Statistical Review

RECOMMENDATIONS FOR ANALYSIS OF REPEATED-MEASURES DESIGNS: TESTING AND CORRECTING FOR SPHERICITY AND USE OF MANOVA AND MIXED MODEL ANALYSIS

Richard A. Armstrong

School of Life and Health Sciences: Ophthalmic Research Group, School of Optometry, Aston University, Birmingham, B4 7ET, UK.

Running head: Repeated-measures design

Abstract

Purpose: A common experimental design in ophthalmic research is the repeated-measures design in which at least one variable is a within-subject factor. This design is vulnerable to lack of 'sphericity' which assumes that the variances of the differences among all possible pairs of within-subject means are equal. Traditionally, this design has been analysed using a repeated-measures analysis of variance (RM-ANOVA) but increasingly more complex methods such as multivariate ANOVA (MANOVA) and mixed model analysis (MMA) are being used. This article surveys current practice in the analysis of designs incorporating different factors in research articles published in three optometric journals, viz. ophthalmic and physiological optics (OPO), optometry and vision science (OVS), and clinical and experimental optometry (CXO), and provides advice to authors regarding the analysis of repeated-measures designs.

Recent findings: Of the total sample of articles, 66% used a repeated-measures design. Of those articles using a repeated-measures design, 59% and 8% analysed the data using RM-ANOVA or MANOVA respectively and 33% used MMA. The use of MMA relative to RM-ANOVA has increased significantly since 2009/10. A further search using terms to select those papers testing and correcting for sphericity ('Mauchly's test', 'Greenhouse-Geisser', 'Huynh and Feld') identified 66 articles, 62% of which were published from 2012 to the present.

Summary: If the design is balanced without missing data then MANOVA should be used rather than RM-ANOVA as it gives better protection against lack of sphericity. If the design is unbalanced or with missing data then MMA is the method of choice. However, MMA is a more complex analysis and can be difficult to set up and run, and care should be taken first, to define appropriate models to be tested and second, to ensure that sample sizes are adequate.

Key Words: Repeated-measures design, Sphericity, Mixed model analysis (MMA), Mauchly's test, Greenhouse-Geisser, Huynh and Feld

Introduction

A repeated-measures design is commonly used in many fields of research including plant science[1], ecology and evolutionary science[2], the training of athletes[3], cognitive neuroscience[4], and genetics[5] and is also frequent in ophthalmic research.[6-8] In a repeated-measures design at least one factor is a within-subject variable, i.e., different values or levels of a factor are measured on the same subject or replicate and differs from a between-subject factor in which different levels of the factor are measured on different subjects. Common examples include when the repeated-measure is an ordered longitudinal variable such as time, trial number, or drug dosage, the objective being to determine how the dependent variable may change with the repeated measure. Alternatively, the repeated-measure may include different categories such as treatment or drug type and the objective may be to determine whether differences among levels of a between-subject variable vary among levels of the repeated-measure. A repeated-measures design may contain multiple within-subject factors in addition to between-subject factors resulting in complex 'mixed model' designs.[9]

A repeated-measures design is vulnerable to a number of assumptions, most significantly to lack of 'sphericity' in which the variances of the differences among all possible pairs of within-subject means are assumed to be equal.[10,11] In more complex designs, the degree of sphericity may also differ among factors; in neuroimaging studies for example, lack of sphericity often varies among brain regions.[12] Lack of sphericity is also related to the number of levels of the within-subject factor, being minimal when only two are present. If the problem of sphericity is not taken into account, the most likely consequence is larger variance ratios (F) than expected when carrying out an analysis of variance (ANOVA) and the likelihood that an inaccurate conclusion could be drawn from the results. Lack of sphericity also has significant implications for the use of *post-hoc* tests for more detailed analysis of group mean differences.[10,11]

If sphericity cannot be assumed, then the data may not exhibit 'compound symmetry', a closely-related concept, but which specifically assumes that the pattern of co-variance or correlation among levels of the repeated-measure is constant. Hence, with four successive trials it is assumed that the correlation between Trial 1 and 2 is the same as that between Trial 1 and 4. Furthermore, it is assumed that the response of all subjects in the trial would change in the same way, i.e., if the response was an increase with trial number that the regression lines would have the same positive slope for each subject. In addition, there are two further assumptions of repeated-measures ANOVA. First, that the trials are equally spaced in time, a consequence of the compound symmetry assumption, and second, that in most statistical programs such as SPSS, there are no missing values in the data set.

