Ophthalmol.* 2004; 122:525–530. [PubMed: 15078670]
10. Leske MC, Wu S-Y, Nemesure B, Li X, Hennis A, Connell AM. Incidence and progression of lens opacities in the Barbados Eye Studies. *Ophthalmology.* 2000; 107:1267–1273. [PubMed: 10889096]
11. Hiller R, Sperduto RD, Ederer F. Epidemiologic associations with nuclear, cortical, and posterior subcapsular cataracts. *Am J Epidemiol.* 1986; 124:916–925. [PubMed: 3776974]
12. Hiller R, Sperduto RD, Podgor MJ, Wilson PW, Feris FL III, Colton T, D'Agostino RB, Roseman MJ, Stockman ME, Milton RC. Cigarette smoking and the risk of development of lens opacities: the Framingham Study. *Arch Ophthalmol.* 1997; 115:1113–1118. [PubMed: 9298050]
13. Hiller R, Podgor MJ, Sperduto RD, Nowroozi L, Wilson PWF, D'Agostino RB, Colton T. A longitudinal study of body mass index and lens opacities: the Framingham Studies. *Ophthalmology.* 1998; 105:1244–1250. [PubMed: 9663229]
14. Nirmalan PK, Robin AJ, Katz J, Tielsch JM, Thulasiraj RD, Krishnadas R, Ramakrishnan R. Risk factors for age related cataract in a rural population of southern India: the Aravind Comprehensive Eye Study. *Br J Ophthalmol.* 2004; 88:989–994. [PubMed: 15258010]
15. Cumming RG, Mitchell P. Alcohol, smoking and cataracts: the Blue Mountains Eye Study. *Arch Ophthalmol.* 1997; 115:1296–1303. [PubMed: 9338677]
16. Kuzniarz M, Mitchell P, Cumming RG, Flood VM. Use of vitamin supplements and cataract: the Blue Mountains Eye Study. *Am J Ophthalmol.* 2001; 132:19–26. [PubMed: 11438049]
17. Kuang T-M, Tsai S-Y, Hsu W-M, Cheng C-Y, Liu J-H, Chou P. Body mass index and age-related cataract: the Shihpai Eye Study. *Arch Ophthalmol.* 2005; 123:1109–1114. [PubMed: 16087846]
18. Leske MC, Chylack LT Jr, He Q, Wu SY, Schoenfeld E, Friend J, Wolfe J. Antioxidant vitamins and nuclear opacities: the longitudinal study of cataract. *Ophthalmology.* 1998; 105:831–836. [PubMed: 9593382]
19. Hankinson SE, Willett WC, Colditz GA, Seddon JM, Rosner B, Speizer FE, Stampfer MJ. A prospective study of cigarette smoking and risk of cataract surgery in women. *JAMA.* 1992; 268:994–008. [PubMed: 1501325]
20. Weintraub JM, Willett WC, Rosner B, Colditz GA, Seddon JM, Hankinson SE. A prospective study of the relationship between body mass index and cataract extraction among women and men. *Intern J Obesity.* 2002; 26:1588–1595.
21. Mukesh BN, Le A, Dimitrov PN, Ahmed S, Taylor HR, McCarty CA. Development of cataract and associated risk factors: the Visual Impairment Project. *Arch Ophthalmol.* 2006; 124:79–85. [PubMed: 16401788]
22. Sperduto RD, Hu TS, Milton RC, Zhao JL, Everett DF, Cheng QF, Blot WH, Bing L, Taylor PR, Li JY, et al. The Linxian cataract studies: two nutrition intervention trials. *Arch Ophthalmol.* 1993; 111:1246–1253. [PubMed: 8363468]
23. Age-Related Eye Disease Study Research Group. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E and beta carotene for age-related cataract and vision loss: AREDS report no. 9. *Arch Ophthalmol.* 2001; 119:1439–1452. [PubMed: 11594943]
24. Christen WG, Manson JE, Seddon JM, Glynn RJ, Buring JE, Rosner B, Hennekens CH. A prospective study of cigarette smoking and risk of cataract. *JAMA.* 1992; 268:989–993. [PubMed: 1501324]
25. Seddon JM, Christen WG, Manson JE, LaMotte FS, Glynn RJ, Buring JE, Hennekens CH. The use of vitamin supplements and the risk of cataract among US male physicians. *Am J Public Health.* 1994; 84:788–792. [PubMed: 8179050]

