BIOMETRICS 38, 105–114
March 1982

Statistical Methods in Ophthalmology: An Adjustment for the Intraclass Correlation between Eyes

Bernard Rosner

Channing Laboratory, Department of Preventive Medicine and Clinical Epidemiology, Harvard Medical School and Peter Bent Brigham Hospital Division of Brigham and Women's Hospital, 180 Longwood Avenue, Boston, Massachusetts 02115, U.S.A.

SUMMARY

For the cases of normally- and binomially-distributed outcome variables, methods are presented for analyzing ophthalmologic data to which a person may have contributed two eyes worth of information, the values from the two eyes being highly correlated. A frequently-used method of analysis, where each eye is treated as an independent random variable, is shown to be invalid in the presence of intraclass correlation: it yields true p-values two to six times as large as nominal p-values when realistic assumptions are made about the degree of correlation between eyes. These results may be applicable to other medical specialities, such as otolaryngology, where highly-correlated replicate observations are obtained from individuals.

1. Introduction

In ophthalmologic studies, the fundamental unit for statistical analysis is often the *eye* rather than the *person*. Frequently, descriptive statistics are computed over distributions of eyes, and tests of hypotheses are constructed to compare such distributions, as for intraocular pressure measurements (Armaly, 1965), or for measurements of refractive error (Dunphy, Stoll and King, 1968; Sorsby *et al.*, 1960). If only one eye is used for a given person in any such distribution, as might be the case in a clinical trial where for a given person one eye is used as the treated eye and the other as the control, then this practice is entirely appropriate and standard methods of estimation and hypothesis testing are valid. However, if the purpose is to compare two different types of people on some finding in an ocular examination, as would occur in a typical epidemiological investigation such as a comparison of intraocular pressures in persons in different age groups, then an individual contributes *two* eyes worth of information to such an analysis, their values being generally highly correlated. If the values are highly correlated, then methods of analysis in which each eye is considered as an independent random variable are not valid. It is the purpose of this paper to explore the consequences of the use of such methods of analysis on ophthalmologic data and to recommend an improved method which takes into account the intraclass correlation between eyes of the same person.

2. Normally-Distributed Outcome Variable

2.1 *Intraclass Correlation Model*

Suppose we wish to compare g groups of persons on some ocular finding y, where there are P_i persons in the ith group, $i = 1, \ldots, g$; $P \equiv \sum P_i$ persons over all groups; and each

Key words: Ophthalmology; Analysis of variance; Nested design; Intraclass correlation; Retinitis pigmentosa.

Table 1
Nested ANOVA data layout under the intraclass correlation model

Source of variation	Sum of squares	Degrees of freedom (df)	Mean square	Expected mean square
Between groups	$\sum P_i(\bar{y}_{i..}-\bar{y}_{...})^2 = SSG$	$g-1$	$SSG/(g-1) = MSG$	$\sum_{i=1}^{g} P_i(\alpha_i-\bar{\alpha})^2/(g-1) + \sigma_\beta^2 + (\sigma^2/\bar{N})$
Between persons within groups	$\sum\sum(\bar{y}_{ij.}-\bar{y}_{i..})^2 = SSP$	$P-g$	$SSP/(P-g) = MSP$	$\sigma_\beta^2 + (\sigma^2/\bar{N})$
Between eyes within persons	$\sum\sum\sum(y_{ijk}-\bar{y}_{ij.})^2 = SSE$	$N-P$	$SSE/(N-P) = MSE$	σ^2

person contributes N_{ij} eyes to the analysis where $N_{ij} = 1$ or 2, $i = 1, \ldots, g$, $j = 1, \ldots, P_i$; and $N \equiv \sum\sum N_{ij}$ = total number of eyes over all persons over all groups. An appropriate model for such a design is given by a nested mixed effects analysis of variance (ANOVA) model of the form:

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + e_{ijk}, \qquad i = 1, \ldots, g, \quad j = 1, \ldots, P_i, \quad k = 1, \ldots, N_{ij}, \tag{2.1}$$

where $\beta_{ij} \sim N(0, \sigma_\beta^2)$, $e_{ijk} \sim N(0, \sigma^2)$ and μ, $\{\alpha_i\}$ are constants. Thus, the group effect represented by $\{\alpha_i,\ i = 1, \ldots, g\}$ is considered to be a fixed effect, while the effect of persons within groups represented by $\{\beta_{ij},\ i = 1, \ldots, g,\ j = 1, \ldots, P_i\}$ and of eyes within persons represented by $\{e_{ijk},\ i = 1, \ldots, g,\ j = 1, \ldots, P_i,\ k = 1, \ldots, N_{ij}\}$ are considered to be random effects. The number of eyes contributed by different individuals might not be the same, due to either missing values or the presence in only one eye of the condition under study.