Traditionally, a conventional univariate repeated-measures ANOVA (RM-ANOVA) followed by a standard *post-hoc* test such as Fisher's protected least significant difference (PLSD), Tukey-Kramer honestly significant difference (HSD) test, or the Bonferroni test[7,13,14] has been used to analyse a repeated-measures design. Nevertheless, the use of this approach has been questioned in various research fields[1-5] and more complex methods suggested including multivariate ANOVA (MANOVA) and 'mixed model analysis' (MMA).[15-17] MMA in particular is regarded as more flexible and less sensitive to various assumptions, including that of sphericity and compound symmetry, than RM-ANOVA. Hence, the aims and objectives of this article are: (1) to estimate the frequency of use of the repeated-measures design in the ophthalmic literature and the extent to which the problem of sphericity has been considered, (2) to discuss the relative merits of univariate RM-ANOVA, MANOVA, and MMA, and (3) to provide some statistical advice to authors wishing to analyse repeated-measures designs.

Frequency of the repeated-measures design

Methods

To estimate the frequency of use of the repeated-measures design in the ophthalmic literature, a sample of research articles published in the period 2003-2015, in which the experimental design incorporated one or more factors, was studied in three optometry journals: (1) ophthalmic and physiological optics (OPO), (2) optometry and vision science (OVS), and (3) clinical and experimental optometry (CXO). To obtain these articles, searches were made using various terms such as 'ANOVA', 'within-subject factor', 'repeated-measures' design, 'MANOVA', and 'MMA'. Whether the data were analysed using univariate ANOVA, MANOVA, or MMA was determined for each article. A total of 193 articles were reviewed for this study. In addition, a further survey was conducted using 'Mauchly's test', 'Greenhouse-Geisser', and 'Huynh and Feld' as search terms to estimate the frequency of articles in the three journals which tested and corrected for sphericity. A more detailed manual review of all articles was also made in 2016 to determine whether the search terms identified all of the papers of interest. This review indicated that all relevant articles were identified although the search terms frequently selected articles which were not relevant and which were subsequently eliminated from the sample.

Results and discussion

Of the total sample of articles, 128/193 (66%) used a repeated-measures design and of these 76/128 (59%) analysed the data using RM-ANOVA, 10/128 (8%) used MANOVA, and 42/128 (33%) used a MMA. Hence, the majority of articles that used a repeated-measures design, the data were analysed using a conventional RM-ANOVA rather than MANOVA or MMA. Results were essentially similar in all three journals studied. In a similar analysis of articles in athletic training journals, 24% of papers reviewed used a repeated-measures design and of these papers 96% used RM-ANOVA.[3] With regard to the proportion of articles using MMA rather than RM-ANOVA, 21% used MMA prior to 2009/10 and 42% from 2009/10 to 2016. The data also suggested some articles using RM-ANOVA did not explicitly test for sphericity using Mauchly's test[18,19] or subsequently correct for this problem using either the Greenhouse and Geisser[20] or Huynh and Feld[21] adjustments.[22,23] The Greenhouse and Geisser adjustment estimates a value of 'epsilon' (ε) in order to correct the degrees of freedom (DF) of the 'F' distribution, thus enabling a more accurate 'P' value when the sphericity assumption is violated. The Huynh and Feld adjustment is similar but is regarded as more liberal than Greenhouse and Geisser. When search terms were used to specifically identify those articles which tested and corrected for sphericity, 14/66 (21%) used Mauchly's test[24], 47/66 (71%) the Greenhouse-Geisser correction[25-27], and 5/66 (8%) the Huynh and Feld correction[28], 62% of which were published in the period 2012 to the present. A problem with these procedures, however, is that they may not protect against multiple testing using *post-hoc* tests after RM-ANOVA especially if an overall error term is used.[10,11] These considerations suggest that some caution should be used regarding the conclusions of papers using RM-ANOVA, either if there has been little consideration of sphericit or where *post-hoc* testing has been used and raises the question as to the best method of analysis of a repeated-measures design, ANOVA, MANOVA, or MMA?

