




# Statistical Methods for Best and Worst Eye Measurements

Kaustav Banerjee <sup>1</sup>, Subhasish Pramanik <sup>2</sup>, Lakshmi Kanta Mondal <sup>3</sup>

<sup>1</sup>Decision Sciences Area, Indian Institute of Management Lucknow, Uttar Pradesh, 226013, India; <sup>2</sup>Department of Endocrinology & Metabolism, Institute of Post Graduate Medical Education & Research and SSKM Hospital, Kolkata, West Bengal, 700020, India; <sup>3</sup>Department of Ophthalmology, Regional Institute of Ophthalmology, Medical College Campus, Kolkata, West Bengal, 700073, India

Correspondence: Lakshmi Kanta Mondal, Department of Ophthalmology, Regional Institute of Ophthalmology, Medical College Campus, Kolkata, West Bengal, 700073, India, Tel +9830830216, Email Lakshmi.mondal62@gmail.com

**Introduction:** The focus is often on the best and worst eyes to detect early predictive and non-invasive biomarkers of diabetic retinopathy. Typically, such data have been dealt with in a case-control setting, which applies two-sample tests and ignores the correlation between the fellow eyes. Practitioners are mostly unaware that such measurements hide the labels of the fellow eyes, which rules out standard tools, such as paired  $t$  or signed-rank tests.

**Methods:** This report discusses the problems with such data on best and worst eye measurements, and illustrates alternative paired tests for equality of means or locations using a case-control dataset.

**Results:** This report illustrates that methods which ignore the correlation between fellow eyes result in grossly conservative tests. A battery of  $Z$ -tests which consider this correlation can resolve this issue.

**Discussion:** This finding emphasizes the importance of selecting an appropriate control group for the detection of possible markers. Further, it cites an example to show that using data from fellow eyes and adjusting for their correlation may not always be the best option, contrary to common perception.

**Keywords:** best eye, worst eye, unordered pairs

## Introduction

Clinical or experimental studies of diabetic retinopathy (DR) often collect paired data from the fellow eyes of a subject. Typically, observations are made using the following three options: (a) left and right eyes, (b) randomly selected eyes, and (c) best and worst eyes. Although the first two options are easy to follow and well-documented,<sup>1-4</sup> the third option has received much less attention<sup>5</sup> and offers severe challenges. The issue is that the best and worst eye measurements hide the labels of fellow eyes: for some subjects, the best and worst could be the left and right eye, while for the rest, it could be the other way round. How can we test for equality of means or locations using such data? To the best of our knowledge, literature on ophthalmology and diabetes research is lacking in this regard.

Suppose  $(X_i, Y_i)_{i=1}^n$  are the visual acuity scores of the left and right eyes of  $n$  subjects, respectively. The acuity scores for fellow eyes are generally not equal in the presence of certain macular lesions.<sup>5</sup> Furthermore, only one eye could be affected for some subjects, while for others, both eyes could be affected to a different extent. Therefore, the focus is on the best eye,  $U_i = \min\{X_i, Y_i\}$  and the worst eye,  $V_i = \max\{X_i, Y_i\}$ . Note that for some subjects,  $(U_i, V_i)$  could be  $(X_i, Y_i)$ , whereas for others,  $(U_j, V_j)$  could be flipped as  $(Y_j, X_j)$ , giving rise to unordered pairs.<sup>6</sup> To detect a marker, suppose we wish to determine whether the acuity scores are significantly different between the best and worst eyes. If one applies a paired  $t$  or signed-rank test with  $(X_i, Y_i)_{i=1}^n$  pairs, it would detect possible differences between the left and right eyes, *not* between the best and worst eyes. Here, we discuss alternative tests for equality of means or locations tailored to address such challenges.

Practitioners are barely aware of the issues involving unordered pairs. As a result, when it comes to assessing whether the best and worst visual (acuity) scores are equal, they typically resort to two-sample tests following a case-control setting. Armstrong<sup>7</sup> analysed 230 articles published in three optometric journals over a span of three years. Of these, 64% (148/230) collected data from one eye only. In these one-eyed studies, 23% (34/148) had better or worse/diseased eyes. Among the studies involving both eyes, 52% (43/82) considered either one eye (typically the right eye), analysed both eyes separately, or treated one eye as diseased and the other as treated (control). 29 out of 82 studies (35%) conducted on both eyes were carried out without correction for correlation, or it was not clear how the data were analysed. Despite this heavy reliance on two-sample tests in practice, here we show that such a procedure may not always identify the markers of eye diseases, such as DR, for which early detection is a key step for its prevention.

