Statistical Analyses of Correlated Eye Data

Gui-shuang Ying, PhD
Professor of Ophthalmology

Center for Preventive Ophthalmology and Biostatistics
Scheie Eye Institute, Perelman School of Medicine
University of Pennsylvania

DCPO Seminar Series 12/10/2020

Outline

- **Features of data from ophthalmic and vision research**
- **Inter-eye correlation and impact on statistical analysis**
- **Rationale and practice for adjustment for inter-eye correlation**
- **Appropriate analysis of correlated eye data**
  - Mixed effects model
  - Marginal model-Generalized estimating equation
  - Cluster bootstrap
- **Examples**
  - Continuous eye data
  - Binary eye data
  - Sensitivity, specificity
  - ROC Analysis

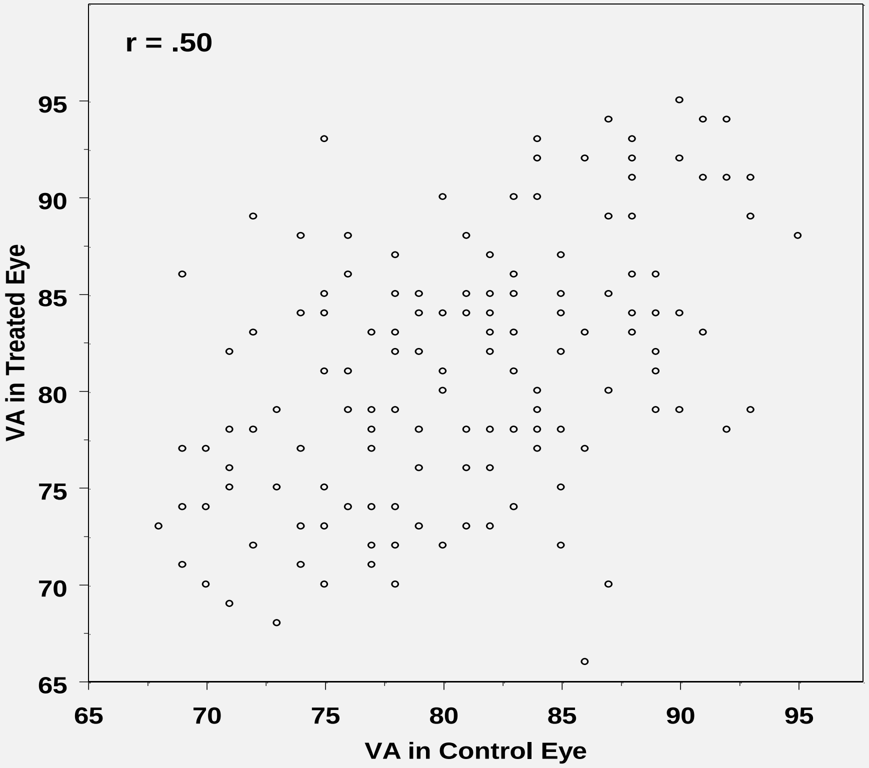
Data from Ophthalmic and Vision Research

- Observational studies: commonly measure 2 eyes of the same subject
- Clinical trials:
    - **Eye specific treatment:** two eyes receive different treatment and inter-eye difference is of interest – CAPT
    - **Systemic treatment:** effect on both eyes, treatment effect is evaluated by comparison of ocular outcome between subjects in different treatment groups – AREDS
- Vision screening of both eyes for eye disorders-Vision In Preschooler Study
- Lab research: measures taken from both eyes of animal

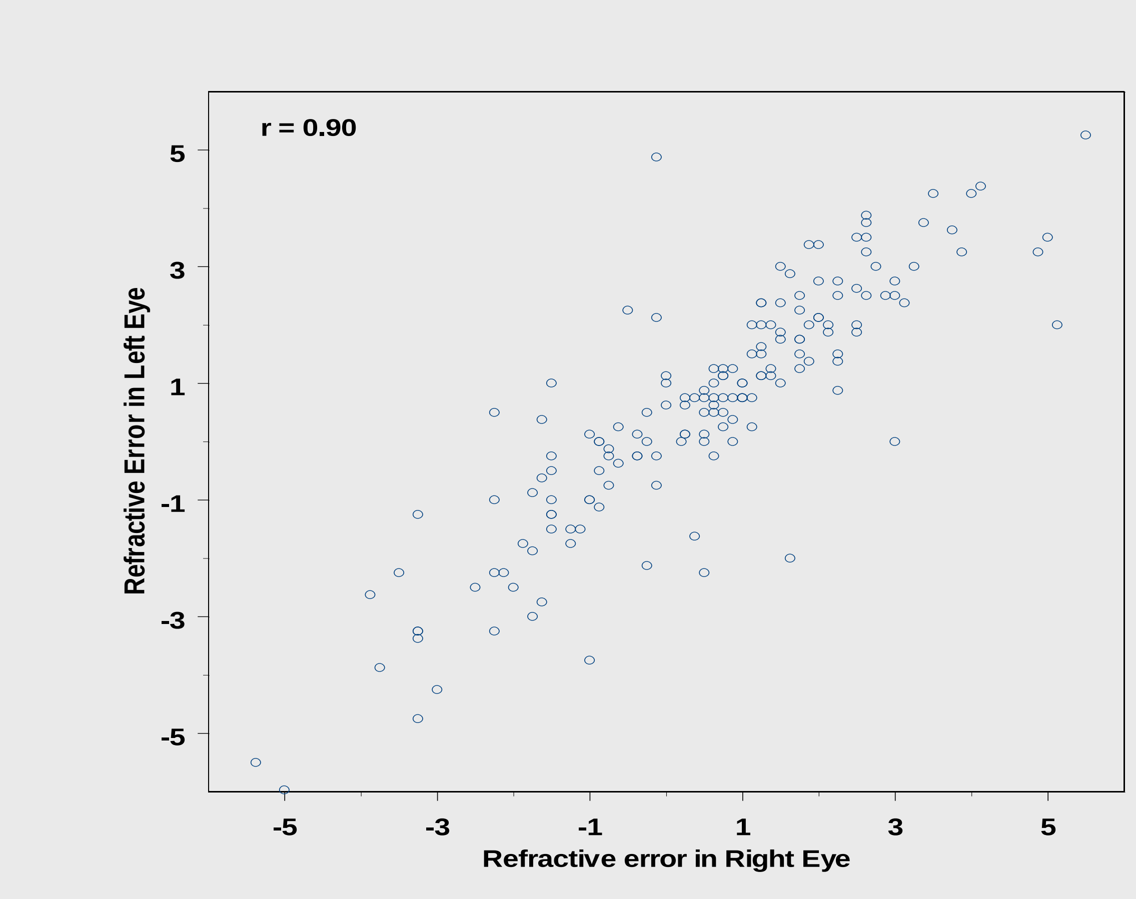
Correlation in Eye Data

- **Positively correlated**: finding in one eye is likely to be more similar to that in the fellow eye of the same subject than to that in eye from different subject
  - ➢ Common environment factors
  - ➢ Genetic factors
- Inter-eye correlation varies: depending on the disease and measurement
  - ➢ High correlation in ROP: >80% are bilateral
  - ➢ Visual acuity
  - ➢ Refractive error