Analysis of repeated-measures designs

Univariate ANOVA

Most commonly, ophthalmic researchers have used conventional 'univariate' ANOVA to analyse a repeated-measures design.[29-37] An example of a RM-ANOVA using a fictitious data set based on that of Winer[38] and analysed using STATISTICA software (general linear models option) is shown in Table 1. The design is a 2 x 3 x 3 factorial in which there are two categories of a between-subject factor (Two groups A, B), three 10-minute time intervals (T1, T2, T3), and three tasks (Task 1, Task 2, Task 3), the latter two variables being repeated-measures. An obvious application of this design in ophthalmic research would be two patient groups and three visual perception tasks measured over three time intervals. The dependent variable is the number of errors made on each task. First, a basic 'univariate' analysis illustrating the significance of all main effects and interactions is shown below the data in Table 1. Note for this analysis that the data are in 'wide' format which is usual in most statistical software when carrying out RM-ANOVA, i.e., each subject is represented by a complete row and the levels of the within-subject factor are separate columns. There are three significant effects, viz, the main effects of 'time' ($F = 63.39$, $P < 0.001$) and 'task' ($F = 89.82$, $P < 0.001$) and a significant 'group x time' interaction ($F = 5.67$, $P = 0.029$). In both groups, performance on the tasks improved, fewer errors being made in successive 10-minute periods. Fig 1 illustrates this interaction showing that the responses were similar at T1 but then diverged statistically at T2, group B performing better than group A. To explore this specific interaction effect further, a contrast analysis can be used. In this circumstance, there are probably too few time periods to make a regression or polynomial 'trend' analysis worthwhile.[10,11] However, it is likely to be the differential response at T1 and T2 which accounts for the interaction effect. This contrast can be tested separately and the analysis is shown at the bottom of Table 1. Hence, the interaction effect due to the difference between T1 and T2 and T2 and T3 is significant ($F = 17.76$, $P = 0.013$) and the two-factor interaction between 'group' and 'time' is therefore, largely due to a differential improvement from T1 to T2 in group B. Note at this stage, the effect of a possible lack of sphericity has not been taken into account. Lack of sphericity is likely in any repeated-measures design with more than two levels and the Greenhouse-Geisser adjustment can be applied, this correction being shown at the bottom of Table 1. This adjustment does not affect the significance of the main effects of 'time' and 'task' but the 'group x time' interaction is not quite significant ($P = 0.057$). Hence, in this example, lack of sphericity could have led to the erroneous acceptance of an interaction effect.

Multivariate ANOVA (MANOVA)

An extension to the method above is a multivariate approach to RM-ANOVA which does not rely to the same extent on the assumption of sphericity[39] but has been relatively little used in ophthalmic research.[40-43] At one time, programs such as SPSS analysed a repeated-measures design within the MANOVA option and it was not possible to carry out a traditional univariate RM-ANOVA. The univariate analysis incorporating the Greenhouse-Geisser adjustment is a strong indication that the sphericity assumption is unlikely to be valid for the data in Table 1 and therefore, a multivariate ANOVA (MANOVA) in which the levels of the repeated-measure are analysed as multiple dependent variables, should be carried out. There are various multivariate test criteria available including 'Wilks' Lambda', 'Hotelling-Lawley Trace', and 'Roy's Largest Root' and an analysis using STATISTICA software from the data in Table 1 using the first of these tests is shown in Table 2. In this analysis, the 'group x time' interaction is no longer quite significant suggesting violation of the sphericity assumption did lead to an erroneous acceptance of the interaction effect in this instance. However, the highly significant main effects of 'time' and 'task' remain significant even after taking sphericity into account. Results from the MANOVA are therefore similar to those of the RM-ANOVA but in many circumstances, and especially with more than two levels of the repeated measure, the outcomes will be different and the multivariate option is the preferred analysis as it gives better protection against lack of sphericity.

MMA

Given the problems associated with the repeated-measures design, there has been considerable interest in alternative methods of analysis. An obvious choice is MMA (also known as 'linear mixed models' or 'hierarchical linear models') especially if multiple factors incorporating more than a single within-subject factor, missing data, or an unbalanced design is present.[15-17] Various examples of the use of this analysis have been published in the ophthalmic literature including as an alternative to RM-ANOVA.[44-47] Other studies have used MMA in combination with conventional ANOVA.[48,49] MMA has also been used in studies which use right and left eye as a within-subject factor[50,51], in the estimation of inter-observer reliability[52], and in calculation of the inter-class correlation coefficient (ICC).[53]

