PERSPECTIVE

People and eyes: statistical approaches in ophthalmology

I E Murdoch, S S Morris, S N Cousens

Clinical or epidemiological studies in ophthalmology generally involve the examination of eyes for a variety of clinical signs or measurements. Unlike cardiology, hepatology, or many other clinical specialties, however, individuals have two eyes. In addition, unlike kidneys or lungs, both eyes are easily accessible for assessment and are frequently both assessed. In some instances a number of participants in a study may refuse complete assessment in both eyes, or the data from one eye may be unavailable for other reasons, resulting in data sets with information on one eye for some individuals and on two eyes for others.

If information is available on both eyes of the same individual, the findings in the left eye are generally likely to be more similar to those in the right eye of the same individual than to those in another eye from a different individual. This is because a multitude of factors, including environmental and genetic factors, act at the level of the individual and thus have an impact on the probability of the finding occurring in both eyes.

Clearly, however, the degree to which a given finding in the right eye predicts the probability of the same finding in the left eye varies considerably for different conditions. Some conditions characteristically only ever occur in one eye. This may be because the disease is rare. An example of this is choroidal melanoma which occurs in only one eye in 99% of cases.[1] Other examples are corneal herpes simplex infection in the immunocompetent or severe ocular trauma (98% cases[2]).

At the other extreme are conditions such as blepharitis which almost always affect both eyes (proportion bilateral 95%[3]). The finding in the right eye almost perfectly predicts the finding in the left eye (fellow eye). The majority of ocular conditions lie between these two extremes.

Correlations between observations are a recognised statistical problem in many areas of medical research,[4] including ophthalmology.[5] If such correlations are not taken into account in statistical analyses, there may be errors in the results obtained, usually resulting in falsely precise confidence intervals and falsely small p values.

Table 1 Predominant analytical approach employed in 79 clinical science papers published in the BJO Jan–June 1995

Predominant mode of analytical approach	*Number of articles*	*(%)*
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 included in analysis		
Random selection of eye for inclusion	5	(6)
Right/left selection of eye for inclusion	7	(9)
"Clinical selection" criteria—worse eye, first eye with disease, etc	13	(16)
Overall summary of ocular findings per individual		
"Pooled" findings	13	(16)
Average taken of results in the two fellow eyes	6	(8)
Analysis at ocular level		
No correction for correlation between eyes	16	(20)
Correction for correlation between eyes	2	(3)
Paired comparison (fellow eye used as "control")	5	(6)

Current practice

What methods of statistical analysis are most commonly employed in ophthalmic research? To investigate this the 79 original clinical science articles published in the *BJO* during the first 6 months of 1995 were reviewed. The findings of this review are presented in Table 1, which shows that a wide variety of approaches are employed in the analysis of ophthalmic data. In the remainder of this paper each of these approaches is discussed with examples. To illustrate some of these approaches we consider the two fictional patients described in the box below.

> *Patient A* has an intraocular pressure (continuous variable) of 18 mm Hg in his right eye and 23 mm Hg in his left eye.
> *Patient B* is diabetic and has retinopathy (binary variable) in his right eye and cataract (binary variable) in his left eye.

ANALYSIS AT THE LEVEL OF INDIVIDUAL BECAUSE OF NATURE OF OBSERVATION

In some cases both eyes are required for a diagnosis—concomitant strabismus and visual impairment being typical examples. In these circumstances it is natural to perform analyses at the level of the individual. Similarly, if the disease or intervention only ever occurs in one eye the natural level for the analysis is the individual. An example of such a situation is provided by a study of compliance with patching for amblyopia.[6]

When an analysis based on individuals is performed correlation between observations (that is, individuals) is not a problem and such an analysis is statistically valid. In the above circumstances it is also efficient.