Inter-eye Correlation in Visual Acuity Score



Inter-eye Correlation in Refractive Error



Impact of Inter-eye Correlation on Statistical Analysis

- Existence of inter-eye correlation means each data point does not represent an independent observation
- Two data points from two eyes of a subject should not be treated as the same way as two data points from one eye of two subjects
  - ➢ Two data points from independent two subjects provides more information than those from two eyes of a subject
- Most standard statistical methods assume **independence** of data points
- Point estimate for mean or proportion is still valid without considering correlation
- Variability estimates (SD, SE) and statistical inferences (95% CI, P-value) are invalid if ignoring correlation

Impact of Ignoring Inter-eye Correlation on Statistical Inference

- Depends on
  - 2 eyes in the **same** or **different** comparison groups
  - Strength of inter-eye correlation
- Two eyes in **same** comparison group
  - Variance estimate too low -> p-value too small; confidence interval too narrow
- Two eyes in **different** comparison group
  - Variance estimate too high -> p-value too large; confidence interval too wide

Two Eyes in the Same Group – Impact of r

- N = number of eyes
- r = inter-eye correlation
- Effective sample size = N/(1+r)
- % Under-estimation of SE = $1/\sqrt{(1+r)}$

<u>Example:</u> 200 eyes of 100 people in one comparison group

r	Effective Sample Size	% Under-Estimation of SE
0.0	200	0%
0.2	167	9%
0.4	143	15%
0.6	125	21%
0.8	111	25%
1.0	100	29%

Unit of Analysis – Per Subject

- **Collapse data from paired eyes of a patient into a summary measure**
  - Continuous data: using average of two eyes
  - Binary data: either eye has a condition
- **Advantage:**
  - Simple, standard statistical method can be applied
  - Easy for interpretation
  - Statistically valid (independent assumption met)
- **Disadvantage:**
  - Loss of information when pooling data from paired eyes, statistical analysis is not efficient
  - Amount of information loss depends on the degree of inter-eye correlation

Unit of Analysis - Per Eye

- **Single eye per patient**
  - Left eye only, right eye only, randomly selected eye
  - **Advantage**: Convenient
  - **Disadvantage:** Inefficient, and potential bias (when not all patients have data from both eyes for selection)
- **Two eye analysis using data from both eyes**
  - Analysis at eye-level, while still account for the inter-eye correlation
  - **Advantage**: Make full use of data
  - **Disadvantage**: Need advanced statistical procedure, may not be easily understood and acceptable by clinician

A Review of Practice in 1997 (Murdoch IE, Morris SS, Cousens SN. *BJO* 1998;82:971-73.)

Analytical Approach	# articles (N=79)	(%)
Analysis at level of individual because of nature of observation		
Uniocular disease or therapy	9	(11)
Disease entity requires both eyes for diagnosis	3	(4)
One eye per individual		
Random selection of eye	5	(6)
Right/left selection of eye	7	(9)
Clinical selection of eye (worst eye, first eye with disease etc)	13	(16)
Overall summary of ocular findings per individual		
Pooled findings	13	(16)
Average taken of results from two eyes	6	(8)
Analysis two eyes per individual		
No correction for inter-eye correlation	16	(20)
Correction for correlation between eyes	2	(3)
Paired comparison (fellow eye used as “control”)	5	(6)

Review of Practice in 2017 (Zhang H, Ying GS, BJO 2018)

Analytical Approach	# articles in 1997 (N=79)	# articles in 2017 (N=112)
Analysis at level of individual because of nature of observation		
Uniocular disease or therapy	**9 (11%)**	**44 (39%)**
Disease entity requires both eyes for diagnosis	**3 (4%)**	**12 (11%)**
One eye per individual		
Random selection of eye	**5 (6%)**	**3 (3%)**
Right/left selection of eye	**7 (9%)**	**4 (4%)**
Clinical selection of eye (worst eye, first eye with disease etc)	**13 (16%)**	**9 (8%)**
Overall summary of ocular findings per individual		
Pooled findings	**13 (16%)**	**0 (0%)**
Average taken of results from two eyes	**6 (8%)**	**1 (1%)**
Analysis two eyes per individual		
No correction for inter-eye correlation	**16 (20%)**	**33 (30%)**
Correction for correlation between eyes	**2 (3%)**	**3 (3%)**
Paired comparison (fellow eye used as “control”)	**5 (6%)**	**3 (3%)**

Improve the Practice of Statistical Analyses of Correlated Eye Data

- **ARVO short course**
  - 2011, 2019, 2020, 2021
- **Tutorial papers** (Ying GS, Maguire MG, Glynn RJ, Rosner B)
  - Cross-sectional analysis of continuous correlated eye data (*Ophthalmic Epi*, 2017)
  - Cross-sectional analysis of binary correlated eye data (*Ophthalmic Epi*, 2018)
  - Longitudinal analysis of continuous correlated eye data (*Ophthalmic Epi*, 2020)
  - Sensitivity and specificity analysis of correlated eye data (*IOVS*, 2020)
  - Receiver-Operating Characteristic (ROC) analysis for correlated eye data (*in preparation*)

Mixed-Effects Model

- $$y_{ij} = \beta_0 + \beta_1 x_{ij1} + \ldots\ldots + \beta_k x_{ijk} \quad \leftarrow \text{Fixed effects}$$
  $$+ b_{0i} + b_{1i} z_{ij1} + \ldots\ldots + b_{mi} z_{ijm} \quad \leftarrow \text{Random effects}$$
  $$+ e_{ij}$$
- Assumes random effect follows a normal distribution
- Provides conditional mean of outcome measure for given covariates and random effects
  - ➢ Interpretation of covariate effects is conditional on random effect
- Mixed effects model requires correct specification of both fixed effects and random effects
- Executed using
  - ➢ **PROC MIXED** with **RANDOM** statement in SAS
  - ➢ **LM( )** or **LMER( )** in R
  - ➢ **XTMIXED** in STATA

Random Intercept Mixed-Effects Model

- $y_{ij} = \beta_0 + \beta_1 x_{ij1} + \ldots\ldots + \beta_k x_{ijk}$ ← **Fixed effects**