There are a number of considerations before attempting a MMA of a repeated-measures design. First, the variables in the investigation are usually classified as ‘fixed-effect’ or ‘random- effect’ factors. To distinguish between them, consider the effect of removing one of the levels from a factor and whether this would effectively change the detail of the experiment. Hence, in Table 2, patient group and task are fixed effect factors since changing one patient group or task to another fundamentally changes the nature of the experiment. By contrast, if the factor is time, altering one time period may not fundamentally change the nature of the experiment and hence, subjects or repeated-measures factors such as time can be usefully modelled as random-effect factors in some circumstances. Second, although closely related to sphericity, MMA usually considers compound symmetry to be the more important assumption. This assumption is unlikely to be valid if the responses of different subjects with time or trial number are different. Hence, it is useful to plot the responses for each subject individually to determine the extent to which the trends with time are consistent among subjects. An example drawn from some of the subjects in Table 1 is shown in Fig 2 illustrating the fact that although the response of all three subjects illustrated declined with time, the response of subjects 1 and 3 were similar but subject 2 declined more rapidly suggesting that the compound symmetry assumption may not be valid in this instance.

There are a number of advantages to MMA.[54-56] First, MMA is more flexible and powerful[57,58] and there are no limits to the number of factors that can be included as long as sample sizes are large enough. Note that to fit several models to a data set may require significant numbers of DF and the question of sample size needs to be carefully considered.[59] Second, although not a specific advantage of MMA, co-variates can be included along with one or more within-subject factors and various testing methodologies are usually available.[12] Third, MMA does not require complete or balanced data, has less stringent assumptions, and exhibits increased power to detect treatment effects.[60] Hence, if some data are missing for a subject, the remaining data from that patient can still be used. As a result, MMA does not use a least-squares solution to calculate its parameters as in ANOVA but a ‘maximum likelihood’ (ML) solution which does not require complete data. Nevertheless, if a design is completely balanced then similar results are likely to be obtained using ANOVA and MMA. Fourth, in longitudinal designs, the data collection does not have to be at regularly spaced sample points or consistent regarding when each subject is measured. Fifth, MMA does not assume compound symmetry and the analysis often allows the user to select whether this assumption is valid or alternatively, to select from a set of possible co-variance structures, or to specify their own co-variance structure.

The first stage of a MMA of the data in Table 1 using SPSS software and which for simplicity models all three factors as fixed-effect factors is shown in Table 3. MMA can be more complex to set up and carry out than traditional ANOVA and potential users should work through several examples in the literature.[17,61-63] In addition, reference[63] is linked to videos which illustrate the 'mechanics' of carrying out MMA. The basic rationale of a MMA is to find the simplest model which best fits the data. To carry out the analysis requires a number of requirements and several decisions. First, most software programs including SPSS, require that data be tabulated in 'long format' in which each observation occurs in a separate row. The data for subject 1 from Table 1 are illustrated in this format at the top of Table 3. Second, there are usually two methods available to estimate the parameters of the model, viz., the restricted maximum likelihood (REML) method or the full maximum likelihood (ML method).[17,60] If both fixed and random-effect factors are to be modelled, the ML method should be used and is also necessary if the intention is to statistically compare different models as in the present example.[60] Third, most software packages provide information on the parameters that are to be estimated and should be checked carefully to ensure that the model has been specified correctly. Fourth, the 'information criteria' table includes several measures of model fit. The simplest measure is the -2log likelihood (-2LL) method, but this can be difficult to interpret because it does not take into account differences in the number of parameters. The 'Akaikes information criteria' (AIC) is regarded as more straight forward and together with the 'Schwarz's Bayesian criterion' (BIC) are commonly used to evaluate the models. The values of these models are not interpretable in themselves but only relevant when comparing models, 'lowest' value indicating best fit to a model. Fifth, in repeated-measures designs, the different time measures are likely to be correlated but the initial analysis assumed no correlations are present (compound symmetry assumption). This problem can be taken into account by allowing the repeated-measure co-variance to be 'unstructured', i.e., the repeated measure to be correlated and to have unequal variances. As a result, this model can be compared with one that assumes compound symmetry. In the present example, an unstructured model could not be fitted as it required too many parameters relative to the size of the data set. However, to illustrate the comparison of models, a ‘diagonal’ model was fitted which allows the variances to be heterogeneous but the correlations to be zero. The two models can be compared using the AIC and BIC information criteria. In this circumstance, the compound symmetry model was a better fit to the data as it has the lowest scores on all information criteria and the fewest model parameters, i.e., 16 rather than 23. Once an appropriate model has been specified, and there are several potential further models which could be tested, then the various main effects and interactions can be estimated and tested as before and individual group means tested The main effects and two-factor interactions, assuming a compound symmetry model, are shown at the bottom of Table 3. The conclusions are essentially similar to those of the RM-ANOVA in which the main effects of ‘task’ and ‘time’ are highly significant but there is a significant group x time interaction effect. This result differs slightly from that given by the less powerful RM-ANOVA after Greenhouse-Geiser adjustment and MANOVA in that the group x time interaction did not quite reach significance. However, very different conclusions could be obtained using an MMA compared with RM-ANOVA and MANOVA if the design was unbalanced or with missing values.