ONE EYE PER INDIVIDUAL INCLUDED IN ANALYSIS

Single eye analyses were used in 25 reports. In five reports the eye was chosen at random for inclusion in the study analysis and in another seven only right or only left eyes were chosen. In a paper on the Pulsair tonometer, the reproducibility of intraocular pressure readings was calculated for right eyes and then for left eyes separately and both reported.[7] In 13 other reports the choice of which eye was used in the analysis was on clinical grounds such as the operated eye, the worse/better eye, or the first eye with disease.

Data for inclusion in analysis using one eye per individual

	Right eye	*Left eye*	*Random eye*
Patient A	18 mm Hg	23 mm Hg	18 *or* 23 mm Hg
Patient B	Retinopathy	Cataract	Retinopathy *or* cataract

In ocular disease the rule is for unilateral pathology to favour neither right nor left eyes. Even in ocular trauma, where the dominant eye has been said to be more frequently injured, there is no significant laterality.[2] One possible exception is normal tension glaucoma which may favour left eyes. Studies of this condition which use data from only one eye per individual will need to consider carefully how they choose which eye to study.

For conditions that are equally likely to affect either eye, analyses based on only right eyes, or on only left eyes, or on a randomly selected eye are statistically equivalent. Analyses based on a single eye per individual are convenient in that standard statistical methods can be employed. The major disadvantage of this approach is a loss of information. In the extreme case when data have been collected on both eyes in all individuals and there is no correlation between right and left eye findings, half of all the information collected is unused in the analysis. Thus, the use of data from only one eye per individual is statistically valid but is likely to be *inefficient* (the *power* and *precision* of the analysis are less than optimal).

A second disadvantage of this approach is the potential for *bias* arising through the choice of which data to use. Often, if data are only available on a single eye in an individual because of incomplete data or for other reasons, that single eye is included in the analysis. Bias could occur if some individuals have data on both eyes, only one of which is "randomly" selected for analysis, while other individuals have data on only one eye which is automatically included (that is, *not* randomly selected). The non-random selection of eyes can introduce bias.

Similar caveats apply to the choice of the first eye with disease, worse/better eye, or operated eye.

In conclusion, the use of one eye per individual in the analysis means that between eye correlation is not a problem, but that there may be considerable "waste" of available data. In addition, incomplete data collection or selection of eyes for inclusion in a study on clinical grounds have the potential to introduce bias. This potential bias should be acknowledged and, when possible, its likely extent assessed.

OVERALL SUMMARY OF OCULAR FINDINGS PER INDIVIDUAL

In the review of papers, the data from an individual's two eyes were pooled in 13 reports. For example, this method was used in the study of therapies for blepharitis[3] and in the investigation of the effect of smoking beedies on cataract development.[8] In a further six reports the average result for the two eyes was taken. For example, the average measurement of corneal surface quality in right and left eyes was used in a study of contact lens wearers.[9] In four of the six reports the use of averaging was not stated explicitly but was left to the reader to surmise.

Data for inclusion in analysis using sum of ocular findings in individual

	Pooling results	*Averaging results*
Patient A	Raised IOP >21 mm Hg	20.5 mm Hg
Patient B	Retinopathy and cataract	NA

One disadvantage of pooling or averaging results from right and left eyes is a "loss" of information and hence a loss of efficiency, analogous to that which may occur when only a single eye from each individual is used in the analysis (see above). However, the degree of information "wasted" by the pooling or averaging of findings is likely to be less than that which occurs when only one eye is included.[5] It will depend on the degree of correlation between the right and left eyes, with the greatest loss occurring when this correlation is low.

Pooling or averaging the findings from the two eyes may be a more or less reasonable approach depending on the circumstances. It may be appropriate in studies of risk factors which operate at the individual level—for example, a study of diet and IOP. It would not be appropriate for studies of factors which operate at the ocular level—for example, in studies of the impact of a topical therapy or ocular surgery.

