HHS Public Access

Author manuscript

Ophthalmic Epidemiol. Author manuscript; available in PMC 2018 June 04.

Published in final edited form as:

Ophthalmic Epidemiol. 2018 February ; 25(1): 1–12. doi:10.1080/09286586.2017.1320413.

Tutorial on Biostatistics: Statistical Analysis for Correlated Binary Eye Data

Gui-shuang Ying[a], **Maureen G. Maguire**[a], **Robert Glynn**[b], and **Bernard Rosner**[b]

[a]Center for Preventive Ophthalmology and Biostatistics, Department of Ophthalmology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[b]Division of Preventive Medicine and the Channing Lab, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

Abstract

Purpose—To describe and demonstrate methods for analyzing correlated binary eye data.

Methods—We describe non-model based (McNemar's test, Cochran-Mantel-Haenszel test) and model-based methods (generalized linear mixed effects model, marginal model) for analyses involving both eyes. These methods were applied to: (1) CAPT (Complications of Age-related Macular Degeneration Prevention Trial) where one eye was treated and the other observed (paired design); (2) ETROP (Early Treatment for Retinopathy of Prematurity) where bilaterally affected infants had one eye treated conventionally and the other treated early and unilaterally affected infants had treatment assigned randomly; and (3) AREDS (Age-Related Eye Disease Study) where treatment was systemic and outcome was eye-specific (both eyes in the same treatment group).

Results—In the CAPT ($n = 80$), treatment group (30% vision loss in treated vs. 44% in observed eyes) was not statistically significant ($p = 0.07$) when inter-eye correlation was ignored, but was significant ($p = 0.01$) with McNemar's test and the marginal model. Using standard logistic regression for unfavorable vision in ETROP, standard errors and p-values were larger for person-level covariates and were smaller for ocular covariates than using models accounting for inter-eye correlation. For risk factors of geographic atrophy in AREDS, two-eye analyses accounting for inter-eye correlation yielded more power than one-eye analyses and provided larger standard errors and p-values than invalid two-eye analyses ignoring inter-eye correlation.

Conclusion—Ignoring inter-eye correlation can lead to larger p-values for paired designs and smaller p-values when both eyes are in the same group. Marginal models or mixed effects models using the eye as the unit of analysis provide valid inference.

Keywords

Correlated binary data; generalized estimating equations; generalized linear mixed effects model; inter-eye correlation; marginal model

CONTACT: Gui-shuang Ying, PhD, gsying@mail.med.upenn.edu, Center for Preventive Ophthalmology and Biostatistics, Department of Ophthalmology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the writing and content of this article.

Introduction

In ophthalmic research, the primary outcome may be binary (e.g. presence or absence of an eye disease) and is often obtained from both eyes of a patient to evaluate the effect of a systemic or ocular treatment or to evaluate the association of person-level or eye-level factors with the eye-level binary outcome. Because the outcome measures from a person's two eyes are usually positively correlated, the appropriate statistical analysis requires accounting for the inter-eye correlation.

Similar to the analysis of correlated continuous eye data discussed in our previous paper,[1] correlated binary eye data are often analyzed using one of the following approaches:[2,3]

1. Ignoring the correlation by using the standard chi-square test for comparing proportions between groups or using standard logistic regression models to assess associations with a binary eye outcome. Because these methods are based on the assumption of independence among observations, estimators of treatment effect or association may be unbiased, but estimators of variances and *p*-values are not accurate. As with continuous eye data,[1] the impact of ignoring inter-eye correlation on variance estimators and *p*-values depends on whether two eyes of a subject are in the same or different comparison groups.
2. Using data from only one eye per subject, where the eye is chosen randomly or as the left eye or as the right eye. These methods provide estimators of treatment effect or association that are unbiased, but because half of the data are not used, power is reduced because the estimator of variance is greater than if all of the data are used with appropriate accommodation of the correlation.
3. Creating a person-specific binary outcome by using presence of an outcome in either eye or in both eyes. Results from this analysis are usually unbiased and the variance estimate is similar to using all of the data with appropriate accommodation of the correlation; however, eye-specific covariates cannot be accommodated.
4. Performing two separate analyses, one for left eye data, another for right eye data. Each of the analyses has the limitations noted in (2). In addition, estimates may differ between the two analyses when there is no biological basis for the associations to be affected by laterality, complicating interpretation.
5. Using McNemar's test for comparing paired proportions. This test is valid when the two eyes of each subject are in two different comparison groups. However, McNemar's test cannot be used when two eyes are in the same comparison group or when adjustment for covariates is needed.

The development of statistical methodology and statistical software for correlated binary data allows the use of generalized regression models for analyzing binary ocular level data that can adjust for the inter-eye correlation. The computational procedures for generalized linear mixed models (GLMM)[4] and population-average (marginal) modelling using the Generalized Estimating Equations (GEE)[5] approach are available in a number of statistical software packages (e.g. SAS, R, Stata, etc.).

In this article, we introduce various statistical approaches applicable to correlated binary eye data and illustrate their use for analyzing data from three real clinical studies of eye disease.

Methods

Non-model based analysis for correlated binary data

When two eyes of a subject are in different comparison groups (i.e. paired design), McNemar's test can be used to compare paired proportions of success in two groups.[6] McNemar's test uses the discordant pairs (one eye a success and the other eye a failure) to test whether the proportions in the two discordant cells of a 2 × 2 table are the same. An equivalent test is the Cochran-Mantel-Haenszel (CMH) test where each subject is a stratum.[7] When data are from a mixture of paired (some subjects have two eyes in the study and are in two different groups) and unpaired (some subjects have one eye in the study) observations, the modified CMH test[8] can be used.