Concluding remarks and advice

A repeated-measures design is one of the most frequently used experimental designs in ophthalmic research. Most commonly, the repeated-measure represents time and longitudinal data frequently lack sphericity and compound symmetry, problems which can significantly affect the analysis and the conclusions drawn from a study. Recently, an increasing number of articles have tested for sphericity using Mauchly’s test and have used the Greenhouse-Geisser or Huynh and Feld corrections. A problem with this approach is that it may not protect against multiple testing using *post-hoc* tests after RM-ANOVA especially if an overall error term is used.[10,11] Hence, the following advice to authors is recommended:

1. If the design is balanced without missing data MANOVA is preferable to RM-ANOVA as it gives better protection against lack of sphericity.[39-43] There are two exceptions to this advice: (a) if only two levels of the repeated measure are present or (b) there are more than two levels but sphericity can be assumed (Mauchly’s test). In both cases a RM-ANOVA is appropriate.

2. With more complex designs with multiple factors, especially if unbalanced or with missing data, then MMA is increasingly regarded as the method of choice.[15-17] Nevertheless, MMA is a more complex analysis and can be difficult to set up and run. Care should be taken to ensure that sample sizes are adequate and clarity is needed to define appropriate models to be tested.[56] Investigators inexperienced with the method should work through the examples in the literature[17,61-63] before carrying out MMA.

Disclosure statement

Dr Armstrong has nothing to disclose.

References

1: Alves RM, Madruga MR, Tavares HR, Lobato TD & De Oliveira TF. Fixed effect models with repeated measures applied to genetics improvement of cupuasu tree. Rev Bras de Fruticul 2015; 37: 993-1000.

2: Kain MP, Bolker BM & McCoy M. A practical guide and power analysis for GLMMs: detecting among treatment variation in random effects. Peerj 2015; 3: e1226.

3: Lininger M, Spybrook J & Cheatham CC. Hierarchical linear model: Thinking outside the traditional repeated-measures analysis of variance box. J Athl Train 2015; 50: 438-441.

4: Magezi DA. Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). Front Psychol 2015; 6: article 2.

5: Lee Y, Park S, Moon S, Lee J, Elston RC, Lee W & Won S. On the analysis of a repeated measure design in genome-wide association analysis. Int J Env res Pub Health 2014; 11: 12283-12303.

6: Armstrong RA, Eperjesi F & Gilmartin B. The application of analysis of variance (ANOVA) to different experimental designs in optometry. Ophthalmic Physiol Opt 2002; 22: 1-9.

7: Armstrong RA, Davies L, Dunne MCM & Gilmartin B. Statistical guidelines for clinical studies of human vision. Ophthalmic Physiol Opt 2011; 31: 123-126.

8: Armstrong RA. Statistical guidelines for the analysis of data obtained from one or both eyes. Ophthalmic Physiol Opt 2013; 33: 7-14.

9: Snedecor GW & Cochran WG. Statistical methods. 7th Ed. Iowa State University Press, Ames Iowa, 1980.

10: Howell DC. Multiple comparisons with repeated measures. 2002a www.uvm.edu/~dhowell/StatPages, (accessed 11/11/2014).

11: Howell DC. Statistical Methods for Psychology. 5th ed., Belmont, CA, Duxbury Press, 2002b.

12: Chen G, Adleman NE, Saad ZS, Leibenluft E & Cox RW. Applications of multivariate modelling to neuroimaging group analysis: A comprehensive alternative to univariate general linear model. Neuroimage 2014; 99: 571-588.

13: Armstrong RA. When to use the Bonferroni correction. Ophthalmic Physiol Opt 2014; 34: 502-508.