While pooling information may appear a useful approach for presenting a clinical picture of disease in individuals, care needs to be taken in the interpretation of the findings. If uniocular findings are given the same weight as binocular findings the results will not provide a good index of threat to vision.

The use of an overall summary of ocular findings for each individual is often an inefficient use of data and will, in some circumstances, be an inappropriate use of data. Careful interpretation of the results will be required.

ANALYSIS AT OCULAR LEVEL

Analysis was performed primarily at the level of individual eyes in 18 (23%) of the studies reviewed. Only two of these studies presented results with some form of correction for between eye correlation.

Data for inclusion in analysis using analysis at ocular level

	No allowance for correlation
Patient A	18 mm Hg
and	23 mm Hg
Patient B	Retinopathy
and	Cataract

Analyses which fail to take account of between eye correlation suppose that we have more information than we really do and leads to a falsely high degree of precision. (Confidence intervals get smaller as the number of observations increased.) Differences between two groups may appear to be more statistically significant than they really are. In a simulation exercise based on a comparison of IOPs,[5] it was found that failure to take account of between eye correlation led to an increase in the rate of "false positive" results from the appropriate 5% to 20%.

ASSESSMENT OF CORRELATION BETWEEN EYES

For binary exposures or outcomes Thompson[10] proposed assessing the between eyes correlation using the formula

$$\frac{P_{2eyes} - P_{eye}^2}{P_{eye}(1 - P_{eye})}$$

where P_{2eyes} is the proportion of individuals with the finding in both eyes and P_{eye} is the proportion of eyes with the finding. Thompson called this the intraclass correlation coefficient and it is identical to the kappa statistic.[11] If there is little or no correlation between eyes the kappa statistic will be close to 0. If the kappa statistic is greater than about 0.3 then between eye correlation is likely to be an important problem.

Other approaches to assessing correlation have been proposed by Rosner[12] and Dallal.[13]

ALLOWANCE FOR CORRELATION BETWEEN EYES

Among the methods of analysis which will cope with the problem of correlation between eyes are the following.[14]

Random effects models—the disease outcome in the eye is modelled as a function of measurable characteristics of the eye and individual. The effect of the unmeasurable characteristics of the individual which give rise to the between eye correlation is modelled implicitly. This approach to the analysis of ophthalmic data has been proposed by Rosner.[15]

Generalised estimating equations—the outcome in each eye is modelled as a function of the risk factors and the correlation between the fellow eyes is modelled separately and explicitly.[16]

If used appropriately, these two approaches give similar results in most, but not all, circumstances. They represent an efficient use of available data.

PAIRED COMPARISON (FELLOW EYE USED AS "CONTROL")

In this case the fellow eye is used as a "control" for the eye of interest. Five of the studies reviewed utilised this approach. Paired designs such as the "double crossover" design advocated by Newcombe and Duff[5] are a particularly powerful approach for comparing the effects of localised treatments on bilateral eye disease. A sign test has been proposed for comparing survivorship curves in clinical trials with correlated or paired data.[17] An example of the paired approach applied outside the context of a clinical trial is provided by a study of aqueous flare in cases of choroidal melanoma in which the fellow eye was used as the control to determine if presence of the tumour resulted in increased flare.[1]

When appropriate, paired designs offer a very powerful approach. To benefit fully from the approach the statistical analysis must take account of the pairing.

Data comparison for paired analysis

Patient A	18 mm Hg compared with 23 mm Hg
Patient B	Retinopathy compared with no retinopathy *or* Cataract compared with no cataract

Summary

In conclusion, when an observation by its nature involves two eyes, as for blindness, statistical analyses should be conducted on individuals rather than eyes and between eye correlation is not a problem. In other circumstances, if information on only one eye per individual is used in the analysis there is a potential "waste" of information leading to less precise estimates of effect and less power. In addition, bias may be introduced into a study if there is non-random selection of the eye for inclusion in the analysis. The use of an overall summary of ocular findings for an individual may result in "wastage" of information in a similar fashion to the use of only one eye per individual. On the other hand, an analysis of individual eyes with no allowance made for between eye correlation may result in falsely narrow confidence intervals around estimates of effect. Between eyes correlation may be assessed empirically using the kappa statistic or similar means. If between eye correlation is substantial, statistical techniques exist which can utilise all available data while allowing for the correlation. In some circumstances a powerful design may be to use the fellow eye as a "control".