An exact treatment of this problem is difficult due to the unbalanced nature of the design whereby different persons contributed either one or two eyes to the analysis. However, the method of unweighted means is considered a reasonable approximate method for analyzing such data when $\max(N_{ij}^{-\frac{1}{2}})/\min(N_{ij}^{-\frac{1}{2}}) \leq 2$ (Searle, 1971, p. 367). This criterion will always be satisfied for ophthalmologic data since a person must contribute data from either one or two eyes to the analysis, and thus $\max(N_{ij}^{-\frac{1}{2}})/\min(N_{ij}^{-\frac{1}{2}}) = 1/2^{-\frac{1}{2}} = 2^{\frac{1}{2}} \leq 2$. Thus, if we assume that each person contributes approximately $\bar{N} = [\{\sum\sum(1/N_{ij})\}/P]^{-1}$ eyes to the analysis, then we have the nested ANOVA data layout given in Table 1, where $\bar{y}_{ij.} = \sum_k y_{ijk}/N_{ij}$, $\bar{y}_{i..} = \sum_j \bar{y}_{ij.}/P_i$, $\bar{y}_{...} = \sum P_i\bar{y}_{i..}/P$.

An appropriate test procedure to test the hypothesis H_0: all α_i are equal versus H_1: some of the α_i are unequal, is given by computing the test statistic $\lambda = MSG/MSP$ which follows an $F_{g-1,P-g}$ distribution under H_0, rejecting H_0 if $\lambda > F_{g-1,P-g,1-\alpha} \equiv$ 100(1-α)% percentile of an $F_{g-1,P-g}$ distribution, and accepting H_0 otherwise. The intraclass correlation ρ can also be estimated from Table 1 and is given by $\hat{\rho} = \hat{\sigma}_\beta^2/(\hat{\sigma}_\beta^2 + \hat{\sigma}^2)$ where $\hat{\sigma}_\beta^2 = \max\{0,\ MSP\text{-}(MSE/\bar{N})\}$, $\hat{\sigma}^2 = MSE$. If the above null hypothesis is rejected, then comparisons of specific groups can be accomplished using ordinary ANOVA t tests where for a comparison of Groups i_1 and i_2 we compute the test statistic $u_{i_1,i_2} = (\bar{y}_{i_1..} - \bar{y}_{i_2..})/\{MSP(P_{i_1}^{-1} + P_{i_2}^{-1})\}^{\frac{1}{2}}$ and reject if $|u_{i_1,i_2}| > t_{P-g,1-\frac{1}{2}\alpha}$, $i_1, i_2 = 1, \ldots, g$, $i_1 \neq i_2$. Alternatively, one of several multiple comparisons procedures can be used for this purpose.

2.2 *Independence Model*

If we assume that two eyes from the same person are independent random variables, then we have an ordinary one-way ANOVA model of the form:

$$y_{ijk} = \mu^* + \alpha_i^* + e_{ijk}^*, \tag{2.2}$$

where $i = 1, \ldots, g$, $j = 1, \ldots, P_i$, $k = 1, \ldots, N_{ij}$, $e_{ijk}^* \sim N(0, \sigma^{*2})$, and μ^*, $\{\alpha_i^*\}$ are constants. If we let $N_{i.} = \sum_j N_{ij}$, $i = 1, \ldots, g$, $\bar{y}_{i..}^* = \sum_j \sum_k y_{ijk}/N_{i.}$, $\bar{y}_{...}^* = \sum N_{i.}\bar{y}_{i..}^*/N$, then the data layout for this model is as given in Table 2.