Unordered pairs are not unique to diabetes research or ophthalmology. They have biomedical applications in genetics,<sup>8–10</sup> clinical trials,<sup>11</sup> twin studies<sup>12</sup> and social studies with dyads.<sup>13</sup> This report briefly discusses the underlying issues and illustrates appropriate methodologies using a case-control dataset. Furthermore, it presents an example to show that using data from fellow eyes and adjusting for correlation may not always be the best option, contrary to common perception.

## Materials and Methods

Suppose that the left and right-eye acuity scores of  $n$  subjects,  $(X_i, Y_i)_{i=1}^n$ , are random samples from a bivariate normal distribution with means  $(\mu_x, \mu_y)$ , standard deviations  $(\sigma_x, \sigma_y)$  and correlation coefficient  $\rho$ . As  $(U_i, V_i)$  can be  $(X_i, Y_i)$  or  $(Y_i, X_i)$ , their density function is an equal mixture of bivariate normal densities:

$$\psi(u, v; \theta) = \frac{1}{2}\phi(u, v; \theta_1) + \frac{1}{2}\phi(u, v; \theta_2)$$

If the means and/or standard deviations are unequal, notice that  $\psi(u, v; \theta)$  remains unaltered for  $\theta_1 = (\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$  and  $\theta_2 = (\mu_y, \mu_x, \sigma_y, \sigma_x, \rho)$ , the two distinct parameters. For this loss of identifiability,<sup>14</sup> (a) we cannot estimate the parameters of the bivariate normal model, (b) we cannot test for equality of means against one-sided alternatives – only two-sided alternatives can be tested, and (c) the classical likelihood ratio test for means is ruled out. This calls for alternative tests for means/locations, which can be broadly classified based on their assumptions, most of which assume that the  $(X_i, Y_i)$  pairs are uncorrelated ( $\rho = 0$ ). Next, we first describe two such tests for equality of means and locations.

## When Left and Right Eye Scores are Uncorrelated

### Moore's Exact F-Test

*F*-tests compare two independent variance estimates by examining their ratios. Presently, one of these estimates, which appears in the denominator of (1), is the sample variance based on the pairwise sum  $S_i = V_i + U_i = X_i + Y_i$ . In the numerator, there is an average of the squared and unsigned paired differences  $|D_i| = V_i - U_i = |X_i - Y_i|$ . Under normality, and assuming  $\sigma_x = \sigma_y = \sigma$ , these two are independent and competing estimates of  $\sigma^2$ , provided  $\rho = 0$ . The estimate in the numerator is also unbiased if  $H_0 : \mu_x = \mu_y$  holds true. If the null hypothesis is not true, the numerator overestimates  $\sigma^2$ . This leads to the following *F* test<sup>10</sup> rejecting the null if:

$$F = \frac{n^{-1} \sum D_i^2}{(n-1)^{-1} \sum (S_i - \bar{S})^2} > F_{\alpha, n, n-1} \quad (1)$$

where  $F_{\alpha, n, n-1}$  is the upper  $100\alpha\%$  cut-off point. This test has some desirable properties: (a) it is the only exact test available to date, perfectly controlling the risk of false discovery, and (b) it can be computed using pocket calculator/Excel. See the key points below for further details.

### Key Points for Equation (1)

- The numerator is an average of the squared and unsigned paired differences.
- The denominator is the sample variance based on the pairwise sum.
- If, on an average, the best and worst eye scores are equal, so the null hypothesis holds true, the ratio will follow  $F$  distribution as indicated above; provided the eye scores are not correlated.
- The test can be carried out in Excel, comparing the observed value of the ratio, with the appropriate cut-off points from  $F$  distribution.