  $+ b_{0i}$ ← **Random intercept**

  $+ e_{ij}$ ← **Error term**

- b_{0i} assume to be $N(0, D_i)$
- e_{ij} = error term assume to be $N(0, \sigma^2)$
- Explicitly accounts for inter-eye correlation by adding random effect
  - ➢ random intercept: intercept is the same for both eyes of a subject, but different across different subjects

Marginal Model: Generalized Estimating Equation (GEE)

- Developed by Liang KY & Zeger SL, 1986
- Account for inter-eye correlation by estimating the covariance among residuals from two eyes of a subject, assuming residuals from same subject are correlated
  - ➢ Standard linear regression model assumes independence in residuals
- Provides estimate of change of population mean corresponding to change of covariates
- Estimation of marginal model depends only on correctly specifying the linear function relating the mean outcome to the covariates
- Uses a robust variance estimator (i.e., sandwich estimator) for the regression coefficients

Marginal Model in Statistical Softwares

- Executed using
  - **PROC GENMOD** in SAS (using quasi-likelihood approach, without normality assumption)
  - **PROC MIXED** using **REPEATED** Statement in SAS (using likelihood approach, assuming normality of outcome)
  - **GEE( )** in R
  - **XTGEE** in STATA

Covariance/Correlation Structure

- Mixed effects model or marginal model requires specification of a covariance/correlation structure to account for inter-eye correlation
- For cross-sectional study, most commonly used covariate structure are:
  - Unstructured: $\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$

    σ_1^2, σ_2^2 are the variance in two eyes respectively (allowing different), and σ_{12} is their covariance

  - Compound symmetry: $\begin{pmatrix} \sigma^2 & \sigma_{12} \\ \sigma_{12} & \sigma^2 \end{pmatrix}$

    σ^2 is the variance in two eyes (assuming equal) and σ_{12} is their covariance

  - Independence: $\begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$

Working Independence Covariance in GEE

- Independence: $\begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$
- Used in GEE to calculate robust variance estimator of regression coefficients for accounting for inter-eye correlation
- Regression coefficients under “Working independence covariance” are the same as standard linear regression models, but standard errors differ
- Most useful when there is little knowledge available to choose between unstructured and compound symmetry covariance structure

Cluster bootstrap

- A resampling technique for generating the distribution of a statistic of interest (e.g., mean, proportion, sensitivity, specificity, AUC etc.)
- Repeatedly taking a random sample of the same size as original sample **with replacement**
  - ➢ Some subjects were selected in the same sample more than once, while some were never selected
  - ➢ Sampling at subject level
  - ➢ Eligible eyes of the sampled subjects are all included
- From each of bootstrapped samples, a statistic of interest is calculated, generating the distribution of statistic of interest
- SD of the bootstrapped statistic represents the SE of the estimate
- 95% CI of the statistic of interest can be derived based on 2.5^{th} and 97.5^{th} percentile

Example 1: Cross-sectional analysis of continuous correlated eye data

Example 1: Analysis of Refractive Error Data from CATT

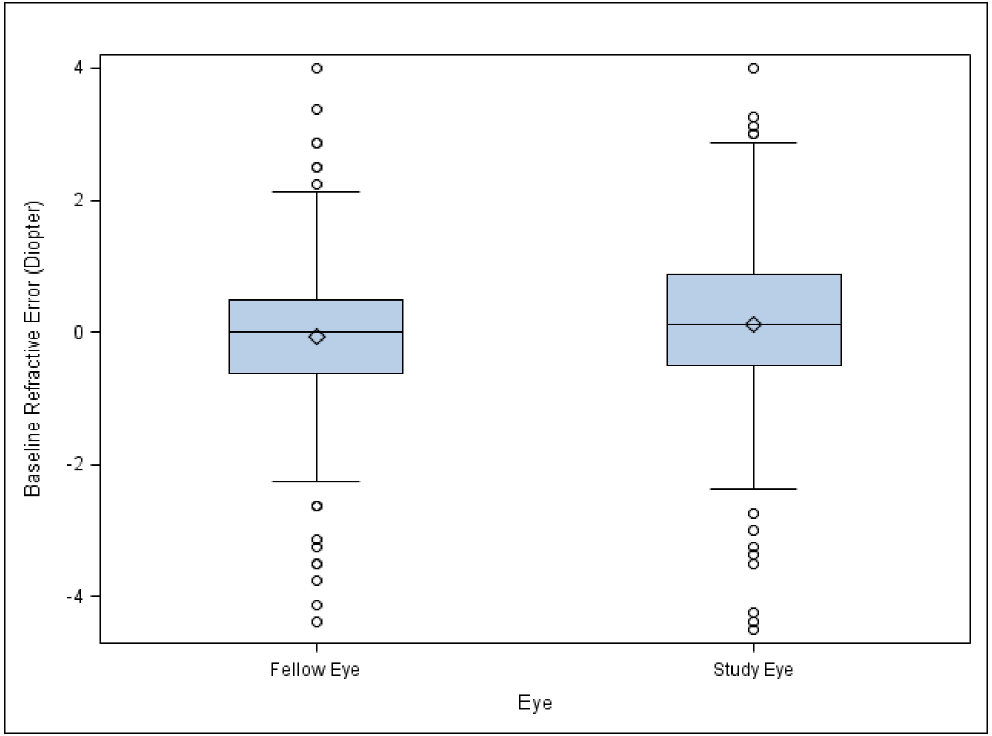
- Comparison of Age-related Macular Degeneration Treatment Trials (CATT)
  - RCT to compare efficacy and safety of ranibizumab vs. bevacizumab
  - Study eye had untreated active choroidal neovascularization (CNV) due to AMD
  - Fellow eye could have or not have CNV

- **Hypothesis:** Morphological changes in retina from active CNV would impact refractive error by changing the axial length of an eye

- Among patients without CNV in fellow eye at baseline, compare baseline spherical equivalent between **study eye <u>with active CNV</u> vs. fellow eye <u>without CNV</u>**

- Restricted to 355 patients who had pseudophakic eyes to eliminate the effect of lens status on refractive error

Refractive Error in Study eye and Fellow eye



Mean (SD) = -0.03 (1.21) D **Mean (SD) = 0.12 (1.17) D**

Inappropriate Analysis: Two-sample t-test

```
proc ttest data=bs_ref_sub;
    class CNV;
    var bs_sphe;
run;
```

The TTEST Procedure

Variable: bs_sphe

CNV	N	Mean	Std Dev	Std Err	Minimum	Maximum
0	355	-0.0338	1.2065	0.0640	-4.3750	9.8750
1	355	0.1165	1.1682	0.0620	-4.5000	4.0000
Diff (1-2)		-0.1503	1.1875	0.0891		