An appropriate test procedure under the model (2.2), to test the hypothesis H_0: all α_i^* are equal versus H_1: some of the α_i^* are unequal, $i = 1, \ldots, g$, is given by computing the test statistic $\lambda^* = MSG^*/MSE^*$ which follows an $F_{g-1,N-g}$ distribution under H_0 and rejecting H_0 if $\lambda^* > F_{g-1,N-g,1-\alpha}$.

2.3 *Performance of the Independence Model under the Assumption of Dependence between Eyes*

The question arises as to the distribution of the test statistic λ^* under the intraclass correlation model (2.1). We have studied this question under the simplifying assumption that $N_{ij} = 2$, $i = 1, \ldots, g$, $j = 1, \ldots, P_i$, i.e. each person contributes two eyes to the analysis. It is shown in the Appendix that this distribution is given by

$$\lambda^*(g, P, \rho) \sim \{(1+\rho)U/(g-1)\}/[\{(1-\rho)V + (1+\rho)W\}/(2P-g)], \tag{2.3}$$

where $U \sim \chi^2_{g-1}$, $V \sim \chi^2_P$, $W \sim \chi^2_{P-g}$, and U, V, W are independent random variables. We wish to study the true α-level of the test procedure appropriate for (2.2), namely $\alpha^*(g, P, \rho, \alpha) = \text{pr}\{\lambda^*(g, P, \rho) > F_{g-1,2P-g,1-\alpha}\}$, under the model (2.1). The exact distribution of λ^* is difficult to obtain for $\rho < 1$. However, since P is generally large relative to g, good lower and upper bounds to α^* are given by approximating the distribution of the denominator of (2.3) by $2\chi^2_P/(2P-g)$ and $2\chi^2_{P-g}/(2P-g)$, respectively, thus yielding the following:

$$\alpha_1^*(g, P, \rho, \alpha) < \alpha^*(g, P, \rho, \alpha) < \alpha_2^*(g, P, \rho, \alpha), \tag{2.4}$$

where

$$\alpha_1^*(g, P, \rho, \alpha) = \text{pr}\{F_{g-1,P} > \frac{2P}{(1+\rho)(2P-g)} F_{g-1,2P-g,1-\alpha}\}$$

and

$$\alpha_2^*(g, P, \rho, \alpha) = \text{pr}\{F_{g-1,P-g} > \frac{2(P-g)}{(1+\rho)(2P-g)} F_{g-1,2P-g,1-\alpha}\}.$$

Furthermore, for $\rho = 1$, it follows immediately from (2.3) that α^* is given exactly by

Table 2
Nested ANOVA data layout under the independence model

Source of variation	Sum of squares	Degrees of freedom (df)	Mean square
Between groups	$\sum N_{i.}(\bar{y}_{i..}^* - \bar{y}_{...}^*)^2 = SSG^*$	$g-1$	$SSG^*/(g-1) = MSG^*$
Within groups	$\sum\sum\sum (y_{ijk} - \bar{y}_{i..}^*)^2 = SSE^*$	$N-g$	$SSE^*/(N-g) = MSE^*$

Table 3

Bounds on exact p-values for λ^* *under the intraclass correlation model*; $\alpha = .05$, $g = 2(1)5$, $P = 25, 50, 100, 200, \infty$, $\rho = 0(.2)1$