### Davies & Phillips' Approximate Z Test

If the normality assumption does not hold, this is the best non-parametric approach. Suppose the median of the paired differences is zero, and the null hypothesis that the best and worst scores are marginally similar is true. We expect that the best score of the  $i$ 'th pair ( $U_i, V_i$ ) will be close to the worst score of the  $j$ 'th pair ( $U_j, V_j$ ). If the null is false and the median deviates grossly from zero, the best score of the  $i$ 'th pair should be far away from the worst score of the  $j$ 'th pair. This is assessed using a modified Mann–Whitney test statistic  $W$ , where a smaller value of  $W$  is evidence against the null hypothesis. Therefore, the following Z test<sup>11</sup> rejects the null hypothesis if

$$Z = \frac{W - n(n-1)/6}{\sqrt{n(n-1)(2n+1)/45}} < -Z_\alpha, \quad (2)$$

where  $Z_\alpha$  is the upper  $100\alpha\%$  cut-off point from the standard normal distribution.  $W$  can be computed by comparing the best and worst scores using the Wilcox test function of the stats package in R, which reasonably controls the risk of false discovery. See the key points below for further details.

### Key Points for Equation (2)

- This test is applicable in case the normality assumption is suspect. This is about comparing the best score of the  $i$ 'th pair with the worst score of  $j$ 'th pair, done through the two-sample Mann–Whitney test, available in the Wilcox test function of the stats package in R.
- The pairwise best and worst scores are to be supplied in the Wilcox test function, which in turn will produce the observed value of  $W$ .
- The resulting test as indicated above is a Z-test, the observed value of  $Z$  is to be compared with the appropriate cut-off value obtained from the standard normal distribution.
- If marginally the best and worst eye scores are not equal, so the null hypothesis is false, it would be reflected by a smaller observed value of the Z test statistic, provided the eye scores are not correlated. This test has to be carried out through R.

## When Left and Right Eye Scores are Correlated

Fellow eye scores are typically positively correlated.<sup>5</sup> Even if the correlation is marginally positive,  $F$  and  $Z$  tests in (1) and (2) are too conservative, often the mean/location difference cannot be detected even if it exists. If the correlation is marginally negative, they are too liberal: often, falsely detect significant differences. Most of the available tests in the literature dealing with unordered pairs are practically useless owing to their assumptions  $\rho = 0$  and  $\sigma_x = \sigma_y$ . The following battery of Z tests<sup>15</sup> seems to be the only option which works reasonably well when these assumptions are violated.

### Test for Equality of Variance

When the left and right-eye scores vary similarly ( $\sigma_x = \sigma_y$ ), the regression line of the unsigned paired differences  $|D|$  on the pairwise sum  $S$  is constant, and the line is horizontal to x-axis. If  $\sigma_x \neq \sigma_y$ , the regression line is non-constant and not independent of  $S$ . This can be verified by the scatter plot of  $|D|$  against  $S$ . Formally, one can apply a likelihood-ratio test<sup>16</sup> in a roundabout manner, although for ease of application, we will go through the scatter plot approach here.

### Test for Equality of Means When Variances are Equal

If the scatter plot of  $|D|$  and  $S$  indicates constant regression, we infer  $\sigma_x = \sigma_y$ . This entails that  $D$  and  $S$  are independent. Then, to test for equality of means, we employ either of the following  $Z$  tests based on  $|D|$  alone. One is the maximum likelihood estimator (MLE) of the square of the effect size  $\mu_D^2/\sigma_D^2$  ( $\mu_D$  and  $\sigma_D$  are mean and standard deviation of the pairwise differences), while the other is its heuristic estimator. Both are right-tailed tests which can be computed using a calculator or Excel. The likelihood-based test rejects the null hypothesis if

$$Z = \sqrt{1.5n} \times MLE > Z_{\alpha}; MLE = \max\left\{\sqrt{3/m_4} \times m_2 - 1, 0\right\} \quad (3)$$

where  $m_r = n^{-1} \sum (|D_i|)^r$  is  $r$ -th-order sample moment. See the key points below for further details.

#### Key Points for Equation (3)

- First, one has to compute the MLE, which is a function of the second and fourth order sample moment of the unsigned paired differences, as defined above.
- A larger observed value of the resulting  $Z$ -test as defined above, rejects the null hypothesis of equality of means of the best and worst eye scores, even if the fellow eyes are correlated. The observed value has to be compared with the appropriate cut-off points from the standard normal distribution.
- The computation is straightforward, and can be carried out in Excel.

A test based on the heuristic estimator  $T$  rejects the null hypothesis if

$$Z = \frac{\sqrt{n} \times \{T - (2/\pi)\}}{\{(8/\pi) \times (1 - 3/\pi)\}^{1/2}} > Z_{\alpha}; T = m_1^2/m_2 \quad (4)$$

They have good control over the risk of false discovery and there is little to choose between them. See the key points below for further details.