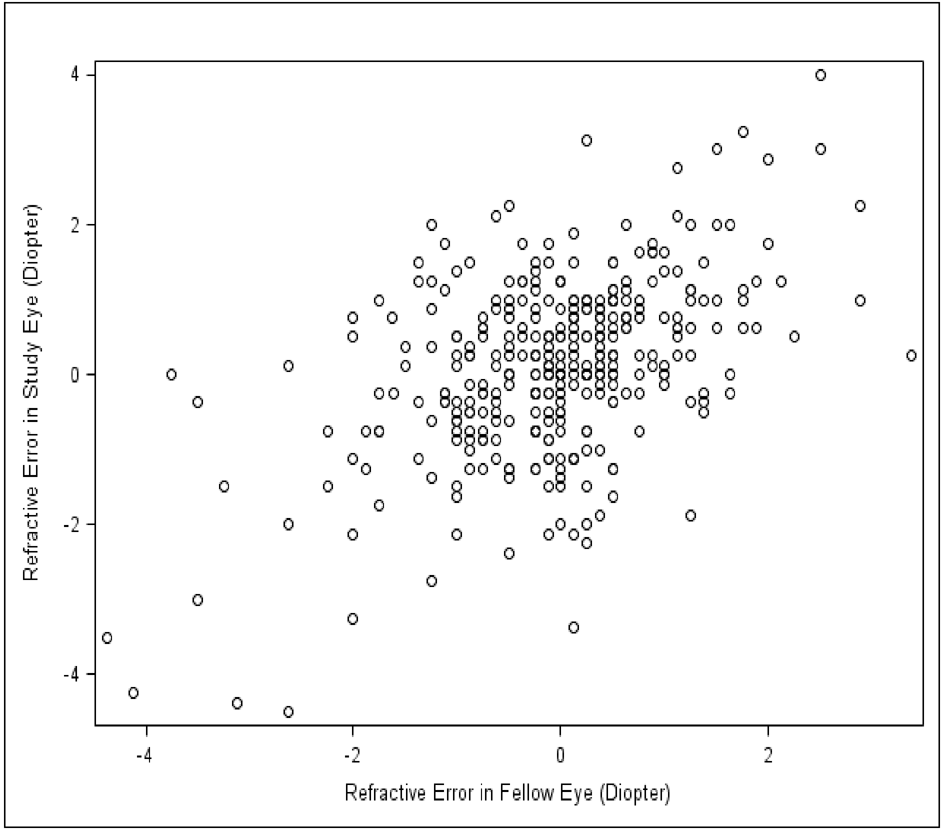
CNV	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		-0.0338	-0.1597	0.0921	1.2065	1.1238	1.3024
1		0.1165	-0.00539	0.2385	1.1682	1.0881	1.2611
Diff (1-2)	Pooled	-0.1503	-0.3253	0.0247	1.1875	1.1287	1.2527
Diff (1-2)	Satterthwaite	-0.1503	-0.3253	0.0247			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	708	-1.69	0.0921
Satterthwaite	Unequal	707.27	-1.69	0.0921

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	354	354	1.07	0.5444

Inter-eye Correlation in Refractive Error



r = 0.43

Paired t-test

```
proc ttest data=CNV01;
   paired sphe1*sphe0;
run;
```

The TTEST Procedure

Difference: sphe1 - sphe0

N	Mean	Std Dev	Std Err	Minimum	Maximum
355	0.1503	1.2629	0.0670	-9.0000	3.7500

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0.1503	0.0185	0.2822	1.2629	1.1764	1.3634

DF	t Value	Pr > |t|
354	2.24	0.0255

Mixed Effects Model: Unstructured

```
proc mixed data=bs_ref_sub noclprint;
    class id CNV;
    model bs_sphe=CNV/s CL;
    random intercept/sub=id type=un;
run;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	id	0.6126
Residual		0.7975

The Mixed Procedure

Fit Statistics

-2 Res Log Likelihood	2190.2
AIC (smaller is better)	2194.2
AICC (smaller is better)	2194.3
BIC (smaller is better)	2202.0

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
1	74.05	<.0001

Solution for Fixed Effects

Effect	CNV	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept		0.1165	0.06303	354	1.85	0.0653	0.05	-0.00740	0.2405
CNV	0	-0.1503	0.06703	354	-2.24	0.0255	0.05	-0.2822	-0.01851
CNV	1	0	.	.	.	.	.	.	.

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
CNV	1	354	5.03	0.0255

Mixed Effects Model: Compound Symmetry

```
Proc mixed data=bs_ref_sub noclprint;
    class id CNV;
    model bs_sphe=CNV/s CL;
    random intercept/sub=id type=cs;
run;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
Variance	id	0.6126
CS	id	0
Residual		0.7975

Fit Statistics

-2 Res Log Likelihood	2190.2
AIC (smaller is better)	2196.2
AICC (smaller is better)	2196.3
BIC (smaller is better)	2207.9

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
2	74.05	<.0001

Solution for Fixed Effects

Effect	CNV	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept		0.1165	0.08303	354	1.85	0.0653	0.05	-0.00740	0.2405
CNV	0	-0.1503	0.06703	354	-2.24	0.0255	0.05	-0.2822	-0.01851
CNV	1	0	.	.	.	.	.	.	.

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
CNV	1	354	5.03	0.0255

Marginal Model: GEE Using Working Independence Covariance

```
proc genmod data=bs_ref_sub;

class id CNV;
model bs_sphe=CNV/dist=normal;
repeated sub=id/type=ind corrw;
run;
```

The GENMOD Procedure

Working Correlation Matrix

	Col1	Col2
Row1	1.0000	0.0000
Row2	0.0000	1.0000

GEE Fit Criteria

QIC	711.9887
QICu	712.0000

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept		0.1165	0.0619	-0.0048	0.2379	1.88	0.0598
CNV	0	-0.1503	0.0669	-0.2815	-0.0191	-2.25	0.0247
CNV	1	0.0000	0.0000	0.0000	0.0000	.	.