Regression models for correlated binary data

Two commonly used modeling approaches that can accommodate correlated binary data are the generalized linear mixed model (GLMM) using maximum likelihood estimation[4] and the marginal model using GEE.[5] GLMM combines the properties of linear mixed effects models for normally-distributed correlated data and generalized linear models for non-normal data.

GLMM was developed to analyze correlated data with either a normal or non-normal distribution (e.g. Binomial, Poisson distributions). The GLMM may contain both fixed effects and random effects. In addition to situations when outcome assessments are made on clusters of related units (such as two eyes of a subject), GLMM is useful in settings where repeated assessments of outcome are made on the same subjects (e.g. repeated assessment of an eye condition in a longitudinal study).[9] In contrast to the standard logistic regression model that assumes the intercept and the effect of each covariate on an outcome are the same across all subjects, GLMM assumes that the effects of some factors are the same for all subjects (i.e. the fixed effects) while the effect of other covariates may vary for different subjects (i.e. the random effects). The GLMM explicitly accounts for the correlation in outcomes between paired eyes of a subject by adding a random effect (such as a random intercept, assuming that the intercept is the same for the two eyes of a subject, but different across different subjects). Other person- or eye-specific effects may be fixed or random. Sometimes in the GLMM, the iterative algorithms used to estimate regression coefficients do not converge or do not provide a positive-definite covariance matrix. Exploring different computational options available in statistical programs (e.g. SAS) may be needed when this happens.

In SAS, the GLMM is executed using PROC GLIMMIX with a RANDOM statement (see the SAS codes in the Appendix). A likelihood ratio test, based on a mixture of chi-square distributions, can be performed to test whether the variance of the random effect of the intercept or covariates is 0 by comparing the fit (i. e. −2 times the log-likelihood ratio) for models with and without considering the intercept or covariates as random effects.[10]

The marginal model using the GEE approach provides an estimate of average log odds of developing a binary outcome corresponding to differences in covariates. Although GEE was initially developed to analyze correlated data from longitudinal repeated measures,[5] it has been extended to other types of correlated data, including observations from paired eyes.[11] In the GEE approach, a correlation structure for the subunits (eyes) within a cluster (person) is selected and a robust estimator of the variance of the regression coefficients, the "sandwich" estimator, is employed. In contrast to the GLMM which has a subject-specific interpretation, GEE is a marginal model approach, as it does not incorporate any random effects into the model. The marginal model takes account of the inter-eye correlation by estimating the covariance among all the residuals from a single subject, assuming the residuals from a subject are correlated, while the standard logistic regression model assumes the residuals are independent. The marginal model can be executed in SAS using PROC GENMOD.

In fitting the GLMM or marginal model, the specification of the covariance or correlation structure is needed to account for the inter-eye correlation. We have described different commonly used types of covariance or correlation structure for analysis of correlated continuous eye data in a companion paper.[1] These descriptions also apply to correlated binary eye data.

We demonstrate the application of the GLMM model and marginal models to analyze binary correlated eye data from three clinical studies as described below.

Our analyses are purely for the purpose of demonstrating various statistical analysis approaches, thus results presented in this paper are not necessarily the same as previously published results from these studies, and should not be used for clinical care or research. The institutional review board associated with each clinical center approved the study protocol and informed consent was obtained from each patient. Each study adhered to the tenets of the Declaration of Helsinki.

All statistical analyses were performed in SAS 9.4 (SAS Institute Inc., Cary, NC), and the SAS codes for analyzing data from these three example studies are included in the Appendix.

Example 1: Analysis of binary data from the CAPT using paired design

We first demonstrate the analysis for correlated binary data by analyzing data from the Complications of Age-related Macular Degeneration Prevention Trial (CAPT).[12,13] CAPT was a multi-center randomized clinical trial to evaluate whether low-intensity laser treatment for eyes with drusen could prevent vision loss from age-related macular degeneration. The study enrolled 1052 participants aged at least 50 years, with visual acuity (VA) 20/40 or better and at least 10 large drusen in each eye. One randomly selected eye was assigned to laser treatment, and the contralateral eye had no treatment. For demonstration purposes, we compared the proportions of loss of 2-lines or more in VA from baseline to Year 5 between eyes treated with laser and the contralateral untreated eye among the participants enrolled from two clinical centers ($N = 80$ subjects). The data for this example is available at: https://sites.google.com/a/channing.harvard.edu/bernardrosner/channing.

Example 2: Analysis of correlated binary data from the ETROP study with a mixture of paired and unpaired data

The Early Treatment for Retinopathy of Prematurity (ETROP) Study was designed to test the hypothesis that earlier treatment in selected high risk cases of acute retinopathy of prematurity (ROP) results in better visual outcomes than conventionally timed treatment.[14,15] Infants with bilateral high-risk pre-threshold ROP ($n = 317$) had one eye randomized to early treatment and the fellow eye managed conventionally. In infants with unilateral high-risk pre-threshold ROP ($n = 84$), the eye with high-risk pre-threshold ROP was randomized to early treatment or to conventional management. The primary outcome was grating visual acuity (defined as favorable or not favorable) at 9 months. For demonstration purposes, we first compared the unfavorable treatment outcome between two treatment groups among bilateral cases only, unilateral cases only, and bilateral cases and unilateral cases combined. We then evaluated the demographic and ocular factors associated with an unfavorable vision outcome.

Example 3: Analysis for correlated data from AREDS with both eyes in the same treatment group

The Age-Related Eye Disease Study (AREDS) was a multi-center clinical trial designed to evaluate the effect of antioxidant vitamins and zinc supplements on progression of age-related macular degeneration (AMD) and visual acuity.[16,17] Participants were randomly assigned to receive daily oral tablets containing one of the following: (1) antioxidants; (2) zinc; (3) antioxidants plus zinc; or (4) placebo. For demonstration purposes, we evaluated the effect of systemic treatment and person-level or eye-level risk factors on development of central geographic atrophy (GA) in an eye by Year 5 among a high-risk subgroup.

