

**Title:** Multilevel multivariate modelling of the effect of gender and patient co-morbidity on spherocylindrical refractive outcome following cataract surgery

**Authors:** Rachael A Hughes<sup>1,2</sup>, Petros Aristodemou<sup>2,3</sup>, John M Sparrow<sup>2,4</sup>, Stephen Kaye<sup>5</sup>

<sup>1</sup>MRC Integrative Epidemiology Unit, University of Bristol, UK, <sup>2</sup>Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK, <sup>3</sup>Vitreous Retina Macula Consultants of Cyprus, <sup>4</sup>Bristol Eye Hospital, <sup>5</sup>Dept. Eye and Vision Science, University of Liverpool

**Running head:** Error of Expected Refractive Outcome

**Correspondence:** Stephen Kaye, Tel: 0044 1517063997, Fax: 0044 151 7065861  
s.b.kaye@liverpool.ac.uk

**Key words:** Refractive outcome, cataract, co-morbidity

**Abstract**

**Purpose** To investigate effects of co-morbidities on refractive outcomes following cataract surgery.

**Design** Case series database study

**Methods** Study population: patients on UK national ophthalmic cataract database. Variables examined included gender, age, diabetic retinopathy, glaucoma, high myopia, inherited eye disease, optic nerve disease, uveitis, pseudoexfoliation, vitreous opacities, retinal pathology, cataract type, previous surgery and posterior capsular rupture. A multivariate normal cross-classified model was fitted to the refractive outcome using Markov Chain Monte Carlo (MCMC) methods with diffuse priors to approximate maximum likelihood estimation. A MCMC chain was generated with a burn-in of 5000 iterations and a monitoring chain of 50,000 iterations.

**Main Outcome Measure** Difference between postoperative and expected refractive outcome

**Results** 490,987 cataract operations were performed on 351,864 patients by 2567 surgeons. Myopic and astigmatic errors were associated with female gender, glaucoma, previous trabeculectomy or vitrectomy, high myopia and increasing age. Hyperopic and astigmatic errors were associated with diabetic retinopathy, pseudoexfoliation and male sex. A large proportion of the unexplained variance in EERO was attributed to unobserved eye-level factors such as biometry measurements and IOL prediction formulae.

**Conclusions** There was minimal variation between surgeons. Co-morbidities and gender accounted for small but significant differences in expected outcomes.

## Introduction

Cataract surgery remains the most frequently undertaken operation with significant patient benefit (1). Surgery involves the replacement of the cloudy cataractous lens with a clear intraocular lens implant (IOL). As outcomes following cataract surgery have improved, patient expectations have correspondingly increased, including in terms of achieving good distance vision without glasses. Postoperative visual acuity (VA) and refractive error are widely used as the principle visual outcome measures to gauge success following cataract surgery. In general, the lower the postoperative refractive error the better the unaided VA. Progress has been made in minimising postoperative refractive errors and in achieving spectacle independence through improvements in surgery and refinements in the acquisition of biometric data and intraocular lens (IOL) power formulae (2,3,4,5). Unintended and uncorrected postoperative spherocylindrical refractive errors, however, are not uncommon. These have a far greater adverse effect on unaided VA or central blur than may be evident using a spherical equivalent (6,7,8,9) and uncompensated refractive errors, particularly those containing oblique cylinder axes are especially destructive on stereopsis and vision (10,11). Most studies on cataract refractive outcomes use spherical equivalent (converting spherocylindrical error into the nearest equivalent spherical error) as an outcome, as it is a scalar variable amenable to standard statistical analysis. Spherical equivalent is however an insensitive measure with potential systemic bias (12). Cylindrical errors are conventionally analysed separately due to their vectorial nature, but this leads to significant errors when analysed separately from sphere (13,14). Treating the refractive outcome as a spherocylinder (15,16) has been shown to be an appropriate and more sensitive and specific approach for identifying refractive outliers than using the spherical equivalent or the mean absolute error of nearest equivalent sphere (17) and / or cylinder (18). We reported a two to five-fold increase in patients who had clinically significantly outlying outcomes following cataract surgery and these patients would not have been identified as outliers had the analysis depended on using the spherical equivalent, and / or on treating the components of the refractive error such as the cylinder independently (18).