Two conclusions may be drawn from this review of analytical approaches to the analysis of clinical data in the *BJO*. Firstly, the analytical approaches employed in many studies fail to use all the data available. In other words the analysis is less than "optimal". Secondly, in a proportion of studies, inappropriate statistical methods are used which may lead the investigator to draw inappropriate conclusions. In other words, the analysis is invalid. Ophthalmic data, by their very nature, present particular statistical challenges. We emphasise the need to involve appropriate statistical expertise in the design and analysis of ophthalmic studies.

Mr Murdoch was supported by the Wellcome Trust at the time this article was written.

I E MURDOCH

Department of Preventive Ophthalmology, Institute of Ophthalmology, London EC1V 9EL

S S MORRIS

Food Consumption and Nutrition Division, International Food Policy Research Institute, Washington DC, USA

S N COUSENS

Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT

Correspondence to: Mr Murdoch.

1 Castella AP, Bercher L, Zografos L, *et al.* Study of the blood-aqueous barrier in choroidal melanoma. *Br J Ophthalmol* 1995;**79**:354–7.
2 Dunn ES, Jaeger EA, Jeffers JB, *et al.* The epidemiology of ruptured globes. *Ann Ophthalmol* 1992:**24**:405–10.
3 Seal D, Wright P, Ficker L, *et al.* Placebo controlled trial of fusidic acid gel and oxytetracycline for recurrent blepharitis and rosacea. *Br J Ophthalmol* 1995;**79**:42–5.
4 Qaqish BF, Liang KY. Marginal models for correlated binary responses with multiple classes and multiple levels of nesting. *Biometrics* 1992;**48**:939–50.
5 Newcombe RG, Duff GR. Eyes or patients? Traps for the unwary in the statistical analysis of ophthalmological studies. *Br J Ophthalmol* 1987;**71**:645–6.
6 Fielder A, Irwin M, Auld R, *et al.* Compliance in amblyopia therapy: objective monitoring of occlusion. *Br J Ophthalmol* 1995;**79**:585–9.
7 Vernon S. Reproducibility with the Keeler Pulsair 2000 non-contact tonometer. *Br J Ophthalmol* 1995;**79**:554–7.
8 Ramakrishnan S, Sulochana KN, Selvaraj T, *et al.* Smoking of beedies and cataract: cadmium and vitamin C in the lens and blood. *Br J Ophthalmol* 1995;**79**:202–6.
9 Weale RA. Contact lenses and the cornea: age and length of wear. *Br J Ophthalmol* 1995;**79**:163–5.
10 Thompson JR. The χ^2 test for data collected on eyes. *Br J Ophthalmol* 1993; **77**:115–17.
11 Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960;**20**:37–46.
12 Rosner B. Statistical methods in ophthalmology: an adjustment for the intraclass correlation between eyes. *Biometrics* 1982;**38**:105–14.
13 Dallal GE. Paired Bernoulli trials. *Biometrics* 1988;**44**:253–7.
14 Neuhaus JM, Jewell NP. Some comments on Rosner's multiple logistic model for clustered data. *Biometrics* 1990;**46**:523–31.
15 Rosner B. Multivariate methods in ophthalmology with application to other paired-data situations. *Biometrics* 1984;**40**:1025–35.
16 Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;**42**:121–30.
17 Seigel D, Pogdor M. A sign test for significance of differences in survivorship curves from paired truncated data. *Controlled Clinical Trials* 1982;**3**:69–71.