14. Pinto LRM, Sanches CLG & Loguercio LL. Advantages of multivariate analysis of profiles for studes with temporal variation of treatment effects in plants. Int J Plant Sci 2013; 174: 85-96.

15: Senn S. Statistical Issues in Drug Development. Wiley, Hoboken and Chichester, 2007.

16: Fitzmaurice GM, Laird NM & Ware JH. Applied Longitudinal Analysis. Wiley, Hoboken and Chichester, 2004.

17: Twisk JWR. Applied longitudinal data analysis in epidemiology. Cambridge Medicine, 2013.

18: Johnson L, Elliot DB & Buckley G. Effects of gaze strategy on standing postural stability in older multifocal wearers. Clin Exp Optom 2009; 92: 19-26.

19: Horwood AM & Riddell PM. Differences between naïve and expert observers' vergence and accommodative responses to a range of targets. Ophthalmic Physiol Opt 2010; 30: 152-159.

20: Greenhouse H & Geisser S. On methods in the analysis of profile data. Psychometrika 1959; 24: 95-112.

21: Huynh H & Feldt LS. Conditions under which mean square ratios in repeated measurement designs have exact F-distributions. J Am Stat Assoc 1970; 65: 1582-1589.

22: Richdale K, Bertsen D, Mack CJ, Merchea MM & Barr JT. Visual acuity with spherical and toric contact lenses in low- moderate astigmatic eyes. Optom Vis Sci 2007; 84: 969-975.

23: Culham LE, Chabra A & Rubin GS. User's subjective evaluation of electronic vision enhancement systems. Ophthalmic Physiol Opt 2009; 29: 138-149.

24: Khan KA, Cufflin MP & Mallen EAH. The effect of interrupted defocus on blur adaptation. Ophthalmic Physiol Opt 2016; 36: 649-656.

25: Woods CA, Bentley SA & Fonn D. Temporal changes in contact lens comfort over a day of wear. Ophthalmic Physiol Opt 2016; 36: 643-648.

26: Read SA, Alonso-Caneiro D, Free KA, Labuc-Spoors E, Leigh JK, Quirk CJ, Yang ZY-L & Vincent SJ. Diurnal variation of anterior scleral and conjunctival thickness. Ophthalmic Physiol Opt 2016; 36: 279-289.

27: Vincent SJ, Alonso-Caneiro D & Collins MJ. Miniscleral lens wear influences corneal curvature and optics. Ophthalmic Physiol Opt 2016; 36: 100-111.

28: O'Connell RA, Anderson AJ, Hosking SL & Bui BV. Provocative intraocular pressure challenge preferentially decreases venous oxygen saturation despite no reduction in blood flow. Ophthalmic Physiol Opt 2015; 35: 114-124.

29: Jinabhai A, Radhakrishnan H, Tromans C & O'Donnell C. Visual performance and optical quality with soft lenses in keratoconus patients. Ophthalmic Physiol Opt 2012; 32: 100-116.

30: Fortuin MF, Schilperoort J, Evans BJW, Edgar DF, van Den Berg TJTP & Kiers H. randomized controlled trial of the effects of two rigid gas permeable (RGP) contact lens materials and two surface cleaners on straylight values Ophthalmic Physiol Opt 2009; 29: 497-508.

31: Lam AK, Wong Y, Cheng S. Corneal volume measures for monitoring contact lens induced corneal swelling: a pilot study. Clin Exp Optom 2011; 94: 93-97.

32: McCulloch DL, Loffler G, Colquhoun K, Bruce N, Dutton GN & Bach M. The effects of visual degradation on face discrimination. Ophthalmic Physiol Opt 2011; 31: 240-248.

33: Pseudovs K, Weisinger HS & Coster DJ. Cataract surgery and changes in quality of life measures. Clin Exp Optom 2003; 86: 34-41.

34: Chan HL, Chu HW, Ng YF et al. The detection of small relative simulated field defects using multifocal VEPs. Ophthalmic Physiol Opt 2005; 25: 224-232.

35: Berntsen DA, Mutti DO & Zadnik K. Validation of aberrometry-based relative peripheral refraction measurements Ophthalmic Physiol Opt 2008; 28: 83-90.

36: Ghergel D, Hosking SL, Cunliffe IA & Heitmar R. Transient cardiac ischaemia and abnormal variations in systemic blood pressure in unselected primary open angle glaucoma patients. Ophthalmic Physiol Opt 2010; 30: 175-181.