There are many well-recognised variables which may lead to an unintended refractive outcome following cataract surgery, such as surgically induced changes in the eye, errors in biometry and predicted effective IOL position. What is not clear, however, is whether risk factors such as co-morbidities are associated with less predictable and poorer refractive outcomes. Whilst risk factors have been identified in patients undergoing cataract surgery which predispose to surgical complications and poorer visual outcomes (19,20), this information is not available in terms of refractive outcome. Identifying and quantifying risk factors which are associated with unexpected refractive outcomes would provide surgeons with information to better inform the patient and to enable potentially better surgical planning to improve refractive outcomes.

A key feature of refractive data is its multilevel structure, with each surgeon operating on many patients, and many patients receiving an operation on each eye. Dependencies between observations (due to the multilevel structure) must be appropriately modelled in order to obtain correct standard errors [ref]. Ignoring such multilevel data structures can lead to incorrect inference [ref]. Multilevel modelling (also known as random effects, hierarchical and mixed-effects modelling) is a flexible approach that can model a wide range of multilevel structures, and can be extended to jointly model multivariate outcomes such as the spherocylindrical refractive outcome (i.e. simultaneously model its three components). This paper illustrates the application of a multilevel multivariate model to refractive data, estimated using Markov Chain Monte Carlo (MCMC) methods to identify ocular factors and co-morbidities that may have significant associations with refractive outcomes following cataract surgery.

## **Methods**

### *Study sample*

Cataract surgery data were obtained through a data sharing agreement with the Health Quality Improvement Partnership who acted as the data controller for the National Ophthalmology

Database cataract audit. De-identified data were derived from 104 UK centres undertaking National Health Service (NHS) funded cataract surgery. Database analyses of this type which use de-identified data do not require ethical permission and are viewed as audit or service evaluation (see <http://www.hra.nhs.uk/research-community/beforeyou-apply/determine-whether-your-study-is-research/>). This study was conducted in accordance with the declaration of Helsinki, and the UK's Data Protection Act. Analyses were based on data on cataract operations undertaken between 1<sup>st</sup> April 2010 and 31<sup>st</sup> August 2018. Patients were eligible for analysis if they were aged 18 years or older and had, a cataract operation using phacoemulsification (where the primary reason for the eye operation was cataract surgery for visual improvement), preoperative keratometry measurements, an intended refractive outcome and a postoperative refraction measurement. The presence of any co-morbidity was noted at surgery by ticking the relevant boxes indicating the presence of a number of concurrent diagnoses where applicable. This was an essential item of the surgery proforma and the surgical record cannot be saved unless this part was completed.

#### *Multilevel structure*

The data had a two-way cross-classified multilevel structure (21). A surgeon operated on several patients and a patient either received a single cataract operation (to the left or right eye) or two cataract operations (one on each eye) predominantly on separate occasions. Among patients who received two operations, these were sometimes conducted by different surgeons. Therefore, cataract operations for individual eyes were nested within cells of a two-way cross-classification of surgeons by patients.

#### *Analysis outcome*

The outcome of interest was the difference between the postoperative and the expected refractive outcome, defined as the Error of Expected Refractive Outcome (EERO). This term was used rather than terms such as surgically induced refractive error or surgically induced refractive change as it is

not always possible to assign the change to the surgery itself as there may be patient factors, and instrument or measurement errors. In addition, these terms have been inconsistently used in relation to one or more of the individual components of the refractive error e.g., spherical equivalent or cylinder rather than the spherocylinder as a compound number.