These three studies used different study designs in terms of how two eyes of a subject were assigned to treatment groups. Typically, the correlation (r) between the two eyes of a subject are positively correlated ($r > 0$). In the first example using CAPT data and the second example using ETROP data, when two eyes of a subject were eligible for the study, they were assigned to different treatment groups (one eye treated and fellow eye not treated; i.e. a paired design). The variance of the difference in group proportions is $\mathrm{Var}(p_1 - p_2) = \mathrm{Var}(p_1) + \mathrm{Var}(p_2) - 2r*\sqrt{\mathrm{Var}(p_1)}*\sqrt{\mathrm{Var}(p_2)}$ (where p1 and p2 are the proportion, n, with occurrence of the binary outcome in two treatment groups 1 and 2 respectively). Thus when the correlation is ignored (i.e. treated as $r = 0$) the estimated variance is too large, yielding p-values that are too large. In the third example using AREDS data, two eyes of a subject were in the same comparison group (from systemic treatment). When the responses of the two eyes are included in the group proportion, their contribution to the variance of the group proportion is:

$$Var[(p1 + p2)/2] = [2p(1 - p) + 2rp(1 - p)]/4 = (1 + r)p(1 - p)/2$$

where p is the proportion with occurrence of the binary outcome in a treatment group. When the correlation is ignored (i.e. treated as $r = 0$), the estimated variance is too small, yielding p-values that are too small.

Results

Results from analysis of binary data from the CAPT

Among 80 participants from two clinical centers who completed the 5-year follow-up visit, 24 (30.0%) treated eyes and 35 (43.8%) untreated eyes lost 2-lines or more from baseline at Year 5 (Table 1a). Because of the paired design (one eye treated, fellow eye untreated), the binary VA outcomes between the two eyes were likely to be correlated even when receiving different treatments. In fact, the outcomes were correlated in that 76% of patients had the same VA outcome in each eye when only 53% would be expected if the outcomes were independent (70.0% * 56.3% + 30.0% * 43.8%) and the kappa statistic for agreement in outcomes beyond chance was 0.50 (95% CI: 0.31–0.69). The simplest analysis is to use McNemar's test to compare the paired proportions. The McNemar's test for comparing counts in two discordant cells of Table 1a (i.e. 15 patients with loss of two or more lines only in the untreated eye vs. 4 patients with loss of two or more lines only in the treated eye) yields a p-value of 0.01. We can also use the equivalent CMH test that uses the patient identification number as a strata variable which provides the same p-value as McNemar's test. If a standard chi-square test that ignores the inter-eye correlation is inappropriately used to compare the two proportions, a larger p-value ($p = 0.07$) is obtained.

We can also use model-based analyses that account for inter-eye correlation. In these models, the binary outcome is loss of two or more lines or not, and the predictor is treatment group. The results from the standard logistic regression model that ignores the inter-eye correlation, GEE (using either a working independence correlation structure or a compound symmetry correlation structure), and the GLMM using a random intercept are shown in Table 1b. The standard logistic regression approach that ignores the inter-eye correlation provides the same p-value as the standard chi-square test ($p = 0.07$), and the GEE models that account for the inter-eye correlation provide the same p-value as McNemar's test ($p = 0.01$). The SE of the regression coefficient (i.e. log odds ratio) and the p-value for the treatment group comparison is larger from standard logistic regression than that from GEE, due to the lack of adjustment for the inter-eye correlation. In general, the regression coefficients from the standard logistic regression model are the same as GEE using a working independence correlation structure, but with a contralateral design, the SEs and p-values for the regression coefficients are usually smaller from the GEE model after accounting for the inter-eye correlation. For the comparison between treated eye vs. control eye, the GLMM provides a larger regression coefficient and larger SE than GEE (Table 1b). The estimated variance for the random intercept from GLMM is 1.62 ($p = 0.002$), supporting the use of a random intercept to account for the inter-eye correlation.

The odds ratio (OR) for the loss of 2-lines or more in treated eyes as compared to control eyes is 0.27 (0.09–0.80) from McNemar's test, 0.55 (0.35–0.87) from GEE, 0.50 (0.25, 1.01) from the GLMM, and 0.55 (0.29–1.06) from standard logistic regression that ignores the inter-eye correlation. The OR (95% CI) from McNemar's test is very different from that from GEE, as they are estimating different quantities. The OR from McNemar's test is estimating the odds of loss of 2-lines or more for a treated eye vs. a control eye within the same subject in this paired design where one eye was treated and the fellow eye was

assigned to control. The OR from a marginal model (such as GEE) is estimating the OR for a treated vs control eye from different subjects where subject A's eye is treated and subject B's eye is assigned to placebo.

Results from analysis of correlated binary data from the ETROP study with a mixture of paired and unpaired data

Among 372 infants (292 bilateral cases, and 80 unilateral cases) who had grating visual acuity obtained at 9 months, the proportions with unfavorable vision are lower in the early treatment group than in the control group for both bilateral cases (15.4% vs. 21.2%, Table 2a) and unilateral cases (6.8% vs. 8.3%, Table 2b). For the statistical comparison of proportions with unfavorable vision, the standard chi-square test or Fisher's exact test can be used for the unilateral cases because only one eye per subject was included in the analysis. For the bilateral cases, the paired design is accommodated by McNemar's test or the equivalent CMH test, because of the high inter-eye correlation with kappa of 0.62 (95% CI: 0.51–0.74). The p-value for the comparison in the bilateral cases using McNemar's test is statistically significant ($p = 0.003$) and the p-value for the comparison among unilateral cases using Fisher's exact test is not significant ($p = 1.00$) due to the small number of unilateral cases and the smaller difference in the proportions with unfavorable vision.