37: Richdale K, Bertsen D, Mack CJ, Merchea MM & Barr JT. Visual acuity with spherical and toric contact lenses in low- moderate astigmatic eyes. Optom Vis Sci 2007; 84: 969-975.

38: Winer BJ. Statistical Principles in Experimental Design. McGraw-Hill, New York, 1962.

39: Loerts H. Multivariate ANOVA and repeated measures. Loerts-2008-MANOVA-Repeated-Measures_pdf (Accessed 1/10/2016).

40: Hasebo S, Graf EW & Schor CM. Fatigue reduces tonic accommodation. Ophthalmic Physiol Opt 2001; 21: 151-160.

41: Wikman AS & Sunnala H. Aging and time-sharing in highway driving. Optom Vis Sci 2005; 82: 716-723.

42: Schmidt KL, Hilmer KS, Lawrence RA & Morrish LJ. The effect of common reductions in letter size and contrast on accommodation response in young adult myopes and emmetropes. Optom Vis Sci 2005; 82: 602-611.

43: Stanley J & Hollands A. A novel video-based paradigm to study the mechanisms underlying age- and falls risk-related differences in gaze behaviour during walking. Ophthalmic Physiol Opt 2014; 34: 459-469.

44: Rohit A, Ehrmann K, Naduvilath T, Willcox M & Stapleton F. Validating a new device for measuring tear evaporation rates. Ophthalmic Physiol Opt 2014; 34: 53-62.

45: Gifford P & Swarbrick HA. The effect of treatment zone diameter in hyperopic orthokeratology. Ophthalmic Physiol Opt 2009; 29: 584-592.

46: Horwood A & Toor S. clinical test responses to different orthoptic exercise regimes in typical young adults. Ophthalmic Physiol Opt 2014; 34: 250-262.

47: Moschos MM, Brouzas D, chatziralli IP & Ladas I. Ranibizumab in the treatment of choroidal neovascularization due to age-related macular degeneration: an optical electroretinographic study. Clin Exp Optom 2011; 94: 268-275.

48: Backhouse S, Fox S, Ibrahim B & Phillips JR. Peripheral refraction in myopia corrected with spectacles versus contact lenses. Ophthalmic Physiol Opt 2012; 32: 294-303.

49: Jaschinski W, Jainta S, Hoormann J & Walper N. Objective vs subjective measurements of dark vergence. Ophthalmic Physiol Opt 2007; 27: 85-92.

50: Koh V, Tan C, Nah G et al. Correlation of structural and electrophysiological changes in the retina of young high myopes. Ophthalmic Physiol Opt 2014; 34: 658-666.

51: Lim LS, Chua S, Tan PT et al. Eye size and shape in newborn children and their relation to axial length and refraction at 3 years. Ophthalmic Physiol Opt 2015; 35: 414-423.

52: Flores-Rodriguez P, Gili P & Martin-Rios MD. Sensitivity and specificity of time-domain and spectral-domain optical coherence tomography in differentiating optic nerve head drusen and optic disc oedema. Ophthalmic Physiol Opt 2012; 32: 213-221.

53: Lowery JP, Hayes JR, Griffith A & Darcy T. Pacific acuity test: testability, validity, and interobserver reliability. Optom Vis Sci 2014; 91: 76-85.

54: Yang S, Tai Y, Sheedy JE, Kinoshita B, Lampa M & Kern JR. Comparative effect of lens care solutions on blink rate, ocular discomfort and visual performance. Ophthalmic Physiol Opt 2012; 32: 412-420.

55: Sulley A, Young G, Lorenz KO & Hunt C. Clinical evaluation of fitting toric soft contact lenses to current non-users. Ophthalmic Physiol Opt 2013; 33: 94-103.

56: Hayat MY & Hedlin H. Modern statistical modelling approaches for analyzing repeated-measures data. Nursing Res 2012: 61: 188-194.

57: Tango T: On the repeated measures designs and sample sizes for randomized controlled trials. Biostatistics 2016; 17: 334-349.

58: Shin JH. Applications of repeated-measures analysis of variance and hierarchical linear model in nursing research. Nursing Res 2009; 58: 211-217.

59: Guo Y, Logan HL, Glueck DH & Muller KE. Selecting a sample size for studies with repeated measures. BMC Medical Research Methodology 2013; 13: Article 100.