The data provided included the refractive target using a 3<sup>rd</sup> generation formula selected by the surgeon. The Electronic Medical Record (EMR) highlights the most appropriate IOL power formula out of Hoffer Q, Holladay 1 or SRK/T, with respect to the patient's axial length [3]. The intended or refractive target was calculated as a spherocylinder [4, 5], using preoperative keratometry measurements observed closest to the date of operation and the surgeon selected intended sphere measurement as previously described [6]. The difference between the pre-operative steep (K2) and flat (K1) meridians was added to the intended spherical refractive outcome selected by the surgeon to give the intended refractive outcome as a spherocylinder. This approximation did not take into account any correction for posterior corneal or irregular astigmatism. For those operations with multiple postoperative refraction measurements, a single postoperative measurement was selected based on when it was observed and the type of refraction measurement. Appendix table 1 shows the order of preference for selection of the single postoperative measurement. For each operation, the observed measurement that satisfied the highest criterion category was selected. The data were transformed from the sphere/cylinder  $\times$  axis scale onto the 3-components of Long's dioptric power matrix for a thin lens  $(f_{11}, f_{12}, f_{22})$  (22) before the difference between the intended and postoperative refraction was calculated.

### *Statistical analyses*

A multivariate normal cross-classified model was fitted to the EERO on the dioptric power matrix scale  $(f_{11}, f_{12}, f_{22})$ . At each level, the variances and covariances of the random effects could be distinct. The model was fitted using MCMC methods using diffuse priors to approximate maximum likelihood estimation. (23) A MCMC chain was generated with a burn-in of 5000 iterations and a

monitoring chain of 50,000 iterations. MCMC chains generated using different starting values were examined using MCMC diagnostic tools. Similarly, to estimate the mean and spread of the preoperative keratometry and postoperative refraction measurements, separate multivariate normal cross-classified models were fitted to the measurements on the dioptic power matrix scale. Analyses were conducted using Stata/MP (version 15.0; StataCorp LP, College Station, TX, USA), command `runmlwin` (24) and MLwiN software (25).

Potential variables which might affect the EERO were classified into surgeon-level, between patient level (time-independent) and within patient-level (time-dependent patient characteristics) or eye-level variables. Of these, patient's sex, age at the time of the operation, time of the refraction since the operation (within 3 months, between 3 and 6 months, between 6 and 12 months, and more than 12 months) and the refraction type (subjective, autorefraction, focimetry, focimetry with second pair of glasses, and other) were included as covariates. We also considered the following as covariates (at time of surgery): diabetic retinopathy, glaucoma, high myopia (defined as greater than -8.00D determined by the surgeon), inherited eye disease, diseased optic nerve, uveitis or synechiae, pseudoexfoliation or phacodonesis, no fundal view or vitreous opacities, other macular pathology, other retinal pathology, other ocular co-pathology, brunescant or white mature cataract, previous vitrectomy, previous trabeculectomy and the occurrence of posterior capsular rupture during surgery. Covariates were selected for inclusion using backward selection with the likelihood ratio test. As the cross-classified model could not be estimated using maximum likelihood methods we instead conducted this procedure using a nested 3-level multilevel model. Note, since the results for the fixed effects i.e., covariates' results were virtually identical between the cross-classified and nested 3-level model, we would expect covariate selection to be the same for both types of models. A covariate was eliminated from the model if the p-value from the likelihood ratio test was  $\geq 0.01$ . Patient's sex was the only patient-level covariate (i.e. a patient who received two eye operations would have the same value for sex on both occasions) that was included. All remaining covariates

were eye-level covariates (e.g., a patient's age at time of surgery would differ between the two eye operations conducted on separate occasions).

The residuals of the model were examined using diagnostic plots. To improve satisfaction of the normality assumption we excluded postoperative refraction measures where the absolute value of the sphere or cylinder was more than 10D. The final model was used to predict values of the EERO for future operations according to the co-morbidities included in the final model (e.g. whether the eye had an existing co-pathology such as glaucoma or had experienced an intra-operative complication such as posterior capsular rupture during surgery). These predicted values were generated for a single co-morbidity and for relevant combinations. The monitoring chain of 50,000 parameter estimates (of the model) were used to derive a distribution of 50,000 predictions, from which prediction intervals were derived. All analyses were conducted on the dioptric power matrix scale  $(f_{11}, f_{12}, f_{22})$  and the results back transformed to the original scale (sphere, cylinder  $\times$  axis).

## Results

Of the 1,070,601 cataract operations undertaken between 1<sup>st</sup> April 2010 and 31<sup>st</sup> August 2018 in the National Ophthalmology Database sample, 466,711 did not have a returned recorded postoperative refraction measurement. Of the remaining cataract operations, 112,174 did not have preoperative keratometry measurements, 302 did not have an intended refractive outcome recorded in the database and 427 had a post-operative sphere or cylinder measurement with an absolute value greater than 10D, leaving 490,987 operations eligible for analysis.