		P									
g	ρ	25		50		100		200		∞	
		α_1^*	α_2^*	α_1^*	α_2^*	α_1^*	α_2^*	α_1^*	α_2^*	α_1^*	α_2^*
2	0.0	.051	.061	.050	.055	.050	.053	.050	.051	.050	.050
	0.2	.073	.086	.073	.079	.073	.076	.073	.075	.074	.074
	0.4	.095	.110	.096	.103	.097	.100	.097	.099	.098	.098
	0.6	.117	.133	.119	.127	.120	.124	.121	.123	.121	.121
	0.8	.139	.156	.141	.150	.143	.147	.143	.145	.144	.144
	1.0		.177		.171		.168		.167		.166
3	0.0	.049	.071	.050	.060	.050	.055	.050	.052	.050	.050
	0.2	.078	.106	.080	.093	.081	.088	.082	.085	.082	.082
	0.4	.109	.142	.113	.129	.115	.123	.117	.120	.118	.118
	0.6	.141	.178	.147	.165	.150	.159	.152	.157	.154	.154
	0.8	.172	.213	.181	.200	.185	.195	.187	.192	.189	.189
	1.0		.246		.234		.229		.226		.223
4	0.0	.047	.082	.049	.064	.049	.057	.050	.053	.050	.050
	0.2	.079	.126	.084	.106	.087	.097	.088	.093	.089	.089
	0.4	.116	.173	.125	.152	.129	.142	.131	.138	.134	.134
	0.6	.154	.219	.167	.198	.174	.189	.177	.185	.180	.180
	0.8	.194	.264	.210	.244	.218	.235	.223	.231	.227	.227
	1.0		.307		.288		.279		.275		.272
5	0.0	.044	.095	.047	.070	.049	.059	.049	.054	.050	.050
	0.2	.078	.148	.087	.119	.091	.106	.093	.101	.095	.095
	0.4	.118	.204	.133	.174	.141	.160	.144	.154	.148	.148
	0.6	.162	.260	.183	.230	.194	.217	.199	.210	.204	.204
	0.8	.207	.313	.234	.285	.247	.272	.254	.266	.260	.260
	1.0		.364		.337		.326		.320		.314

$\alpha_2^*(g, P, 1, \alpha)$. We have evaluated $\alpha_1^*(g, P, \rho, \alpha)$, $\alpha_2^*(g, P, \rho, \alpha)$ for $g = 2(1)5$; $P = 25, 50, 100, 200, \infty$; $\rho = 0(.2).8$; $\alpha = .05$, and $\alpha_2^*(g, P, 1, \alpha)$ for the same values of g, P and α. We used the IMSL subroutines MDFD and MDFI to compute the cumulative probability distribution and the exact percentage points of the F distribution for finite P, and the corresponding subroutines MDCH and MDCHI for the chi square distribution for $P = \infty$ (International Mathematical and Statistical Libraries, 1979). The results are given in Table 3.

There is clearly a substantial discrepancy between the nominal and actual p-values with the discrepancy increasing as g and ρ increase and decreasing slightly as P increases. It is the author's experience that the estimated intraclass correlation between eyes in data derived from ocular examinations is at least .4; from Table 3 the true p-value is about two to three times the nominal level for $\rho = .4$ and three to six times the nominal level for $\rho = 1$.

Computations similar to those in Table 3 have been performed for nominal α-levels of .01 and .001, with discrepancies between nominal and actual p-values that are at least as large as for $\alpha = .05$. Furthermore, computations similar to those in Table 3 have been performed for the case where each person contributed $\bar{N}$ eyes to the analyses, $1 < \bar{N} < 2$. The true and nominal α-levels are still substantially different but the discrepancy becomes

smaller as $\bar{N}$ approaches 1. For example, if $g=2$, $P=25$ and $\rho=1$, then $\alpha^*=.177$ for $\bar{N}=2$, .154 for $\bar{N}=1.8$, .129 for $\bar{N}=1.6$, .103 for $\bar{N}=1.4$, and .076 for $\bar{N}=1.2$.

3. Binomially-Distributed Outcome Variable

3.1 *Theory*

Let $z_{ijk}=1$ if the kth eye of the jth person in the ith group is affected, and 0 otherwise, $i=1,\ldots,g$, $j=1,\ldots,P_i$, $k=1,2$. We shall assume that

$$\text{pr}(z_{ijk}=1)=\lambda_i,\ \text{pr}(z_{ijk}=1 \mid z_{ij,3-k}=1)=R\lambda_i, \tag{3.1}$$

$i=1,\ldots,g$, $j=1,\ldots,P_i$, $k=1,2$, for some positive constant R. The constant R is a measure of dependence between two eyes of the same person. If $R=1$, then the two eyes are completely independent, while if $R\lambda_i=1$, then the eyes are completely dependent. We wish to test the hypothesis H_0: $\lambda_1=\lambda_2=\cdots=\lambda_g=\lambda$ versus H_1: some of the λ_i are unequal.