60: Ma Y, Mazumdar M & Memtsoudis SG. Beyond repeated-measures analysis of variance advanced statistical methods for the analysis of longitudinal data in anesthesia research. Reg Anesth Pain Med 2012; 37: 99-105.

61: Maxwell SE & Delaney HD. Designing experiments and analyzing data. Laurence Erlbaum Associates. 2ed Edition, 2004).

62: Howell DC. Mixed models for missing data with repeated measures Part 1. www.uvm.edu/~dhowell/StatPages/More_Stuff/Mixed-Model-Repeated/Mixed-Model-for-Repeated-Measures1.html (accessed 1/10/2016).

63: Beaumont R. Analysing repeated measures with linear mixed models (Random effects models). 1. Getting familiar with the Linear Mixed Models (LMM) options in SPSS. D:\web_sites_mine\Hicourseweb.new\stats\statistics2\repeated_measures_1_spss_1mm_intro.docx (accessed 1/10/2016).

64: Singer JD & Willett JB. Applied Longitudinal Data Analysis: Modelling Change and Event Occurrence. Oxford University Press Inc, USA.

Legends to figures

Figure 1. Mean number of errors on three visual perception tasks over three successive 10-minute time intervals in two different subject groups (A, B) illustrating the group x time interaction.

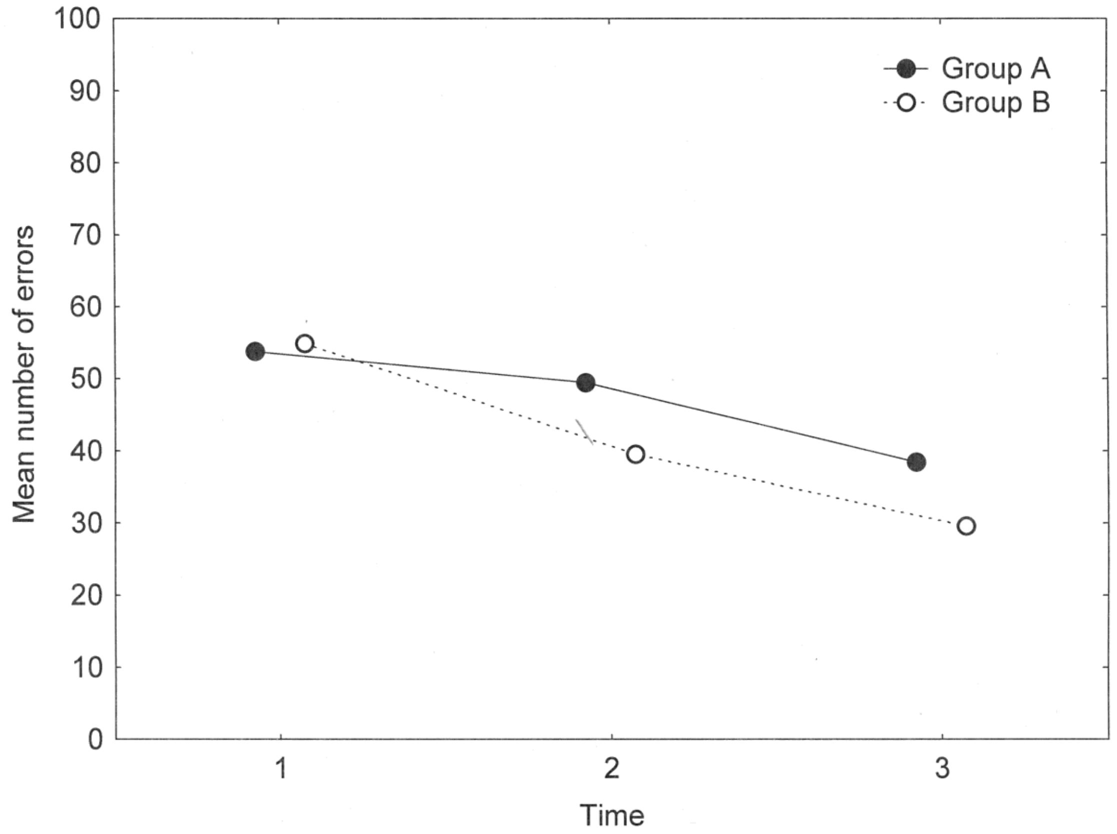

Figure 2. Individual responses of three subjects extracted from the data in Table 1 suggesting different responses over time and therefore, that the assumption of compound symmetry is unlikely to hold.