#### Key Points for Equation (4)

- First, one has to compute the estimator  $T$ , which is a function of the first and second order sample moment of the unsigned paired differences, as before.
- A larger observed value of the resulting  $Z$ -test indicated above, rejects the null hypothesis of equality of means of the best and worst eye scores, even if the fellow eyes are correlated. The observed value has to be compared with the appropriate cut-off points from the standard normal distribution. In many ways, this test is equivalent to the test indicated in Equation 3. Still, for theoretical reasons, they should both be evaluated and compared.
- The computation is straightforward, and can be carried out in Excel.

### Test for Equality of Means When Variances are Not Equal

When the scatter plot of  $|D|$  and  $S$  indicate a non-constant regression, we infer  $\sigma_x \neq \sigma_y$ . Then, testing the equality of means is equivalent to testing for zero correlation between  $|D|$  and  $S$ . The following  $Z$  test rejects the null hypothesis if

$$|Z| = \frac{\sqrt{n} |r(|D|, S)|}{\hat{V}} > Z_{\alpha/2}, \text{ where } \hat{V}^2 = \frac{n \sum (|D_i| - \overline{|D|})^2 (S_i - \bar{S})^2}{\sum (|D_i| - \overline{|D|})^2 \sum (S_i - \bar{S})^2} \quad (5)$$

where  $r(|D|, S)$  is the sample correlation between  $|D|$  and  $S$ , and  $\overline{|D|}$  refers to the mean of the unsigned differences. See the key points below for further details. This test has several desirable properties. If pairs were not unordered, a paired  $t$ -test should be used as the benchmark. This test closely mimics a paired  $t$ -test. Second, it works reasonably well even when the normality assumption is grossly violated, and can be computed using a calculator or Excel.

**Table 1** Paired Tests with Best and Worst Scores for DC Group

Test	Visual Acuity		Contrast Sensitivity	
	Statistic	p-value	Statistic	p-value
F	0.065	1.00	0.549	0.94574702
Z	6.100	1.00	1.951	0.97447128
Likelihood-Z	- 0.447	0.69619	0.633	0.17362
Heuristic-Z	- 4.608	1.00	- 0.112	0.55287
Correlation-Z	0.606	0.544721	1.790	0.07338704

### Key Points for Equation (5)

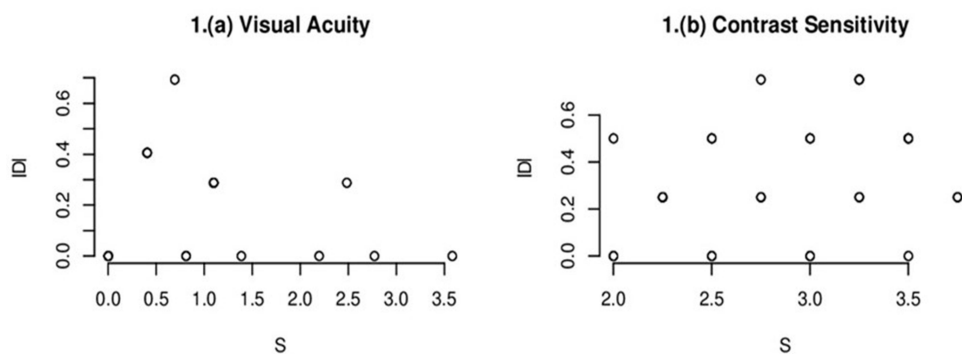
- First, one has to compute the sample (Pearson) correlation coefficient between the unsigned paired differences and paired sums.
- Second, one has to compute the denominator as defined above. This denominator is nothing but the associated standard error term of the distribution of the sample correlation coefficient, obtained in the first step.
- The resulting *Z*-test is a two-sided test: extremely larger or smaller observed value of *Z* is indicative of the fact that the means of the best and worst eye scores are not equal; even if the eye scores are correlated, even if the eye scores do not vary similarly. The observed value has to be compared with appropriate cut-off points from the standard normal distribution.
- The computation is slightly convoluted, but can be easily carried out in Excel.