Marginal Model: Using PROC MIXED with REPEATED

```
proc mixed data=bs_ref_sub noclprint;
class id CNV;
model bs_sphe=CNV/s CL;
repeated /sub=id type=un;
run;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	id	1.4556
UN(2,1)	id	0.6126
UN(2,2)	id	1.3647

Fit Statistics

-2 Res Log Likelihood	2189.8
AIC (smaller is better)	2195.8
AICC (smaller is better)	2195.8
BIC (smaller is better)	2207.4

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
2	74.51	<.0001

Solution for Fixed Effects

Effect	CNV	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept		0.1165	0.06200	354	1.88	0.0610	0.05	-0.00539	0.2385
CNV	0	-0.1503	0.06703	354	-2.24	0.0255	0.05	-0.2822	-0.01851
CNV	1	0	.	.	.	.	.	.	.

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
CNV	1	354	5.03	0.0255

Inappropriate Analysis: Standard Linear Regression Model

```
proc reg data=bs_ref_sub;
 model bs_sphe=CNV/CLB;
run;
```

```
                        The REG Procedure
                          Model: MODEL1
                   Dependent Variable: bs_sphe

              Number of Observations Read          710
              Number of Observations Used          710

                      Analysis of Variance

                                  Sum of           Mean
Source                   DF      Squares         Square    F Value    Pr > F

Model                     1      4.01177        4.01177       2.84    0.0921
Error                   708    998.37698        1.41014
Corrected Total         709   1002.38875

          Root MSE              1.18749    R-Square     0.0040
          Dependent Mean        0.04138    Adj R-Sq     0.0026
          Coeff Var          2869.70469

                      Parameter Estimates

                 Parameter     Standard
Variable    DF    Estimate        Error   t Value   Pr > |t|     95% Confidence Limits

Intercept    1    -0.03379      0.06303     -0.54     0.5920     -0.15753      0.08995
CNV          1     0.15034      0.08913      1.69     0.0921     -0.02466      0.32533
```

Comparison of Results from Unadjusted Analysis

Analysis approaches	Mean difference between study eyes with CNV vs. fellow eyes without CNV (95% CI), Diopters	Width of 95% CI	P-value
Inappropriate Analysis			
Independent- sample t-test	0.15 (-0.03, 0.33)	0.36	0.09
Standard linear regression model	0.15 (-0.03, 0.33)	0.36	0.09
Appropriate Analysis			
Paired t-test	0.15 (0.02, 0.28)	0.26	0.026
Mixed model, compound symmetry or unstructured	0.15 (0.02, 0.28)	0.26	0.026
Marginal model, PROC MIXED REPEATED, unstructured	0.15 (0.02, 0.28)	0.26	0.026
Marginal model-GEE, working independent	0.15 (0.02, 0.28)	0.26	0.025

Need for Regression Models Using Eye as Unit of Analysis

- Evaluate association between factors and ocular outcome measure
  - ➢Person-specific factors (age, smoking status)
  - ➢Eye-specific factors (AMD status, IOP etc.)

- Need to adjust for other covariates

Comparison of Results from adjusted Analysis

-Adjusted by age, gender, smoking status, geographic atrophy, glaucoma

Analysis approaches	Mean difference between study eyes with CNV vs. fellow eyes without CNV (SE), Diopters	Width of 95% CI	P-value
Inappropriate Analysis			
Standard linear regression model	0.15 (-0.03, 0.32)	0.35	0.10
Appropriate Analysis			
Mixed model, compound symmetry or unstructured	0.15 (0.01, 0.28)	0.27	0.03
Marginal model, PROC MIXED REPEATED, unstructured	0.15 (0.01, 0.28)	0.27	0.03
Marginal model, GEE, working independent	0.15 (0.02, 0.28)	0.26	0.03

Summary of Example 1

- Ignoring inter-eye correlation has some impacts on statistical inference (SE, 95% CI, p-value)
- When two eyes are in different comparison groups, ignoring inter-eye correlation inflates SE, 95% CI and p-value
- Mixed effects model and marginal model provide very similar results
  - Consistent with our general experience that when there is only inter-eye correlation and sample size is not small, there is little difference between mixed effects model and marginal models
- Type of covariance structure used in mixed effects model or marginal models has little impact on the results

Example 2: Cross-sectional analysis of binary correlated eye data

Example 2: Early Treatment for Retinopathy of Prematurity (ETROP) Study

- Designed to evaluate whether early treatment of pre-threshold ROP results in better visual outcome than conventionally timed treatment
- **317 bilateral infants**
  - ➢ one eye randomized to early treatment, fellow eye to conventional treatment
- **84 unilateral infants**
  - ➢ randomized to early treatment or conventional timed treatment
- **Primary outcome**: favorable or unfavorable visual acuity at 9 months
  - ➢ restricted to 292 bilateral infants and 80 unilateral infants who completed 9-month follow-up

ETROP Results: Bilateral and Unilateral Separately

Bilateral 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

P=0.003 from McNemar's test

Unilateral Infants

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

P=1.00 from Fisher's exact test

Inappropriate Analysis: Standard Chi-square test: Bilateral and Unilateral Combined

```
/* ignore inter-eye correlation */
proc freq data=comb;
 tables group*outcome/chisq nocol nopercent
measures;
run;
```

Frequency
Row Pct

Table of group by outcome			
	outcome		
group	0	1	Total
0	263 80.18	65 19.82	328
1	288 85.71	48 14.29	336
Total	551	113	664

Statistics for Table of group by outcome

Statistic	DF	Value	Prob
Chi-Square	1	3.5960	0.0579
Likelihood Ratio Chi-Square	1	3.6056	0.0576
Continuity Adj. Chi-Square	1	3.2149	0.0730
Mantel-Haenszel Chi-Square	1	3.5905	0.0581
Phi Coefficient		-0.0736	
Contingency Coefficient		0.0734	
Cramer's V		-0.0736	

Odds Ratio and Relative Risks			
Statistic	Value	95% Confidence Limits	
Odds Ratio	0.6744	0.4481	1.0149
Relative Risk (Column 1)	0.9355	0.8729	1.0026
Relative Risk (Column 2)	1.3872	0.9868	1.9500

Inappropriate Analysis: Standard logistic regression: Bilateral and Unilateral Combined

```
 /* standard logistic regression, ignore correlation **/
proc logistic data=comb descending;
 class group /ref=first;
 model outcome=group;
run;
```

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.5947	0.1043	233.8973	<.0001
group	1	1	-0.1970	0.1043	3.5688	0.0589

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
group 1 vs 0	0.674	0.448	1.015

GEE -independent: Bilateral and Unilateral Combined

```
 /* Using the GEE: independent working correlation **/
proc genmod data=comb descending;
 class id;
 model outcome=group/dist=bin type3;
 repeated subject=id/type=ind;
 estimate "OR" group 1 -1/exp;
run;
```

Score Statistics For Type 3 GEE Analysis			
Source	DF	Chi-Square	Pr > ChiSq
group	1	8.52	0.0035

Contrast Estimate Results										
		Mean					L'Beta			
Label	Mean Estimate	Confidence Limits		L'Beta Estimate	Standard Error	Alpha	Confidence Limits		Chi-Square	Pr > ChiSq
OR	0.4028	0.3417	0.4670	-0.3940	0.1336	0.05	-0.6559	-0.1321	8.70	0.0032
Exp(OR)				0.6744	0.0901	0.05	0.5190	0.8762		

GEE-compound symmetry: Bilateral and Unilateral Combined