To make an overall comparison between eyes treated early vs. conventionally in the proportion with unfavorable vision, the data from bilateral cases and unilateral cases can be combined. The ETROP Study group used the modified CMH test[8] that combines the data from infants with bilateral disease and unilateral disease into one overall chi-square test ($\chi^2 = 8.03$; $p = 0.005$), and the OR for unfavorable vision in early treated eyes as compared to conventionally treated eyes is 0.37 (95% CI: 0.19–0.75). If the inter-eye correlation is ignored by using the standard chi-square test for the bilateral and unilateral combined data, the p-value is not statistically significant ($\chi^2 = 3.60$, $p = 0.058$), and the OR for unfavorable vision in early treated eyes as compared to conventionally treated eyes is 0.67 (95% CI: 0.45–1.01). The GEE models provide a p-value similar to that from the modified CMH test, but an OR similar to that associated with the chi-square test (Table 2c). The GLMM provides a larger regression coefficient, SE and p-value than the GEE models.

We also evaluated risk factors for unfavorable vision using various multivariate regression models (Table 3). Compared with the standard logistic regression model that ignored the inter-eye correlation, the GEE models and GLMM usually provided larger SEs and p-values for both person-specific and eye-specific risk factors including zone 1 ROP, stage 3 ROP and plus disease (Table 3). For example, the association between whether the infant was transferred in and unfavorable vision is significant from the standard logistic regression model ($p = 0.04$), but was not significant in the GEE model under a working independence correlation structure ($p = 0.09$) or under a compound symmetry correlation structure ($p = 0.06$) and GLMM ($p = 0.09$). For eye-level factors, the GLMM generally provides a larger regression coefficient and larger SE than GEE (Table 3). In the GLMM model, the variance estimate for the random intercept is 1.99 ($p < 0.0001$), supporting the use of a random intercept to account for inter-eye correlation.

Results from analysis for correlated data from AREDS with both eyes in the same treatment group

Although the AREDS treatment is systemic, the outcome measurement (i.e. development of central GA by year 5) is eye-specific. Among 4572 eyes from 2286 participants with baseline AMD categories of 2 or 3 in the worse eye and followed-up for 5 years, 325 (7.1%) eyes had central GA by year 5 (7.3% in the placebo group, 5.8% in the antioxidants group, 8.7% in the zinc group, and 7.3% in the Zinc + antioxidants group). The inter-eye correlation in development of central GA is strong (kappa = 0.44 (95% CI: 0.37–0.51)). Among 2286 patients, 78 (3.4%) developed GA in both eyes, whereas only 0.5% (7.1% × 7.1%) would be expected if development of GA were independent between the two eyes.

We first compared the incidence of central GA across the four treatment groups using: (1) a standard chi-square test ignoring inter-eye correlation; (2) logistic regression using the right eye only; (3) logistic regression using the left eye only; (4) logistic regression using a person-level outcome defined as central GA in either eye; (5) a marginal logistic regression model (GEE) for both eyes using a working independence correlation structure; (6) GLMM for both eyes using a compound symmetry correlation structure for the random intercept.

The chi-square test without adjusting for inter-eye correlation provided the smallest p-value ($p = 0.08$, Table 4). The p-values from GEE and GLMM are larger ($p > 0.20$) after adjusting for the inter-eye correlation. The variance for the random intercept from GLMM is 3.32 ($p < 0.0001$) supporting use of the random intercept. The analysis of the left eye only, right eye only, or of central GA determined on a person-specific basis provided larger p-values than analyses using two eyes (Table 4), indicating the loss of information from not considering the data from two eyes simultaneously. Of note, the GLMM using PROC GLIMMIX in SAS did not converge when the default estimation method of RSPL (residual likelihood subjective-specific expansion pseudo-likelihood technique) was used. When the estimation method of RMPL (residual likelihood marginal expansion pseudo-likelihood technique) was used, the algorithm converged.

We also evaluated risk factors for the incidence of central GA by year 5 using multivariable regression models. The results from these eye-level and person-level analyses are shown in Table 5. The standard logistic regression model for two-eye analysis without adjusting for inter-eye correlation provides smaller p-values than GEE and GLMM.

Discussion

In this article, we have described and demonstrated generalized linear mixed models and marginal models for analyzing correlated binary data from two eyes of a subject. We have seen that ignoring the inter-eye correlation between the eyes of the same subject can lead to inaccurate estimates of standard errors of coefficients and p-values. The analysis of data from one eye only (left eye only, right eye only or a randomly selected eye) may produce inefficient estimates of the association. The impact of ignoring inter-eye correlation on the standard error and p-value depends on whether the two eyes of a subject are in the same comparison group or not. When two eyes of a subject are in different comparison groups (i.e. paired design) as in the CAPT example, if the inter-eye correlation is ignored, the

standard error and *p*-value are generally too large. When the two eyes are in the same comparison group as in the AREDS example, the standard error and *p*-value for the treatment effect are generally too small when the inter-eye correlation is ignored. Changing from an analysis approach that ignores the correlation to one that accommodates the correlation can have a substantial impact on inference. In the marginal model using GEE, the impact on inference between choice of a working independence correlation structure and a compound symmetry structure to accommodate the correlation was small. However, the working independence correlation has the feature that the estimates of regression coefficients are the same as those without accounting for inter-eye correlation and only the SE and *p*-value change after adjusting for inter-eye correlation.