Let P_{ij} = number of persons in the ith group with exactly j affected eyes, $i=1,\ldots,g$, $j=0,1,2$. An appropriate test statistic for the above hypothesis is given by $T=\sum(\hat{\lambda}_i-\bar{\lambda})^2/\text{var}(\hat{\lambda}_i)\sim\chi^2_{g-1}$ under H_0, where $\hat{\lambda}_i=\frac{1}{2}(P_{i1}+2P_{i2})/P_i$, $i=1,\ldots,g$, $\bar{\lambda}=\frac{1}{2}\sum(P_{i1}+2P_{i2})/P$, and we reject H_0 if $T>\chi^2_{g-1,1-\alpha}=100(1-\alpha)\%$ percentile of a χ^2_{g-1} distribution. If z_{ij1}, z_{ij2} are independent random variables, $i=1,\ldots,g$, $j=1,\ldots,P_i$, then under H_0, $\text{var}(\hat{\lambda}_i)=\frac{1}{2}\lambda(1-\lambda)/P_i$, $i=1,\ldots,g$, while if they are completely dependent, $\text{var}(\hat{\lambda}_i)=\lambda(1-\lambda)/P_i$, $i=1,\ldots,g$. In general, under the model (3.1), we have that $\hat{\lambda}_i=\frac{1}{2}\sum_j\sum_k z_{ijk}/P_i$, and thus, under H_0,

$$\begin{aligned}\text{var}(\hat{\lambda}_i)&=\tfrac{1}{4}\sum_j \text{var}(\sum_k z_{ijk})/P_i^2\\&=\tfrac{1}{4}\sum_j\{\text{var}(z_{ij1})+\text{var}(z_{ij2})+2\,\text{cov}(z_{ij1},z_{ij2})\}/P_i^2\\&=\tfrac{1}{2}\{\lambda(1-\lambda)+(R-1)\lambda^2\}/P_i\\&=\lambda(1-\lambda)/(eP_i),\end{aligned}$$

where $e=2\lambda(1-\lambda)/\{\lambda(1-\lambda)+(R-1)\lambda^2\}$.

We interpret e as the 'effective number of eyes per person' since if each person contributed e independent eyes to the analysis, then $\text{var}(\hat{\lambda}_i)=\lambda(1-\lambda)/(eP_i)$. Note that $e=2$ under complete independence ($R=1$) and $e=1$ under complete dependence ($R\lambda=1$). We estimate e by $\hat{e}=2\hat{\lambda}(1-\hat{\lambda})/\{\hat{\lambda}(1-\hat{\lambda})+(\hat{R}-1)\hat{\lambda}^2\}$, where $\hat{\lambda}$, $\hat{R}$ are the maximum likelihood estimators of λ, R, respectively, under H_0. It is straightforward but tedious to show that $\hat{\lambda}=\bar{\lambda}$, $\hat{R}=4P\sum P_{i2}/(\sum P_{i1}+2\sum P_{i2})^2$. Thus, an approximate level-α test for the above hypothesis is given by computing the test statistic

$$T=[\hat{e}/\{\bar{\lambda}(1-\bar{\lambda})\}]\sum P_i(\hat{\lambda}_i-\bar{\lambda})^2, \tag{3.2}$$

which $\sim\chi^2_{g-1}$ under H_0, and rejecting H_0 if $T>\chi^2_{g-1,1-\alpha}$.

3.2 *Performance of the Independence Model under the Assumption of Dependence between Eyes*

A frequently-used procedure in this case would be to assume that the outcomes for the two eyes of an individual are independent random variables, in which case we would use the test statistic

$$T_2=[2/\{\bar{\lambda}(1-\bar{\lambda})\}]\sum P_i(\hat{\lambda}_i-\bar{\lambda})^2. \tag{3.3}$$

Table 4
Exact p-values for T_2; $g = 2(1)5$, $e = 1.0(0.2)2.0$, $\alpha = .05$

g	e					
	1.0	1.2	1.4	1.6	1.8	2.0
2	.166	.129	.101	.080	.063	.050
3	.223	.166	.123	.091	.067	.050
4	.272	.196	.140	.100	.071	.050
5	.314	.223	.156	.108	.074	.050