```
 /* Using the GEE: using compound symmetry**/
proc genmod data=comb descending;
 class id;
 model outcome=group/dist=bin type3;
 repeated subject=id/type=cs;
 estimate "OR" group 1 -1/exp;
run;
```

Score Statistics For Type 3 GEE Analysis			
Source	DF	Chi-Square	Pr > ChiSq
group	1	9.00	0.0027

Contrast Estimate Results										
		Mean					L'Beta			
Label	Mean Estimate	Confidence Limits		L'Beta Estimate	Standard Error	Alpha	Confidence Limits		Chi-Square	Pr > ChiSq
OR	0.3999	0.3366	0.4668	-0.4057	0.1392	0.05	-0.6785	-0.1330	8.50	0.0035
Exp(OR)				0.6665	0.0927	0.05	0.5074	0.8755		

Comparison of Results from Various Approaches for Analyzing ETROP Data

Analysis approach	OR (95% CI)	Width of 95% CI	P-value
Inappropriate Analysis			
Chi-square	0.67 (0.45, 1.01)	0.56	0.058
Standard logistic regression	0.67 (0.45, 1.01)	0.56	0.058
Appropriate Analysis			
GEE: working independent	0.67 (0.52, 0.88)	0.36	0.0027
GEE: Compound symmetry	0.67 (0.51, 0.88)	0.37	0.0035

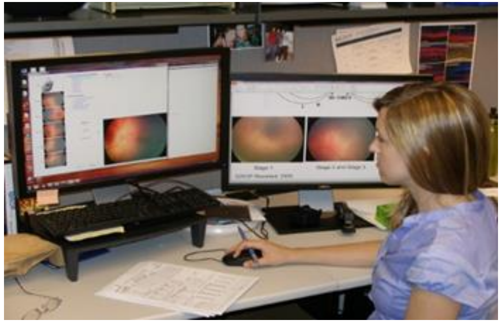
Summary of Example 2

- For correlated binary eye data, the GEE model can properly account for inter-eye correlation, even under the mixture of unilateral and bilateral infants
- Ignoring inter-eye correlation by standard chi-square test or standard logistic regression model inflates 95% CI for OR and p-value
- Type of covariance structure used in the GEE has little impact on the results

Example 3: Sensitivity and Specificity for Correlated Eye data

Example 3: Telemedicine System for the Evaluation of acute-phase retinopathy of prematurity (e-ROP)

- Designed to evaluate the validity of using RetCam images to identify infants with referral-warranted ROP (RW-ROP)
- Infants underwent diagnostic examination and RetCam imaging in both eyes
- Trained non-physician readers in central reading center evaluated images



Analysis for e-ROP Data

- **Primary analysis:** Eye-level analysis

 Comparing image evaluation finding to ophthalmologist clinical examination findings (reference standard)

  - **Sensitivity** = P(T+ | D+) = d/(b+d)
  - **Specificity** = P(T- | D-) = a/(a+c)

	True Disease Status	
Test Result	**D-**	**D+**
T-	**a**	**b**
T+	**c**	**d**

- Enriched sample of 100 infants
  - 29 with RW-ROP
  - 71 without RW-ROP

Inter-eye Agreement in RW-ROP from Clinical Exam

	Right Eye		
Left Eye	**RW-ROP Absent**	**RW-ROP Present**	**Total**
RW-ROP Absent	71	**3**	74
RW-ROP Present	**6**	20	26
Total	77	23	100

Percent agreement=91%

Kappa (95% CI)=0.76 (0.61-0.91)

Cross-tabulation between RW-ROP from image evaluation vs. clinical examination

	Clinical Examination		
Image evaluation	RW-ROP Absent	RW-ROP Present	Total
RW-ROP negative	**131 (86.8%)**	8 (16.3%)	139
RW-ROP positive	20 (13.2%)	**41 (83.7%)**	61
Total	151	49	200

Specificity

Sensitivity

Per-eye analysis: Naïve 95% CI for Sensitivity and Specificity

```
/** get Naïve 95% CI for Sensitivity **/
 proc freq data=subsample;
  tables RWROP_RC/binomial(level='1');
  format RWROP_RC test2f;
  where rwROP_DE=1;
run;

 /** get Naïve 95% CI for Specificity **/
proc freq data=subsample;
   tables RWROP_RC/binomial(level='0');
   format RWROP_RC test2f.;
   where rwROP_DE=0;
run;
```

Sensitivity (95% CI)

Binomial Proportion	
rwROP_RC = RW-ROP +	
Proportion	0.8367
ASE	0.0528
95% Lower Conf Limit	0.7332
95% Upper Conf Limit	0.9402
Exact Conf Limits	
95% Lower Conf Limit	0.7034
95% Upper Conf Limit	0.9268

Specificity (95% CI)

Binomial Proportion	
rwROP_RC = RW-ROP -	
Proportion	0.8675
ASE	0.0276
95% Lower Conf Limit	0.8135
95% Upper Conf Limit	0.9216
Exact Conf Limits	
95% Lower Conf Limit	0.8029
95% Upper Conf Limit	0.9172

SAS Macro for 95% CI of Sensitivity and Specificity Using GEE

```
%macro gee(data=, de=, rc=);
proc genmod data=&data descending;
 class id &de;
 model &rc=&de/dist=bin;
 repeated subject=id/type=ind;
 estimate 'sens' intercept 1  &de 0 1/exp;
 estimate 'spec' intercept 1  &de 1 0/exp;
 ods output Genmod.Estimates=sensdata;
run;

data CI;
 set sensdata (rename=(LBetaestimate=estimate LBetaLowerCL=LowerCL LBetaUpperCL=UpperCL));
 if label='Exp(sens)' then do;
     Parameter='Sensitivity';
     point=estimate/(1+estimate);
     lower=lowerCL/(1+lowerCL);
     upper=upperCL/(1+upperCL); end;

 if label='Exp(spec)' then do;
     parameter='Specificity';
     point=1/(1+estimate);
     upper=1/(1+lowerCL);
     lower=1/(1+upperCL); end;
 if label in ('Exp(sens)', 'Exp(spec)');
run;

proc print data=ci noobs;
 var parameter point lower upper;
run;
%mend;