When covariates need to be considered, the GLMM and GEE models are the primary regression analysis approaches for correlated binary eye data. GLMM and GEE models have different interpretations and one model can be more appropriate than another depending on the study question.[18,19] When the primary exposure variable is a treatment variable that differs by design between two eyes of the same subject (i.e. a paired design), the odds ratios provided by mixed models apply to the change expected for developing the binary outcome within an individual when a different treatment is used. Conversely, if the primary exposure variable is a treatment variable that is the same for each eye (e.g. either from a systemic treatment such as the AREDS formulation, or an ocular treatment which is the same for each eye as in a parallel design where the same eye drop is administered bilaterally to a given patient, but patients are randomized to different types of eye drops), the odds ratios estimated by population-average models (i.e. GEE) apply to differences expected between subjects when a different treatment is used. Unlike models for correlated continuous data that provide similar estimates for regression coefficients, the models for correlated binary data provide different estimates. Neuhaus has shown that there are systematic differences in regression coefficients between the two model approaches, with population-average models (such as GEE) yielding coefficients closer to 0 and OR's closer to 1 than mixed-effects models such as GLMM.[18,19]

In conclusion, there is often moderate to high concordance in binary outcomes between the two eyes of a subject. The standard chi-square test or standard logistic regression models that assume the independence of individual data points do not consider the inter-eye correlation and thus can lead to invalid conclusions. When there are no covariates for consideration, McNemar's test and the Cochran-Mantel-Haenszel test are appropriate for comparing paired proportions, and the modified Cochran-Mantel-Haenszel test is appropriate for a mixture of paired data and unpaired data. When there are other covariates that need to be accounted for, the GLMM and marginal model are the two main appropriate approaches for analyzing correlated binary eye data. Both the mixed effects model and marginal model require specification of the covariance or correlation structure. In marginal models using GEE, even when the structure of the correlation matrix is mis-specified, asymptotically unbiased estimates of the regression coefficients and variance are provided. In our three examples, the GLMM models and marginal models under various specifications of covariance/correlation structures led to similar conclusions.

Acknowledgments

Funding

This study was supported by grants R01EY022445 and P30 EY01583-26 from the National Eye Institute, National Institutes of Health, Department of Health and Human Services and an unrestricted grant from Research to Prevent Blindness to the University of Pennsylvania.

References

1. Ying GS, Maguire MG, Glynn R, Rosner B. Tutorial on biostatistics: linear regression analysis of continuous correlated eye data. Ophthalmic Epidemiol. 2017; 24:130–140. [PubMed: 28102741]
2. Armstrong RA. Statistical guidelines for the analysis of data obtained from one or both eyes. Ophthalmic Physiol Opt. 2013; 33:7–14. [PubMed: 23252852]
3. Murdoch IE, Morris SS, Cousens SN. People and eyes: statistical approaches in ophthalmology. Br J Ophthalmol. 1998; 82:971–973. [PubMed: 9828786]
4. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. J Am Stat Assoc. 1993; 88:9–25.
5. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics. 1986; 42:121–130. [PubMed: 3719049]
6. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 1947; 12:153–157. [PubMed: 20254758]
7. Wittes J, Wallenstein S. The power of the Mantel-Haenszel Test. Biometrics. 1993; 49:1077–1087. [PubMed: 8117902]
8. Duffy SW, Rohan TE, Altman DG. A method for combining matched and unmatched binary data. Application to randomized, controlled trials of photo-coagulation in the treatment of diabetic retinopathy. Am J Epidemiol. 1989; 130:371–378. [PubMed: 2750732]
9. Ten Have TR, Kunselman AR, Tran L. A comparison of mixed effects logistic regression models for binary response data with two nested levels of clustering. Stat Med. 1999; 18:947–960. [PubMed: 10363333]
10. Self SG, Liang K-Y. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Assoc. 1987; 82:605–610.
11. Liang KY, Zeger SL. Regression analysis for correlated data. Annu Rev Public Health. 1993; 14:43–68. [PubMed: 8323597]
12. The Complications of Age-Related Macular Degeneration Prevention Trial (CAPT): rationale, design and methodology. Clin Trials. 2004; 1:91–107. [PubMed: 16281465]
13. Complications of Age-related Macular Degeneration Prevention Trial Research Group. Laser treatment in patients with bilateral large drusen: the Complications of Age-related Macular Degeneration Prevention Trial (CAPT). Ophthalmology. 2006; 113:1974–1986. [PubMed: 17074563]
14. Good WV, Hardy RJ. The multicenter study of Early Treatment for Retinopathy of Prematurity (ETROP). Ophthalmology. 2001; 108:1013–1014. [PubMed: 11382621]
15. Good WV. Final results of the Early Treatment for Retinopathy of Prematurity (ETROP) randomized trial. Trans Am Ophthalmol Soc. 2004; 102:233–248. [PubMed: 15747762]
16. The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no. 1. Control Clin Trials. 1999; 20:573–600. [PubMed: 10588299]
17. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. Arch Ophthalmol. 2001; 119:1417–1436. [PubMed: 11594942]
18. Neuhaus JM, Kalbfleisch JD, Hauck WW. A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. Intl Stat Rev. 1991; 59:25–35.
19. Neuhaus JM. Statistical methods for longitudinal and clustered designs with binary responses. Stat Methods Med Res. 1992; 1:249–273. [PubMed: 1341660]