We wish to study the behaviour of T_2 if the outcomes on the two eyes are in fact dependent, i.e. if the effective number of eyes per person, e, is less than 2. We see that, in this case,

$$T_2 = (2/e)[e/\{\bar{\lambda}(1-\bar{\lambda})\}]\sum P_i(\hat{\lambda}_i - \bar{\lambda})^2 \sim (2/e)\chi^2_{g-1}.$$

Thus, the true α-level for T_2 is given by

$$\alpha^* = \text{pr}(T_2 > \chi^2_{g-1,1-\alpha}) = \text{pr}\{\chi^2_{g-1} > (e/2)\chi^2_{g-1,1-\alpha}\}. \tag{3.4}$$

We have evaluated α^* using the IMSL subroutines MDCH and MDCHI (International Mathematical and Statistical Libraries, 1979) for $g = 2(1)5$, $e = 1.0(0.2)2.0$, $\alpha = .05$, and present these results in Table 4.

There clearly is a substantial discrepancy between the true and nominal p-values which gets larger as e decreases and g increases. If the two eyes are completely dependent ($e = 1$), then the true α-levels are three to six times the nominal level of .05.

4. Examples

We now present examples of the use of these methods on a data set obtained from an outpatient population of 218 persons aged 20–39 with retinitis pigmentosa (RP) who were seen at the Massachusetts Eye and Ear Infirmary from 1970 to 1979. The patients were classified on the basis of a detailed family history into the genetic types of autosomal dominant RP (DOM), autosomal recessive RP (AR), sex-linked RP (SL), and isolate RP (ISO) for a study of differences between these four groups on certain measurements made in a routine ocular examination. In order to simplify the analysis, only one person from this age group was selected from each family, and if more than one affected person was available for analysis, then a randomly-selected affected person from this age group was chosen. Thus, the 218 persons were from 218 unique families. The details of the design of this study and the procedures for genetic classification are given by Berson, Rosner and Simonoff (1980).

We first present an analysis of the difference between groups for spherical refractive error using the methods of §2.1. The sample used for this analysis consists of the subgroup of 212 persons who had information on spherical refractive error for at least one eye. Of these persons, 210 had information for both eyes while two had information for only one eye. All refractive errors were determined with retinoscopy after cycloplegia.

We first present the analysis of the data on a per-person basis using the intraclass correlation model (2.1). There were 28 persons in the DOM group, 20 persons in the AR group, 18 persons in the SL group, and 146 persons in the ISO group. The ANOVA results are given in Table 5.

Table 5

ANOVA results comparing spherical refractive error for different genetic types of RP using the intraclass correlation model (2.1)

(a) Overall ANOVA table

Source of variation	Sum of squares	df	Mean square	F statistic	p-value
Between groups	133.59	3	44.53	3.68	.013
Between persons within groups	2518.45	208	12.11		
Within persons	80.49	210	0.383		

(b) Values of t statistics and p-values for comparisons of specific groups

Group, i	Mean, $\bar{y}_{i..}$	Estimated standard error, $\{MSP/P_i\}^{\frac{1}{2}}$	Number of persons, P_i	Comparison group DOM	AR	SL	ISO
DOM	+0.127	0.658	28	—	0.941 (*NS*)	3.259 ($p = .001$)	1.350 (*NS*)
AR	−0.831	0.778	20		—	2.183 ($p = .030$)	0.013 (*NS*)
SL	−3.299	0.820	18			—	−2.826 ($p = .005$)
ISO	−0.842	0.288	146				—

There was a significant difference between the refractive errors of the four groups ($p = .013$). We then performed t tests as given in §2.1 for comparing specific pairs of groups. The results were that the overall significant difference in Table 5 could be wholly attributed to significant differences between SL and the other three groups. The estimated intraclass correlation over all groups was .969!

We now present the analysis of the data on a per-eye basis using the independence model (2.2). There was a total of 422 eyes, comprised of 54 eyes in the DOM group, 40 eyes in the AR group, 36 eyes in the SL group and 292 eyes in the ISO group. The ANOVA results are presented in Table 6.