## Results

We illustrate these test procedures using a case-control dataset from a study seeking to detect the early markers of DR.<sup>17</sup> There are 30 diabetic subjects without DR, seen as diabetic “control” (DC). 43 subjects were included in the DR (mild non-proliferative DR) group. Visual acuity (VA) and contrast sensitivity (CS) scores are obtained from both eyes of each subject, and suppose the objective is to see if the best and worst scores are substantially different. Table 1 summarises the test statistics and associated *p*-value for all five tests in the DC group.

From Table 1, we see that all these tests unanimously failed to find differences in the best and worst VA scores: three of them had an approximate *p*-value of 1. Should we feel so confident? Recall that the *F* and *Z* tests are appropriate if fellow eye scores are uncorrelated, an assumption which rarely (if ever) holds in this context. As previously discussed, in the presence of even a small positive correlation, *F* and *Z* tests are extremely conservative, as corroborated by their high *p*-value. Therefore, they can be disregarded safely.

Next, to see if  $\sigma_x = \sigma_y$ , we checked the scatter plot of  $|D|$  against *S* for acuity scores and looked for possible non-constant nature. Figure 1a shows that as *S* increased,  $|D|$  decreased. Therefore, the regression line is not independent of *S*:



**Figure 1** Scatter plot with best and worst scores for DC group: (a) depicts the plot for VA and (b) for CS, respectively.

**Table 2** Paired Tests with Best and Worst Scores for DR Group

Test	Visual Acuity		Contrast Sensitivity	
	Statistic	p-value	Statistic	p-value
F	0.118	1.00	0.145	1.00
Z	7.556	1.00	5.635	1.00
Likelihood-Z	-0.373	0.6685	0.589	0.21664
Heuristic-Z	-6.664	1.00	-0.579	0.72476
Correlation-Z	2.672	0.007541628	0.663	0.507402

we infer  $\sigma_x \neq \sigma_y$ , and rule out likelihood and heuristic tests. The correlation-based test did not find a significant difference; however, it was the least confident.

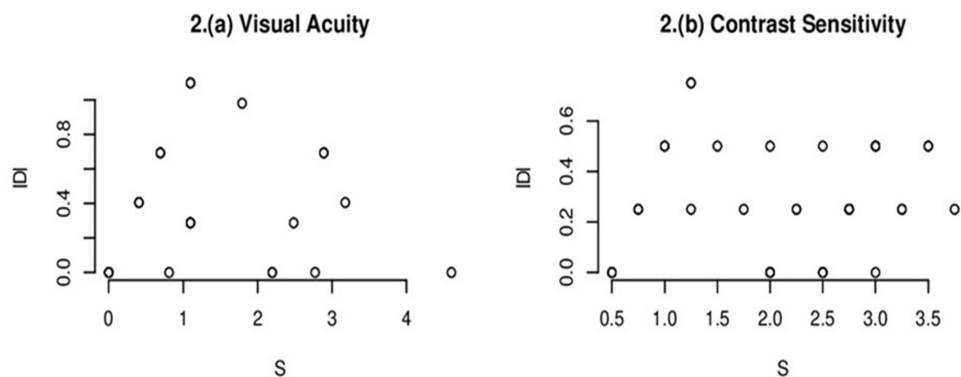
For CS, the scatter plot in Figure 1b suggests a constant regression line, indicating that either the likelihood or heuristic test should be followed. Both tests failed to identify substantial differences in means. Therefore, within the DC group, the best and worst eyes did not substantially differ in either their VA or CS scores.

### Discussion

Practitioners are generally unaware of the paired tests described here. Therefore, while dealing with the best and worst eyes, the only option seems to be to compare, for example, the worst eyes between two independent groups of subjects, as in case-control studies. Thus, the correlation between fellow eyes does not appear in this approach. Simultaneously, researchers seem to be convinced that considering both eyes of a subject and adjusting for the correlation between fellow eyes is preferable. Here, we discuss a counterexample to demonstrate that this may not be true in general.

Recall that, for the DC group, Table 1 shows no substantial difference between the best and worst eyes in their VA and CS scores. What about the DR group? The results are presented in Table 2. The *F* and *Z* tests have extremely high *p*-value, so we ignore them, as before. Comparing Figure 2a with Figure 1a, it is evident that the decrease in  $|D|$  with increasing *S* is more pronounced in the DR group than in the DC group. Therefore, we relied on the correlation-based test which provides strong evidence of substantial differences in VA scores, implying that VA is an important marker for early DR.