Appendix: SAS Codes

SAS Codes for analysis of CAPT data

```
/* for standard chi-square test */
proc freq data=VA;
  tables  group*loss3/nocol  nopercent chisq measures;
run;
/* for Cochran-Mantel-Haenszel test */
proc freq data=VA;
  tables  id*group*loss3/cmh  noprint measures;
run;
/* for McNeMar test */
proc freq data=paired_loss3;
  tables    loss3_trt*loss3_con/norow nocol agree;
run;
/***********************************
  Model-based Analysis
***********************************/
/* Standard logistic regression analysis without account correlation */
proc logistic data=model_data descending;
  class id;
  model loss3=group;
run;
  /* GEE under working independence correlation **/
/* In the model statement, the type 3 is specified to request type 3 score
test */
proc genmod data=model_data descending;
  class id;
  model  loss3=group/dist=bin  type3;
  repeated sub=id/type=ind corrw;
run;
/* GEE under compound symmetry correlation **/
proc genmod data=model_data descending;
  class id;
  model loss3=group/dist=bin type3;
  repeated sub=id/type=cs;
run;
  /* glimmix under compound symmetry **/
proc glimmix data=model_data;
  class id group;
  model loss3=group/dist=bin solution;
  random intercept /sub=id type=cs;
  covtest DiagG;
```

```
  covtest zerog;
  format  hypstat  hyp2fmt.  Diabetic $diabet2f. cigsmk cig2smkf.;
run;
```

SAS codes for analysis of ETROP data

```
/* for unilateral cases **/
proc freq data=comb;
  tables  group*outcome/chisq nocol  nopercent;
  where type='Asymmetric';
run;
  /* for bilateral cases */
data     treated(keep=id  outcome  rename= (outcome=outcome_trt))
    control(keep=id  outcome  rename=(outcome=outcome_con));
  set comb;
  if type='Bilateral ' and group=0 then output treated;
  if type='Bilateral ' and group=1 then output control;
run;
proc sort data=treated; by id;
proc sort data=control; by id;
data comb_bi;
  merge treated control; by id;
run;
  /* for bilateral cases **/
proc freq data=comb_bi;
  tables outcome_trt*outcome_con/agree nocol nopercent;
run;
/** make the data for modified CMH test */
data comb;
  set  bilateral  (in=aa)  asymmetric (in=bb);
  if aa then do; id2=id; type='Bilateral'; end;
  if    bb    then    do;    id2=999; type='Asymmetric'; end; /* for
unilateral cases */
run;
  /* consider correlation using modified CMH */
proc freq data=comb;
  tables id2*group*outcome/cmh noprint measures;
run;
/* ignore inter-eye correlation */
proc freq data=comb;
  tables  group*outcome/chisq  nocol nopercent measures;
run;
/* standard logistic regression, ignore correlation **/
proc logistic data=comb descending;
```

```
  class group /param=reference;
  model outcome=group;
run;
  /* using the GEE: working independence correlation **/
proc genmod data=comb descending;
  class id;
  model outcome=group/dist=bin type3;
  repeated subject=id/type=ind;
run;
  /* using the GEE: compound symmetry **/
proc genmod data=comb descending;
  class id;
  model outcome=group/dist=bin type3;
  repeated subject=id/type=cs;
run;
  /* using GLMM: compound symmetry */
proc glimmix data=comb;
  class id group;
  model  outcome=group/dist=bin  solution g;
  random intercept /sub=id type=cs;
run;
/*********************************
  Analyze the risk factors for the unfavorable visual outcome
  *********************************/
proc logistic data=VA_eye_uni descending;
  class PNO bilateral sex inborn stage3 zone1 plus group/param=reference;
  model    VA_unfav=group    bilateral GSTAGEWK sex inborn zone1 stage3 plus;
run;
proc genmod data=VA_eye_uni descending;
  class PNO bilateral sex inborn stage3 zone1 plus group;
  model    VA_unfav=group    bilateral GSTAGEWK sex inborn zone1 stage3
plus /dist=bin link=logit type3;
  repeated subject=pno/type=ind;
run;
proc genmod data=VA_eye_uni descending;
  class PNO bilateral sex inborn stage3 zone1 plus group;
  model    VA_unfav=group    bilateral GSTAGEWK sex inborn zone1 stage3
plus /dist=bin link=logit type3;
  repeated subject=pno/type=cs;
run;
proc glimmix data=VA_eye_uni;
  class PNO bilateral sex inborn stage3 zone1 plus group;
  model    VA_unfav=group    bilateral GSTAGEWK sex inborn zone1 stage3
plus /dist=bin link=logit solution;
```

```
  random intercept/sub=PNO type=cs g;
  covtest zerog;
run;
```

SAS codes for analyzing the AREDS data

```
/** treatment group comparison without adjust for inter-eye correlation ***/
proc freq data=CGA5yr_eyes;
  tables trtcat*CGA5yr/nopercent nocol chisq;
run;
/** for each eye ***/
proc freq data=CGA5yr_eyes;
  tables  eye*trtcat*CGA5yr/nopercent nocol chisq;
run;
/** for unique person ***/
proc freq data=CGA5yr;
  tables     trtcat*GA5yr_p/nopercent nocol chisq;
run;
/*** using GEE ***/
Proc genmod data=CGA5yr_eyes;
  class id trtcat;
  model cGA5yr=trtcat/dist=bin type3;
  repeated subject=id/type=ind;
run;
  /*** using Glimmix ***/
Proc glimmix data=CGA5yr_eyes method=RMPL;
  class id trtcat;
  model  cGA5yr=trtcat/dist=bin  link=logit;
  random  intercept  /type=cs  subject=id g;
  covtest zerog;
run;
/**********************************
  Multivariate Analysis for Risk factors of CGA
***********************************/
proc freq data=CGA5yr_eyes;
  tables  trtcat  sex  smk_now  hypstat AMDcat_eye;
run;
  /*** without adjust for inter-eye correlation ***/
Proc logistic data=CGA5yr_eyes descending;
  class  id  trtcat  sex  smk_now  hypstat AMDcat_eye/ref=first param=ref;
  model cGA5yr=trtcat enrollage smk_now BMI_R AMDcat_eye;
  format  AMDcat_eye  amdcatf.  hypstat hypstatf.;
run;
  /*** using right eye ***/
```

```
Proc logistic data=CGA5yr_eyes descending;
  class id trtcat sex smk_now hypstat AMDcat_eye/ref=first param=ref;
  model cGA5yr=trtcat enrollage smk_now BMI_R AMDcat_eye;
  format  AMDcat_eye  amdcatf.  hypstat hypstatf.;
  where eye='R';
run;
  /*** person level ***/
Proc logistic data=CGA5yr descending;
  class id trtcat sex smk_now hypstat AMDcat_p/ref=first param=ref;
  model    GA5yr_p=trtcat    enrollage smk_now BMI_R AMDcat_p;
  format  AMDcat_eye  amdcatf.  hypstat hypstatf.;
run;
  /*** using GEE ***/
Proc genmod data=CGA5yr_eyes descending;
  class id trtcat sex smk_now hypstat AMDcat_eye/ref=first param=ref;
  model cGA5yr=trtcat enrollage smk_now BMI_R AMDcat_eye/dist=bin type3;
  repeated subject=id/type=ind;
  format AMDcat_eye amdcatf. hypstat hypstatf.;
run;
  /*** using GLIMMIX ***/
Proc glimmix data=CGA5yr_eyes method=RMPL;
  class  id  trtcat  smk_now  hypstat AMDcat_eye/ref=first;
  model cGA5yr=trtcat enrollage smk_now BMI_R AMDcat_eye/dist=bin link=logit
solution;
  random intercept /subject=id type=cs;
  format  AMDcat_eye  amdcatf.  hypstat hypstatf.;
  covtest zerog;
run;
```