These operations were performed on 351,864 patients by 2567 surgeons. For 71.7% of the operations (n=352,041), this was the patient's first cataract operation. Median patient age at first eye cataract surgery was 76.6 years (interquartile range (IQR) 69.6 to 82.3 years) and 41.7% (n=146,874) of patients were male. Among the 139,123 patients who had surgery on both eyes, 177 patients had immediate sequential bilateral surgery and for the remaining 138,946 patients the median time between the operations was 3.9 months (IQR: 2.3 to 6.1 months). The same surgeon performed both operations for 31.3% (n=43,487) of these patients.

Among all 490,897 cataract operations, the mean preoperative keratometry measurement was K1 43.77 K2 43.81 @ 3 degrees (95% credible interval (CI) K1 43.76 K2 43.80 @ 4 degrees, K1 43.79 K2 43.82 @ 180 degrees), with the spread of the observations indicated by 95% prediction interval K1 39.16 K2 40.48 @ 42 degrees to K1 47.11 K2 48.42 @ 134 degrees. Similarly, the mean postoperative refraction outcome was  $-0.50+0.49\times 5$  (95% CI  $-0.51+0.50\times 5$ ,  $-0.49+0.49\times 4$ ), with the spread of the observations indicated by 95% prediction interval  $-3.00+1.54\times 38$  to  $0.99+1.45\times 147$ .

The majority of the postoperative refraction outcomes were recorded within 3 months of the operation (84.8%; n=416,429), with 7.5% (n=37,015), 4.1% (n=20,234) and 3.6% (n=17,304) recorded between 3 and 6 months, 6 and 12 months, and more than 12 months after the operation respectively. For the type of refraction outcome, 51.3% (n=251,936) were subjective, 39.2% (n=192,245) were autorefraction, 7.8% (n=38,382) were focimetry, 0.3% (n=1,601) were focimetry

with 2<sup>nd</sup> pair of glasses, and the remaining 1.4% classified as other (n=350 cycloplegic, n=14 retinoscopy and n=6,459 of unknown type).

The top five reported co-pathologies, ocular features or intraoperative complications were diabetic retinopathy, glaucoma, high myopia, brunescient/white mature cataract and other ocular co-pathology (Table 1). Reported in table 2 are the population-average effects of an ocular co-pathology, feature or intraoperative complication on the EERO (i.e. the fixed effects of the cross-classified multilevel model). Being male was associated with a hypermetropic and astigmatic shift, whilst being 10 years older (at the time of the operation) was associated with a hypermetropic and astigmatic shift. Having diabetic retinopathy was also associated with a hypermetropic and astigmatic shift, and similarly for pseudoexfoliation or phacodonesis. Conversely, previous vitrectomy or trabeculectomy surgery, high myopia, glaucoma or posterior capsular rupture were associated with a myopic and astigmatic shift. A history of uveitis, presence of synechia, other retinal pathology or the presence of a brunescient or white mature cataract had no effect on EERO. Reported in table 3 are the population-average differences in EERO between patients with two specified ocular co-morbidities compared to patients without either ocular co-morbidities (all else being equal). For example, the population-average difference in EERO of 77 year old female with glaucoma was a myopic and astigmatic shift of  $-0.35+0.41x4$  (95% CI  $-0.37+0.41x5$ ,  $-0.34+0.40x3$ ). In contrast, the population-average difference in EERO of 77 year old male with diabetes was predominantly an astigmatic shift of  $-0.019+0.37x4$  (95% CI  $-0.036+0.38x5$ ,  $-0.0011+0.37x3$ ). The mean population difference between a male with diabetes and a female with glaucoma (for all ages) would be  $+0.30+0.031x96$  [95% CI  $+0.29+0.031x86$ ,  $+0.32+0.035x106$ ].