%gee(data=subsample, de=RWROP_de, rc=RWROP_rc);
```

adjusting for inter-eye correlation for data from two eyes under the same subject ID

Parameter	point	lower	upper
Sensitivity	0.83673	0.68985	0.92193
Specificity	0.86755	0.79329	0.91789

Per-eye analysis: Cluster Bootstrap

- A resampling technique for generating the distribution of sensitivity, specificity
- Taking a random sample of the same size as original sample **with replacement**
  - ➢ Stratified by number of eyes (0, 1, 2) with RW-ROP from clinical exam
  - ➢ Some subjects were selected in the same sample more than once, while some were never selected
- From bootstrapped sample, sensitivity and specificity are calculated
- Repeat process many times (e.g., 2000 times) to generate the distribution of sensitivity and speficity
- The 95% CI for sensitivity and specificity is derived based on 2.5th and 97.5th percentile

Per-eye analysis: Accounting for Inter-eye Correlation Using Cluster Bootstrap

```
%boot_sens(pdata=sub_person, pind=count, edata=subsample, b=2000);
```

95% CI for Sensitivity

Obs	sens95_low	sens_med	sens95_hi
1	71.4286	83.6735	93.8776

95% CI for Specificity

Obs	spec95_low	spec_med	spec95_hi
1	80.7947	86.7550	92.0530

Per-Infant analysis

- In telemedicine of ROP, if image evaluation found RW-ROP positive in either eye, the infant should be referred for clinical eye examination by ophthalmologist
- Desirable to calculate the sensitivity and specificity of image evaluation at infant level
- For infant level analysis, reduce eye-level data into infant level:
  - ➢ Infant RW-ROP present from eye examination if RW-ROP was present in **either** eye
  - ➢ Infant RW-ROP positive if image evaluation found RW-ROP in **either** eye
- Standard statistical methods can be applied for calculating sensitivity and specificity and their 95% CI

Per-Infant Analysis: Sensitivity and Specificity and 95% CIs

```
/** get 95% CI **/
proc freq data=left_right;
 tables
RWROP_RC_infant*RWROP_DE_infant/n
orow nocol nopercent;
run;

 /** get 95% CI **/
proc freq data=left_right;
 tables
RWROP_RC_infant/binomial(level=2)
;
 where rwROP_DE_infant=1;
run;

proc freq data=left_right;
 tables
RWROP_RC_infant/binomial(level=1)
;
 where rwROP_DE_infant=0;
run;
```

Sensitivity (95% CI)

RWROP_RC_infant	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	3.45	1	3.45
1	28	96.55	29	100.00

Binomial Proportion	
RWROP_RC_infant = 1	
Proportion	0.9655
ASE	0.0339
95% Lower Conf Limit	0.8991
95% Upper Conf Limit	1.0000
Exact Conf Limits	
95% Lower Conf Limit	0.8224
95% Upper Conf Limit	0.9991

Specificity (95% CI)

RWROP_RC_infant	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	61	85.92	61	85.92
1	10	14.08	71	100.00

Binomial Proportion	
RWROP_RC_infant = 0	
Proportion	0.8592
ASE	0.0413
95% Lower Conf Limit	0.7782
95% Upper Conf Limit	0.9401
Exact Conf Limits	
95% Lower Conf Limit	0.7562
95% Upper Conf Limit	0.9303

Example 3: 95% CI from Various Analysis Approaches

Analysis Approach	Sensitivity		Specificity	
Per-eye analysis	Estimate	Width of 95% CI	Estimate	Width of 95% CI
Ignoring inter-eye correlation	83.7%	20.7%	86.8%	10.8%
GEE	83.7%	23.2%	86.8%	12.5%
Cluster bootstrap	83.7%	22.5%	86.8%	11.3%
Left eye only	80.8%	35.2%	89.2%	25.4%
Right eye only	87.0%	30.8%	84.4%	17.3%
Per-infant analysis	96.6%	17.7%	85.9%	16.2%

Summary of Example 3

- In calculating 95% CI for sensitivity and specificity, ignoring inter-eye correlation leads to under-estimate their 95% CI (i.e., too narrow in 95% CI)
- Analyzing two eyes separately leads to different estimate of sensitivity and specificity, and makes their 95% CIs too wide
- GEE and cluster bootstrap can properly account for the inter-eye correlation

Example 4: ROC Analysis for Correlated Eye Data

Example 4: ROC analysis for AREDS Severity Scale

- Age-related Eye Disease Study Group (AREDS) developed 9-step AMD severity scale for predicting progression to advanced AMD
  - Based on drusen area and pigmentary abnormalities
  - Larger value indicates more severe AMD

- ROC analysis for performance of baseline AREDS severity scale for predicting 5-year incidence of advanced AMD
  - Completed 5-year followed-up
  - Eyes had baseline AREDS severity scale of 5 to 8
  - Random sample of 135 patients (198 eyes)
    - 63 patients (126 eyes) with both eyes eligible
    - 34 patients with one eye eligible because the fellow eye had a severity scale below 5
    - 38 patients with one eye eligible because the fellow eye had advanced AMD at baseline

Inter-eye Correlation in baseline AREDS severity scale

	Left Eye				
Right eye	5	6	7	8	Total
5	4	5	2	0	11
6	4	7	8	0	19
7	1	8	16	4	29
8	0	0	3	1	4
Total	9	20	29	5	63
	Percent agreement=28/63=44.4% Weight Kappa (95% CI)=0.33 (0.16, 0.49)				

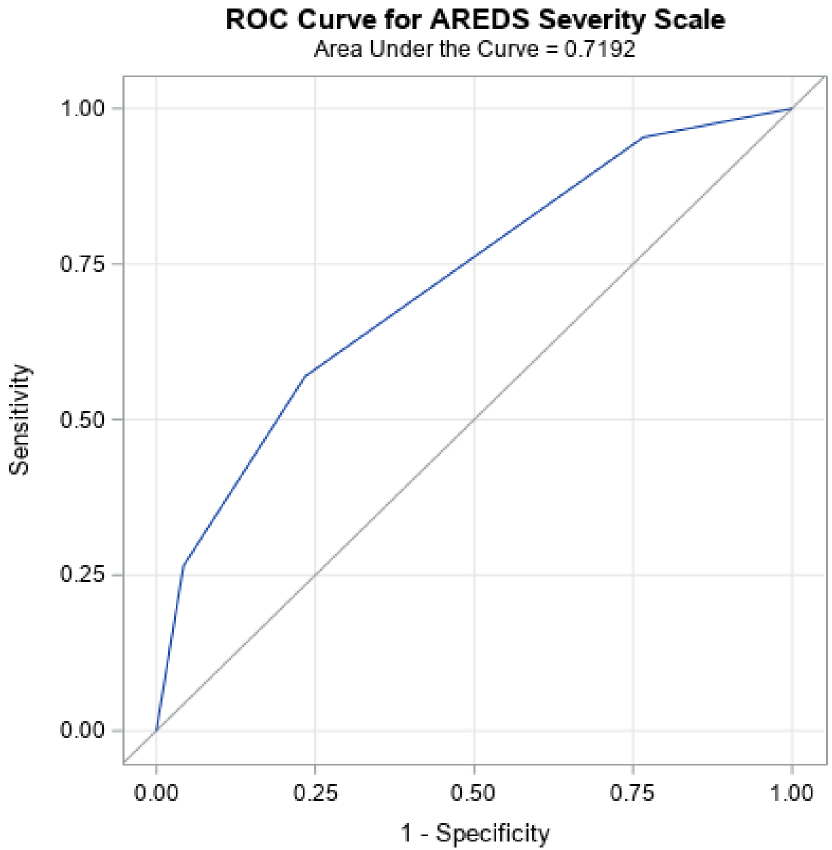
Inter-eye Correlation in 5-year advanced AMD

	Advanced AMD in Right Eye		
Advanced AMD in Left Eye	Absent	Present	Total
Absent	42 (66.7%)	9 (14.3%)	51 (81.0%)
Present	7 (11.1%)	5 (7.9%)	12 (19.1%)
Total	49 (77.8%)	14 (22.2%)	63
	Percent agreement = 47/63=74.6% Kappa (95% CI) = 0.23 (-0.05, 0.50)		

Risk of progression to advanced AMD in 5 years by baseline AREDS severity scale in each group of patients

	Bilateral patients (N=63 patients, 126 eyes)		**Unilateral patients where the fellow eye had severity scale <5 (N=34 patients, 34 eyes)**		**Unilateral patients where the fellow eye had advanced AMD (N=38 patients, 38 eyes)**	
Baseline AREDS Severity Scale	# of eyes	# of eyes progressing to advanced AMD in 5-year (%)	# of eyes	# of eyes progressed to advanced AMD in 5-year (%)	# of eyes	# of eyes progressing to advanced AMD in 5-year (%)
5	20	2 (10.0%)	19	0 (0.0%)	3	0 (0.0%)
6	39	6 (15.4%)	7	0 (0.0%)	9	3 (33.3%)
7	58	14 (24.1%)	6	2 (33.3%)	19	9 (47.4%)
8	9	4 (44.4%)	2	1 (50.0%)	7	6 (85.7%)
Total	126	26 (20.6%)	34	3 (8.8%)	38	18 (47.4%)

ROC Curve for AREDS scale Predicting 5-year Advanced AMD



Naïve ROC Analysis Using Standard Logistic Regression