Table 1a

Proportions at year 5 of ≥ 2-lines loss from baseline by treatment group in CAPT ($N = 80$ patients).[a]

		≥ 2-lines loss in untreated eye		
		No	**Yes**	**Total**
≥ 2-lines loss in treated eye	No	41 (51.3%)	15 (18.8%)	56 (70.0%)
	Yes	4 (5.0%)	20 (25.0%)	24 (30.0%)
	Total	45 (56.3%)	35 (43.8%)	80

[a]Data were used just to exemplify the statistical techniques, and should not be used for clinical care. The main results of CAPT can be found in reference 13.

Table 1b

Comparison of results from various analysis approaches for comparing the proportion at year 5 with ≥ 2-lines loss from baseline between treatment groups in CAPT (N = 80 subjects).[a]

	Log odds ratio (SE) for treated vs. control	Chi-square value	*p*-value
Non-model-based analysis			
Chi-square test, ignoring the inter-correlation	−0.596 (0.332)	3.24	0.07
McNemar's test	−1.322 (0.563)	6.37	0.01
Cochran-Mantel-Haenszel test	−1.322 (0.563)	6.37	0.01
Model-based analysis			
Standard logistic regression, ignoring the inter-eye correlation	−0.596 (0.332)	3.21	0.07
GEE – working independence correlation matrix	−0.596 (0.230)	6.37	0.01
GEE – compound symmetry correlation matrix	−0.596 (0.230)	6.37	0.01
GLMM, random intercept (compound symmetry correlation matrix)	−0.688 (0.357)	3.72	0.06

[a]Data were used just to exemplify the statistical techniques, and should not be used for clinical care. The main results of CAPT can be found in reference 13.

Table 2a

Comparison of eyes by treatment group among bilateral cases in ETROP (N = 584 eyes of 292 infants).

		Early treatment		
		Unfavorable vision	**Favorable vision**	**Total**
Conventional treatment	Unfavorable vision	37 (12.7%)	25 (8.56%)	62 (21.2%)
	Favorable vision	8 (2.74%)	222 (76.0%)	230 (78.8%)
	Total	45 (15.4%)	247 (84.6%)	292 (100.0%)

Table 2b

Comparison of eyes by treatment group among unilateral cases in ETROP (N = 80 eyes of 80 infants).

	Unfavorable vision	Favorable vision	Total
Early treatment	3 (6.82%)	41 (93.2%)	44
Conventional treatment	3 (8.33%)	33 (91.7%)	36

Table 2c

Comparison of results from various analysis approaches of combined data from bilateral cases and unilateral cases in ETROP.

	Log odds ratio (SE)	Chi-square value	*p*-value
Non-model-based analysis			
Chi-square test, ignoring the inter-eye Correlation	−0.394 (0.209)	3.60	0.058
Modified Cochran-Mantel-Haenszel test	−0.985 (0.355)[a]	8.03	0.005
Model-based analysis			
Standard logistic regression, ignoring the inter-eye correlation	−0.394 (0.209)	3.57	0.059
GEE – working independence correlation Matrix	−0.394 (0.134)	8.52	0.004
GEE – compound symmetry correlation matrix	−0.406 (0.139)	9.00	0.003
GLMM, random intercept (compound symmetry correlation matrix)	−0.452 (0.226)	4.00	0.047

[a] ln(OR) = −1.139 for bilateral cases, = −0.217 for unilateral cases and −0.985 overall.

Table 3

Factors associated with unfavorable vision at 9 months in ETROP (N = 664 eyes of 372 infants).[a]