What about CS as a marker? In the DC group, we found no substantial differences in the CS scores of the best and worst eyes. For the DR group, we ignored the *F* and *Z*-tests again and compared the results in Figure 2b with Figure 1b. As before, Figure 2b hints at constant regression and likelihood or heuristic tests found no difference in CS scores. Can we infer that CS is not an important marker? We now consider the data in Table 3.



**Figure 2** Scatter plot with best and worst scores for DR group: (a) depicts the plot for VA and (b) for CS, respectively.

**Table 3** Two-Sample *t*-test Comparing Worst Eyes Between DR and DC Group

Marker	Statistic	p-value
Visual acuity	0.455	0.6502
Contrast sensitivity	-3.401	0.001106

In Table 3, the two-sample *t*-test compares the worst VA scores between the DR and DC groups. We also compared the worst CS scores between the two groups. From Table 3, it appears that CS is a good marker whereas VA is not. If we go by the common perception, applying paired tests as in Tables 1 and 2 is the right approach which says that VA is valuable as a marker and CS is not. However, the two-sample *t*-test says otherwise. Which pieces of evidence should we listen to?

This perception is based on the following fact. A paired *t*-test, taking care of within-pair correlations, is more capable of discovering true differences. This is related to the denominator of the paired *t*-test, the sample standard deviation of the paired differences. As the within-pair correlation  $\rho$  approaches 1, the paired differences hardly vary; therefore, the standard deviation becomes smaller, making the paired *t*-test more powerful in detecting smaller mean differences.

However, this is one side of the story. A paired *t*-test fails to discover true differences when the paired differences are more or less evenly distributed on both sides of the axes or when they are too scattered. This is why, the paired tests were unsuccessful in detecting CS as a useful marker. Looking at the *p*-values of the likelihood and heuristic tests for CS from Tables 1 and 2, and comparing Figures 1b, and 2b, it appears that, in both cases, the standard deviation dominates the average.

The two-sample *t*-test comes as a rescue, with two-fold purpose: (1) efficiently capturing the mean difference, and (2) smoothing out differences in variability between the two groups by pooling them to a weighted average. Therefore, a smaller mean difference can result in a larger test statistic. Perhaps this is why Ray and O'Day<sup>2</sup> emphasised on a two-eye design for routine analyses. Recall that the objective was to identify *early* markers as the subjects suffered from *mild* non-proliferative DR. Therefore, comparing the worst CS scores between the DR and DC groups holds the key: to detect a marker, the key lies in finding an appropriate DC group.

Therefore, VA and CS should both be regarded as useful early markers for DR. While the paired test can identify VA, a two-sample *t*-test can identify CS. Following the common perception of taking into account the correlation between fellow eyes, CS would have been missed from the list of possible markers. Therefore, while collecting the data, if the question is one or both eyes: the answer is definitely both eyes. The more the data, the better. However, if the question is whether a two-eye or paired-eye design should be used; the answer is, allow the evidence to be your guide.

## Conclusion

While detecting early predictive and non-invasive biomarkers of DR, interest often lies in the best and worst eyes which hide the labels of fellow eyes. Problems with such data and statistical methods to test for equality of means or locations are almost unknown to practitioners in general. It may be of some interest to see, for example, that Olkin and Viana<sup>5</sup> recognise that the best and worst eyes hide the labels and assumes that the fellow eyes are exchangeable, so that  $\mu_x = \mu_y$  and  $\sigma_x = \sigma_y$ . Our results indicate that this assumption may not hold in real life.

There should be valid concerns regarding the use of scatter plots to detect equality of variance, as they are not free from subjective assessments. If the scatter plot seems inconclusive, one can always go for the likelihood-ratio test,<sup>15,16</sup> which can be implemented in R. All the tests can be carried out in R, and the R-code is available upon request. We also provided a counterexample to refute the perception that considering data from fellow eyes and adjusting for their correlation is the best option. Instead, we attempt to highlight the importance of selecting an appropriate control group by exploring the data. This exploration requires active participation of a statistician because, the issues discussed here concern current statistical research, where newer results are being obtained.

## Ethics and Consent Statements

This study was approved by the Institutional Ethics Committee (Medical College, Kolkata, Ref. No: MC/KOL/IEC/NON-SPON/181/12-2018). Informed consent was obtained from all patients according to the Declaration of Helsinki.