```
proc logistic data=advAMD5yr_eye_elig_sub;
 class scale0;
 model advAMD5yr=scale0;
 ROC "ROC for Predicting 5-year GA using AREDS Severity Scale" scale0;
run;
```

ROC Association Statistics

ROC Model	Mann-Whitney Area	Standard Error	95% Wald Confidence Limits		Somers' D	Gamma	Tau-a
Model	0.7192	0.0381	0.6446	0.7939	0.4385	0.6131	0.1596
ROC for Predicting 5-year GA using AREDS Severity Scale	0.7192	0.0381	0.6446	0.7939	0.4385	0.6131	0.1596

Cluster Bootstrap for AUC

- Taking a random sample of the same sample size as original sample **with replacement**
- From bootstrapped sample, calculate the AUC from the logistic regression model
- Repeat process many times (e.g., 2000 times) to generate the distribution of AUC
- The 95% CI for AUC is derived based on 2.5th and 97.5th percentile

Nonparametric Clustered ROC analysis

- Developed by Obuchowski for estimating variance of the AUC from clustered data (*Biometrics*, 1997)
- Based on the concept of design effect and effective sample size used in the analysis of data from sample surveys
- Nonparametric, not require specification of the intra-cluster correlation structure
- R functions are available at https://www.lerner.ccf.org/qhs/software/roc_analysis.php

AUC from Various Approaches

Analysis Approach			
Two Eyes Analysis	**AUC**	**95% CI**	**Width of 95% CI**
Ignoring inter-eye correlation	0.719	0.645, 0.794	0.149
Cluster bootstrap	0.722	0.641, 0.793	0.152
Nonparametric clustered ROC analysis	0.719	0.641, 0.797	0.156
Left Eye Analysis (N=102)			
Simple logistic regression	0.691	0.583, 0.801	0.218
Right Eye Analysis (N=96)			
Simple logistic regression	0.745	0.643, 0.848	0.205

Summary of Example 4

- In ROC analysis, ignoring the inter-eye correlation makes 95% CI for AUC too narrow
- Analyzing two eyes separately is not efficient
- Cluster bootstrap and the nonparametric clustered ROC analysis can properly account for the inter-eye correlation

Summary

- When data from two eyes of a subject are available, statistical analysis should consider the unit of analysis (per-eye or per-subject)
- Inter-eye correlation should be accounted for at per-eye analysis
- Several statistical methods (mixed effects model, GEE, cluster bootstrap etc.) available to properly account for the inter-eye correlation
  - Provide similar results

Summary (Cont'd)

- Ignoring inter-eye correlation leads to invalid statistical inference
- Its impact depends on the degree of inter-eye correlation and membership
  - ➢ When two eyes are in different comparison group, ignoring inter-eye correlation leads to over-estimate of variance, 95% CI and p-value
  - ➢ When two eyes are in the same comparison group, ignoring inter-eye correlation leads to under-estimate of variance, 95% CI and p-value
  - ➢ Ignoring the inter-eye correlation makes the 95% CIs of sensitivity, specificity and AUC too narrower

References

1. Ying GS, Maguire MG, Glynn R, Rosner. Tutorial on Biostatistics: Statistical Analysis for Correlated Binary Eye Data. *Ophthalmic Epidemiol*. 2018 Feb;25(1):1-12.
2. Ying GS, Maguire MG, Glynn R, Rosner. Tutorial on Biostatistics: Linear Regression Analysis of Continuous Correlated Eye Data. *Ophthalmic Epidemiol*. 2017 Apr;24(2):130-140.
3. Ying GS, Maguire MG, Glynn RJ, Rosner B. Tutorial on Biostatistics: Longitudinal Analysis of Correlated Continuous Eye Data. *Ophthalmic Epidemiol*. 2020 Aug 2:1-18.
4. Ying GS, Maguire MG, Glynn RJ, Rosner B. Calculating Sensitivity, Specificity, and Predictive Values for Correlated Eye Data. *Invest Ophthalmol Vis Sci*. 2020 Sep 1;61(11):29.
5. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 1986;42 (1):121–130.
6. Obuchowski NA. Nonparametric analysis of clustered ROC curve data. *Biometrics.* 1997;53(2):567-578.
7. Huang FL. Using Cluster Bootstrapping to Analyze Nested Data With a Few Clusters. *Educ Psychol Meas.* 2018;78(2):297-318.

Thank You