The was a highly significant difference between the refractive errors of the four groups ($p = .00026$). The comparable p-value in Table 5 (.013) representing the overall comparison of groups was 50 times as large as this p-value. We also performed standard one-way ANOVA t tests comparing specific pairs of groups. All the significant comparisons in Table 5 became considerably more significant when assessed on a per-eye basis in Table 6. Furthermore, one nonsignificant comparison between the DOM and ISO groups when assessed on a per-person basis in Table 5 became significant ($p = .016$) when assessed on a per-eye basis in Table 6.

We next present an analysis of the difference between groups for best corrected Snellen visual acuity (VA) using the methods of §3.1. An eye was considered affected if VA was 20/50 or worse, and normal if VA was 20/40 or better. The sample used for this analysis consists of the subgroup of 216 persons out of the sample of 218 persons each of whom had complete information for VA on both eyes.

The data were first analyzed on a per-person basis using the model (3.1). The distribution of the number of affected eyes for the persons in each genetic type is given in Table 7.

Table 6

ANOVA results comparing spherical refractive error for different genetic types of RP using the independence model (2.2)

(a) Overall ANOVA table

Source of variation	Sum of squares	df	Mean square	F statistic	p-value
Between groups	297.64	3	99.21	8.45	.00026
Within groups	4906.83	418	11.73		

(b) Values of t statistics and p-values for comparisons of specific groups

Group, i	Mean, $\bar{y}_{i..}$	Estimated standard error, $(MSE^*/N_{i.})^{\frac{1}{2}}$	Number of eyes, $N_{i.}$	Comparison group DOM	AR	SL	ISO
DOM	+0.386	0.466	54	—	1.704 (*NS*)	4.999 ($p<.001$)	2.421 ($p=.016$)
AR	−0.831	0.542	40		—	3.135 ($p=.002$)	0.018 (*NS*)
SL	−3.299	0.571	36			—	−4.059 ($p<.001$)
ISO	−0.842	0.201	292				—

We find that the effective number of eyes per person is 1.207 and there is an overall significant difference between the proportion of affected eyes in the four groups ($p=.010$). Standard methods for decomposition of chi square (Maxwell, 1961, p.52) show that this overall difference can be completely attributed to differences between the SL group and each of the other three groups.

The data were then analyzed on a per-eye basis assuming independence between eyes, using the model (3.3). The test statistic for the overall comparison of groups is then $T_2=[2/\{\bar{\lambda}(1-\bar{\lambda})\}]\sum P_i(\lambda_i-\bar{\lambda})^2=18.82\sim\chi^2_3$, $p=.00030$. Thus, the true p-value (.010) is approximately 30 times as large as the nominal p-value (.00030). If we use standard methods of decomposition of chi square, then we find that this overall difference can be

Table 7

Distribution of the number of affected eyes for persons in each genetic type

	P_{i0}	P_{i1}	P_{i2}	P_i	$\hat{\lambda}_i$
DOM	15	6	7	28	.357
AR	7	5	9	21	.548
SL	3	2	14	19	.789
ISO	67	24	57	148	.466
	92	37	87	216	.488

$\hat{R}=1.688$, $\hat{e}=1.207$,

$T=\{1.207/(.488\times.512)\}\sum_{i=1}^{4} P_i(\hat{\lambda}_i-.488)^2$
$=11.36\sim\chi^2_3$,
$p=.010$

completely attributed to differences between the SL group and the other groups. However, the p-value for this comparison is again much smaller ($\chi_1^2 = 15.11$, $p = .00010$) when the data are analyzed on a per-eye basis, than when they are analyzed on a per-person basis ($\chi_1^2 = 9.12$, $p = .0025$). Finally, in this example, a person who provides two eyes to the analysis is contributing approximately 1.2 independent eyes worth of information!

5. Discussion

We have presented models for normally- and binomially-distributed outcome variables which can be used to compare groups of persons on some finding in an ocular examination in which each person contributes two eyes worth of information to the analysis, the values of which may be highly correlated. It is shown that standard methods of analysis whereby the data are analyzed as if two eyes from the same person are independent random variables are not valid and can result in p-values which are from two to six times the nominal level. This method should probably never be used if the intraclass correlation between eyes, ρ, is at least .4 in the normally-distributed case, or if the effective number of eyes per person, e, is not greater than 1.6 in the binomially-distributed case. Note that in our two examples $\hat{\rho} = .969$ for spherical refractive error and $\hat{e} = 1.208$ for best corrected Snellen visual acuity of 20/50 or worse.