	Logistic regression without adjusting for inter-eye correlation		GEE (working independence correlation matrix)		GEE (compound symmetry correlation matrix)		GLMM, random intercept (compound symmetry correlation matrix)	
Baseline covariates	**Log odds ratio (SE)**	***p*-value**	**Log odds ratio (SE)**	**Wald test *p*-value**	**Log odds ratio (SE)**	**Wald test *p*-value**	**Log odds ratio (SE)**	***p*-value**
Early treated vs. conventionally treated	−0.40 (0.22)	0.06	−0.40 (0.15)	0.006	−0.40 (0.14)	0.004	−0.47 (0.23)	0.047
Infant-level factors								
Bilateral case	0.63 (0.40)	0.12	0.63 (0.40)	0.11	0.61 (0.40)	0.13	0.58 (0.47)	0.22
Gestational age (per week increase)	0.14 (0.08)	0.06	0.14 (0.11)	0.19	0.14 (0.10)	0.16	0.15 (0.10)	0.13
Male	0.39 (0.22)	0.08	0.39 (0.27)	0.15	0.51 (0.27)	0.053	0.46 (0.29)	0.12
Transferred in	0.52 (0.25)	0.04	0.52 (0.31)	0.09	0.55 (0.30)	0.06	0.58 (0.34)	0.09
Eye-level factors								
Zone 1 ROP at first ROP exam	1.57 (0.30)	<0.001	1.57 (0.36)	<0.001	1.34 (0.33)	<0.001	1.68 (0.41)	<0.001
Stage 3 ROP at pre-threshold	0.27 (0.26)	0.29	0.27 (0.32)	0.40	0.20 (0.27)	0.47	0.26 (0.33)	0.43
Plus disease at pre-threshold	1.22 (0.33)	<0.001	1.22 (0.38)	0.001	0.87 (0.32)	0.007	1.30 (0.45)	0.005

[a] Data were used just to exemplify the statistical techniques, and should not be used for clinical care. The main results of ETROP can be found in reference 15.

Table 4

Comparison among treatment groups of the incidence of central geographic atrophy in AREDS from various analysis approaches.[a]

	Treatment group				*p*-value			
Analysis approach	**Placebo (*n* = 1154 eyes, 577 subjects)**	**Antioxidants (*n* = 1146 eyes, 573 subjects)**	**Zinc (*n* = 1144 eyes, 572 subjects)**	**Antioxidant + Zinc (*n* = 1128 eyes, 564 subjects)**	**Overall[b]**	**Antioxidants vs. placebo**	**Zinc vs. placebo**	**Antioxidants + zinc vs. placebo**
Two-eye analysis, without adjusting inter-eye correlation	81 (7.02%)	68 (5.93%)	99 (8.65%)	77 (6.83%)	0.08	0.29	0.14	0.86
Right eye only	42 (7.28%)	33 (5.76%)	50 (8.74%)	41 (7.27%)	0.29	0.30	0.36	1.00
Left eye only	39 (6.76%)	35 (6.11%)	49 (8.57%)	36 (6.38%)	0.35	0.65	0.25	0.80
Person level (either eye with atrophy)	61 (10.6%)	53 (9.30%)	73 (12.8%)	60 (10.6%)	0.29	0.45	0.25	0.97
Two-eye analysis, adjusting inter-eye correlation using GEE	81 (7.02%)	68 (5.93%)	99 (8.65%)	77 (6.83%)	0.23	0.38	0.23	0.88
Two-eye analysis, adjusting inter-eye correlation using GLMM	81 (7.02%)	68 (5.93%)	99 (8.65%)	77 (6.83%)	0.21	0.37	0.23	0.88

[a] Data were used just to exemplify the statistical techniques, and should not be used for clinical care. The main results of AREDS can be found in reference 17.

[b] Test for heterogeneity among four treatment groups.

Table 5

Analysis of risk factors for central geographic atrophy in AREDS from various analysis approaches.[a]

	Two-eyes, without adjusting for inter-eye correlation		Right eye only		Left eye only		Person-level: central atrophy in either eye		Two-eyes, GEE (working independence correlation)		Two-eye, GLIMMIX (compound symmetry)	
Risk factors	**Log odds ratio (SE)**	***p*-value**	**Log odds ratio (SE)**	***p*-value**	**Log odds ratio (SE)**	***p*-value**	**Log odds ratio (SE)**	***p*-value**	**Log odds ratio (SE)**	***p*-value**	**Log odds ratio (SE)**	***p*-value**
Treatment Group		0.11[b]		0.28[b]		0.43[b]		0.26[b]		0.26[b]		0.16[b]
Antioxidants vs. placebo	−0.15 (0.18)	0.39	−0.22 (0.25)	0.38	−0.08 (0.25)	0.75	−0.17 (0.21)	0.40	−0.15 (0.21)	0.46	−0.14 (0.21)	0.48
Zinc vs. placebo	0.24 (0.16)	0.13	0.25 (0.23)	0.28	0.25 (0.23)	0.29	0.21 (0.19)	0.27	0.24 (0.19)	0.20	0.27 (0.19)	0.16
Zinc + antioxidants vs. placebo	−0.04 (0.17)	0.80	−0.01 (0.24)	0.98	−0.09 (0.25)	0.72	−0.04 (0.20)	0.84	−0.04 (0.20)	0.82	−0.05 (0.20)	0.81
Age	0.07 (0.01)	<0.001	0.06 (0.02)	<0.001	0.07 (0.02)	<0.001	0.06 (0.01)	<0.001	0.07 (0.01)	<0.001	0.07 (0.01)	<0.001
Current smoking	0.57 (0.22)	0.009	0.89 (0.29)	0.002	0.19 (0.34)	0.57	0.61 (0.26)	0.02	0.57 (0.25)	0.02	0.55 (0.27)	0.04
BMI	0.04 (0.01)	0.001	0.05 (0.02)	0.002	0.02 (0.02)	0.18	0.04 (0.01)	0.002	0.04 (0.01)	0.004	0.04 (0.01)	0.004
Worse eye AMD category: 3 vs.2[c]	2.64 (0.22)	<0.001	2.94 (0.35)	<0.001	2.41 (0.28)	<0.001	3.09 (0.36)	<0.001	2.64 (0.22)	<0.001	2.57 (0.22)	<0.001

[a] Data were used just to exemplify the statistical techniques, and should not be used for clinical care. The main results of AREDS can be found in reference 17.

[b] Test for heterogeneity among four treatment groups.

[c] Based on the higher of two eyes and excluding participants with grade 1 in both eyes.