26. Glynn RJ, Christen WG, Manson JE, Bernheimer J, Hennekens CH. Body mass index: an independent predictor of cataract. *Arch Ophthalmol*. 1995; 113:1131–1137. [PubMed: 7661746]
27. Schaumberg DA, Glynn RJ, Christen WG, Ajani UA, Sturmer T, Hennekens CH. A prospective study of blood pressure and risk of cataract in men. *Ann Epidemiol*. 2001; 11:104–110. [PubMed: 11164126]
28. Christen WG, Manson JE, Glynn RJ, Gaziano JM, Sperduto RD, Buring JE, Hennekens CH. A randomized trial of beta carotene and age-related cataract in US physicians. *Arch Ophthalmol*. 2003; 121:372–378. [PubMed: 12617708]
29. Marshall RJ, Chisholm EM. Hypothesis testing in the polychotomous logistic regression model with an application to detecting gastrointestinal cancer. *Stat Med*. 1985; 4:337–344. [PubMed: 4059720]
30. Glynn RJ, Rosner B. Methods to evaluate risks for composite end points and their individual components. *J Clin Epidemiol*. 2004; 57:113–122. [PubMed: 15125620]
31. Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics*. 1995; 51:524–532. [PubMed: 7662841]
32. 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]
33. Lee, EW.; Wei, LJ.; Amato, DJ. Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In: Klein, JP.; Goel, PK., editors. *Survival analysis: state of the art*. Dordrecht, the Netherlands: Kluwer Academic Publishers; 1992. p. 237–247.
34. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J Am Statist Assoc*. 1989; 84:1065–1073.
35. Belanger C, Buring JE, Cook N, Eberlein K, Goldhaber SZ, Gordon D, Hennekens CH Chairman, Mayrent SL, Peto R, Rosner B, Stampfer M, Stubblefield F, Willett W. The Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the ongoing Physicians' Health Study. *N Engl J Med*. 1989; 321:129–135. [PubMed: 2664509]
36. Hennekens CH, Buring JE, Manson JE, Stampfer M, Rosner B, Cook NR, Belanger C, LaMotte F, Gaziano JM, Ridker PM, Willett W, Peto R. Lack of effect of long-term supplementation with beta-carotene on the incidence of malignant neoplasms and cardiovascular disease. *N Eng J Med*. 1996; 334:1145–9.
37. Gaynor JJ, Feuer EJ, Tan CC, Wu DH, Little CR, Straus DJ, Clarkson BD, Brennan MF. On the use of cause-specific failure and conditional failure probabilities: examples from clinical oncology data. *J Am Statist Assoc*. 1993; 88:400–409.
38. Pepe MS, Mori M. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Stat Med*. 1993; 12:548–555.
39. Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *Am J Epidemiol*. 1997; 145:72–80. [PubMed: 8982025]
40. Therneau, TM.; Grambsch, PM. *Modeling Survival Data: Extending the Cox Model*. New York: Springer-Verlag; 2000. p. 44–47.
41. Rotnitzky A, Jewell NA. Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika*. 1990; 77:485–497.
42. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982; 69:239–241.
43. Beyersmann J, Dettlenkofer M, Bertz H, Shumacher M. A competing risks analysis of bloodstream infection after stem-cell transplantation using subdistribution hazards and cause-specific hazards. *Stat Med*. 2007; 26:5360–5369. [PubMed: 17634974]
44. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*. 2007; 26:2389–2430. [PubMed: 17031868]

Appendix: SAS program for competing risks with tied event times

*The input file phscat contains the variables id, asa (1=randomized to active aspirin, 0=otherwise), bc (1=randomized to active beta carotene, 0=otherwise), agernd (age at randomization), agefinal (age at cataract diagnosis or final follow-up), bmi (body mass

index in kg/m²), pastsmok (1=former smoker, 0=otherwise), smoklt20 (1=current smoker of <20 cigarettes/day, 0=otherwise), smokge20, (1=current smoker of ≥ 20 cigarettes/day, 0=otherwise), dmhx (1=history of diabetes, 0=otherwise), fmhxmi (1=family history of early myocardial infarction, 0=otherwise), multvit (1=current regular use of multivitamins, 0=otherwise), dx_ns (1=diagnosis of nuclear sclerotic cataract, 0=otherwise), dx_crt (1=diagnosis of cortical cataract, 0=otherwise), dx_psc (1=diagnosis of posterior subcapsular cataract, 0=otherwise);

*first create the augmented dataset with 3 observations per person;

*for this example, we assume common effects of family history of MI, smoking and multivitamin use as in table 4;

*variables with common effects appear in each of the 3 observations per person;

*variables with unique effects on one type are set to 0 within strata for other types;

data augmented; set phscat;

stratum=1;

if dx_ns=1 then fail=1; else fail=0;

nasp=asa; nbc=bc; nbmi=bmi; ndm=dmhx;

casp=0; cbc=0; cbmi=0; cdm=0; pasp=0; pbc=0; pbmi=0; pdm=0;

output;

stratum=2;

if dx_crt=1 then fail=1; else fail=0;

nasp=0; nbc=0; nbmi=0; ndm=0; nmihx=0;

casp=asa; cbc=bc; cbmi=bmi; cdm=dmhx;

pasp=0; pbc=0; pbmi=0; pdm=0;

output;

stratum=3;

if dx_psc=1 then fail=1; else fail=0;

nasp=0; nbc=0; nbmi=0; ndm=0; casp=0; cbc=0; cbmi=0; cdm=0;

pasp=asa; pbc=bc; pbmi=bmi; pdm=dmhx;

output;