Another method of analysis that has been discussed is to tabulate results separately for the left and right eye and compare results (Ederer, 1973). This method is valid but will be less efficient than the methods described here which combine information over both eyes in the analysis.

The results in this paper have been discussed specifically in terms of data derived from an ocular examination. However, these results may be applicable to the analysis of data from other areas of medicine such as otolaryngology, where highly-correlated observations are obtained from two ears of an individual and it is desirable to use all the information in the analysis. Furthermore, it would be desirable to extend these methods to the case where there are more than two highly-correlated observations for a given individual.

Acknowledgements

The author wishes to thank Dr Eliot L. Berson, Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, for permission to use information from his retinitis pigmentosa outpatient population data base for the example in this paper. This work has been supported in part by National Heart, Lung and Blood Institute Training Grant HL07427.

Résumé

Dans le cas de variables résultats de distribution normale ou binomiale, on présente des méthodes pour analyser des données ophtalmologiques pour lesquelles une personne peut avoir fourni de l'information sur ses deux yeus, les valeurs obtenues à partir des deux yeux étant fortement correlées. On montre qu'une méthode d'analyse couramment utilisée, où chaque oeil est considéré comme une variable aléatoire indépendante, est invalidée par la présence de corrélations intraclasses: elle fournit des vraies valeurs de p deux à six fois aussi grandes que les valeurs nominales de p quand on fait des hypothèses réalistes sur le degré de corrélation entre les yeux.

Ces résultats sont applicables à d'autres spécialités médicales, telles que l'otorhinolaryngologie, où l'on obtient des sujets des observations repétées hautement corrélées.

References

Armaly, N. F. (1965). On the distribution of applanation pressure. I. Statistical features and the effect of age, sex, and family history of glaucoma. *Archives of Ophthalmology* **73,** 11-18.

Berson, E. L., Rosner, B. and Simonoff, E. (1980). An outpatient population of retinitis pigmentosa and their normal relatives: Risk factors for genetic typing and detection derived from their ocular examinations. *American Journal of Ophthalmology* **89,** 763-775.

Dunphy, F. B., Stoll, M. R. and King, S. H. (1968). Myopia among American male graduate students. *American Journal of Ophthalmology* **65,** 518-521.

Ederer, F. (1973). Shall we count numbers of eyes or numbers of subjects? *Archives of Ophthalmology* **89,** 1-2.

International Mathematical and Statistical Libraries (1979). IMSL Library Reference Manual, Vol. 2, 7th ed. Houston: IMSL.

Maxwell, A. E. (1961). *Analyzing Qualitative Data.* London: Methuen.

Searle, S. R. (1971). *Linear Models.* New York: Wiley.

Sorsby, A., Sheridan, M., Leary, G. A. and Benjamin, B. (1960). Vision, visual acuity, and ocular refraction of young men. Findings in a sample of 1,033 subjects. *British Medical Journal* **2,** 1394-1398.

Received August 1980; *revised February* 1981

Appendix

Derivation of the Distribution of λ^* under the Intraclass Correlation Model

It follows immediately from (2.1) that, under H_0,

$$MSG^* \sim (2\sigma_\beta^2+\sigma^2)U/(g-1),$$

where $U \sim \chi^2_{g-1}$. We can write

$$\begin{aligned} MSE^* &= \{\sum\sum\sum (y_{ijk}-\bar{y}_{ij.})^2 + 2\sum\sum (\bar{y}_{ij.}-\bar{y}_{i..})^2\}/(2P-g) \\ &\sim \{\sigma^2 V + (2\sigma_\beta^2+\sigma^2)W\}/(2P-g), \end{aligned}$$

where $V \sim \chi^2_P$, $W \sim \chi^2_{P-g}$ and U, V, W are independent random variables. Thus, upon reparametrizing λ^* in terms of $\rho = \sigma_\beta^2/(\sigma_\beta^2+\sigma^2)$ we have that

$$\lambda^* = MSG^*/MSE^* \sim \{(1+\rho)U/(g-1)\}/[\{(1-\rho)V+(1+\rho)W\}/(2P-g)].$$