## Acknowledgments

Professor Asim Kumar Ghosh, Director, Regional Institute of Ophthalmology, for his encouragement and support in the fulfilment of this work. We thank an anonymous reviewer for the insightful comments which improved the presentation of the paper.

## Funding

This study did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

## Disclosure

The authors declare that they have no conflicts of interest in this work.

## References

- Rosner B. Statistical methods in ophthalmology: an adjustment for the intraclass correlation between eyes. *Biometrics*. 1982;38(1):105–114. doi:10.2307/2530293
- Ray WA, O'Day DM. Statistical analysis of multi-eye data in ophthalmic research. *Invest Ophthalmol Vis Sci*. 1985;26(8):1186–1188.
- Newcombe RG, Duff GR. Eyes or patients? Traps for the unwary in the statistical analysis of ophthalmological studies. *Br J Ophthalmol*. 1987;71(9):645–646. doi:10.1136/bjo.71.9.645
- Murdoch IE, Morris SS, Cousens SN. People and eyes: statistical approaches in ophthalmology. *Br J Ophthalmol*. 1998;82(8):971–973. doi:10.1136/bjo.82.8.971
- Olkin I, Viana M. Correlation analysis of extreme observations from a multivariate normal distribution. *J Am Stat Assoc*. 1995;90(432):1373–1379. doi:10.1080/01621459.1995.10476642
- Hinkley DV. Two-sample tests with unordered pairs. *JRSS B*. 1973;35(2):337–346. doi:10.1111/j.2517-6161.1973.tb00963.x
- Armstrong RA. Statistical guidelines for the analysis of data obtained from one or both eyes. *Ophthalmic Physiol Opt*. 2013;33(1):7–14. doi:10.1111/opo.12009
- Matérn B, Simak M. Statistical problems in karyotype analysis. *Hereditas*. 1968;59(2–3):280–288. doi:10.1111/j.1601-5223.1968.tb02177.x
- Carothers AD. On determining the parental origins of homologous chromosomes. *Ann Human Gene*. 1981;45(4):367–374. doi:10.1111/j.1469-1809.1981.tb00350.x
- Moore DH. Do homologous chromosomes differ? Two statistical tests. *Cytogenet Cell Genet*. 1973;12(5):305–314. doi:10.1159/000130469
- Davies P, Phillips AJ. Nonparametric tests of population differences and estimation of the probability of misidentification with unidentified paired data. *Biometrika*. 1988;75(4):753–760. doi:10.1093/biomet/75.4.753
- Ernst MD, Guerra R, Schucany WR. Scatterplots for unordered pairs. *Am Stat*. 1996;50(3):260–265. doi:10.1080/00031305.1996.10474394
- Kenny DA, Kashy DA, Cook WL. *Dyadic Data Analysis*. New York: Guilford Press; 2020.
- Liu X, Shao Y. Asymptotics for likelihood ratio tests under loss of identifiability. *Ann Stat*. 2003;31(3):807–832. doi:10.1214/aos/1056562463
- Banerjee T, Chattopadhyay G, Banerjee K. Two-stage test of means of unordered pairs. *Stat Med*. 2017;36(15):2466–2480. doi:10.1002/sim.7304
- Crainiceanu CM, Ruppert D. Likelihood ratio tests in linear mixed models with one variance component. *JRSS B*. 2004;66(1):165–185. doi:10.1111/j.1467-9868.2004.00438.x
- Pramanik S, Chowdhury S, Ganguly U, Banerjee A, Bhattacharya B, Mondal LK. Visual contrast sensitivity could be an early marker of diabetic retinopathy. *Heliyon*. 2020;6(10):e05336. doi:10.1016/j.heliyon.2020.e05336

Clinical Ophthalmology

Dovepress

### Publish your work in this journal

Clinical Ophthalmology is an international, peer-reviewed journal covering all subspecialties within ophthalmology. Key topics include: Optometry; Visual science; Pharmacology and drug therapy in eye diseases; Basic Sciences; Primary and Secondary eye care; Patient Safety and Quality of Care Improvements. This journal is indexed on PubMed Central and CAS, and is the official journal of The Society of Clinical Ophthalmology (SCO). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/clinical-ophthalmology-journal>