*now fit the proportional hazards model using the covsandwich option in SAS version 9 to account for multiple events per person;

proc phreg multipass data=augmented covsandwich(aggregate);

model (agernd, agefinal)*fail(0) = nasp nbc nbmi ndm casp cbc cbmi cdm

pasp pbc pdm multvit fmhxm1 pastsmok smklt20 smkge20/rl;
strata stratum;
id id;

Figure 1.
Physicians' Health Study Flowchart for Cataract Evaluation

Figure 2.
Cumulative incidence of cataract types in the Physicians' Health Study, 1982–1995

Table 1

Baseline characteristics, by age at study entry, in the Physicians' Health Study (US male physicians, 1982)

Age group (years)	40–49	50–59	60–69	70–84
N	8,842	7,137	3,610	1,010
Body mass index (kg/m ² , mean±SD)	24.6 ± 2.8	25.0 ± 2.8	24.9 ± 2.6	24.6 ± 2.7
Multivitamin use (%)	15.6	19.7	24.8	31.3
Parental history of MI* (%)	15.3	13.2	10.4	6.9
Reported diabetes mellitus (%)	0.9	2.2	3.9	6.8
Former cigarette smoker	33.3	41.8	45.7	48.0
Current smoker, <1 pack/day	3.9	4.0	3.2	4.1
Current smoker, ≥ 1 pack/day	6.7	7.7	7.2	4.0

* Myocardial infarction in either parent before age 60

Table 2

Distribution of cataract types at diagnosis during follow-up (median 13 years) in the Physicians' Health Study, by baseline age

Age group (years)	40-49	50-59	60-69	70-84
N	8,842	7,137	3,610	1,010
Any cataract*, N (%)	140 (1.6)	579 (8.1)	874 (24.2)	384 (38.0)
Nuclear sclerotic only, N (%)	38 (0.4)	182 (2.6)	314 (8.7)	134 (13.3)
Cortical only, N (%)	2 (0.02)	13 (0.2)	29 (0.8)	9 (0.9)
Posterior subcapsular only, N (%)	23 (0.3)	45 (0.6)	32 (0.9)	7 (0.7)
Nuclear sclerotic and cortical, N (%)	15 (0.2)	95 (1.3)	192 (5.3)	106 (10.5)
Nuclear sclerotic and posterior subcapsular, N (%)	41 (0.5)	142 (2.0)	171 (4.7)	60 (5.9)
Cortical and posterior subcapsular, N (%)	8 (0.1)	14 (0.2)	9 (0.2)	6 (0.6)
All three types, N (%)	11 (0.1)	71 (1.0)	110 (3.0)	44 (4.4)

* 54 men with confirmed cataract have missing information on type of cataract

Incidence of nuclear, cortical and posterior subcapsular cataract by age in the Physicians' Health Study (US male physicians, 1982–1995)

Table 3

Age, years	Person-years	Nuclear Cataract		Cortical Cataract		Posterior Subcapsular Cataract	
		N of Events	IR*	N of Events	IR*	N of Events	IR*
40–49	45,939	11	2.4	5	1.1	12	2.6
50–59	99,857	148	14.8	51	5.1	116	11.6
60–69	72,655	584	80.4	231	31.8	298	41.0
70–79	26,693	793	297.1	356	133.4	312	116.9
≥80	3,738	190	508.3	91	243.5	56	149.8
Total	248,882	1,726	69.4	734	29.5	794	31.9

* Incidence rate per 10,000 person-years

Simplified model for the hazard of cataract types after equating comparable effects; 20,599 participants in the Physicians' Health Study, 1982–1995

Table 5

	Cataract Type		
	Nuclear Sclerotic (1,726 events)		Posterior Subcapsular (794 events)
	HR (95% CI)	Cortical (734 events) HR (95% CI)	
BMI, per 1 kg/m ²	1.03 (1.01–1.05)	1.00 (0.97–1.03)	1.06 (1.03–1.08)
Multivitamin use	0.87 (0.78–0.98)	0.87 (0.78–0.98)	0.87 (0.78–0.98)
Family history of MI	1.14 (0.98–1.32)	1.14 (0.98–1.32)	1.14 (0.98–1.32)
Diabetes mellitus	1.23 (0.96–1.57)	1.65 (1.19–2.29)	1.77 (1.29–2.41)
Smoke < 1 pack/day	1.42 (1.11–1.81)	1.42 (1.11–1.81)	1.42 (1.11–2.81)
Smoke ≥ 1 pack/day	1.56 (1.29–1.88)	1.56 (1.29–1.88)	1.56 (1.29–1.88)
Former smoker	1.14 (1.03–1.26)	1.14 (1.03–1.26)	1.14 (1.03–1.26)

Note: model also allows for different effects of randomized treatment assignment, treats age as the time scale and stratifies on cataract type