Table 4 contains the population-average difference in EERO between patients with posterior capsular rupture plus another specified ocular co-morbidity compared to patients without either ocular co-morbidities (all else being equal). The population-average difference in EERO between patients with posterior capsular rupture and high myopia compared to similarly aged female or male

patients without either of these ocular co-morbidities was a myopic shift of  $-0.45+0.064 \times 66$  again with a narrow 95% CI of  $-0.49+0.082 \times 61$  to  $-0.41+0.049 \times 74$ .

### **Predicting the difference between the intended/expected and postoperative refractive error**

To gain an understanding of the spread of the EERO among patients, the cross-classified multilevel model was used to predict an EERO for a new operation according to the presence of an ocular co-pathology or an intra-operative complication, or the absence of any of these. Table 5 contains the corresponding 95% prediction intervals; that is, the range of likely values of an EERO for a future operation. For example, with 95% certainty, a man aged 77 years with an eye with pre-existing diabetic retinopathy will have an EERO of between  $-2.65+1.44 \times 38$ ,  $+1.41+1.36 \times 145$ . The prediction intervals were similar for all ocular co-pathologies, except for posterior capsular rupture which had a slightly wider prediction interval, that is,  $-3.18+1.47 \times 38$ ,  $+0.90+1.34 \times 146$  for a female and  $-3.02+1.45 \times 39$ ,  $+1.04+1.33 \times 145$  for a male both aged 77 years. Note the myopic shift of the prediction interval for female compared to a male. Prediction intervals were similarly generated for the presence of more than one ocular pathology when these were likely to coexist, for example, high myopia and previous vitrectomy or glaucoma and pseudoexfoliation (table 6) and following posterior capsule rupture in the presence of other ocular pathology for example, posterior capsule rupture and pseudoexfoliation (table 7). This gave a slightly wider prediction interval than the presence of a single co-pathology.

### **Residual variance at the surgeon, patient, and eye level**

In our multilevel model the residual variance (i.e. variance not explained by the covariates of the model) was partitioned into a between-surgeon component (the variance of the surgeon-level residuals), a between-patient component (the variance of the patient-level residuals) and a within-patient component (the variance of the eye-level residuals). The surgeon-level residuals represent the unobserved surgeon characteristics that affect the EERO (e.g. whether the surgeon was left- or right-handed). The patient-level residuals represent the unobserved time-independent patient characteristics that might affect the EERO. The eye-level residuals represent the unobserved eye-

level characteristics or time-dependent patient characteristics that affect the EERO, for example, changes in the thickness of the cataractous lens, or the worsening of diabetes between operations on each eye.

Appendix table 3 reports the random effects results of the cross classified multilevel model (reported on the power matrix scale). We can use these random effects to understand how much of the residual variance in the outcome is attributed to (unobserved) differences between surgeons and between patients. In our final model (right-hand side of appendix table 3) differences between surgeons and between patients explain only a small proportion of the residual variance. Differences between surgeons accounted for 4% of the residual variance in  $f_{11}$ , 23% due to differences between patients, leaving 73% of the residual variance at the eye level. Similarly, for outcome  $f_{22}$ . Note, for outcome  $f_{12}$ , the model assumed no differences between patients (see appendix for details), so 4% of the residual variance in  $f_{12}$  was due to differences between surgeons, leaving 96% of the residual variance at the eye level.

Adding covariates measured at the lowest level (in our case the eye level) to a model will always reduce the total amount of residual variance and the remaining variance at the lowest level [7]. For example, adding covariates time since operation and refraction type reduced the total residual variance in  $f_{11}$ ,  $f_{22}$ , and  $f_{12}$  from 1.23, 1.28 and 0.13, respectively, to 1.01, 1.09 and 0.12, respectively (compare the left- and right-hand results of appendix table 3). Similarly, the eye level variances were smaller for the model including these covariates (e.g., at the eye level, residual variance in  $f_{11}$  decreased from 0.92 to 0.72). The variances at the patient and surgeon levels were virtually unaffected by the addition of covariates time since operation and refraction type, implying that the (within-patient and within-surgeon) distributions of these variables were similar across patients and surgeons, respectively [7]. Although including eye-level factors refraction type and time of refraction since surgery reduced the residual variance at the eye level by 19% to 23% for outcomes  $f_{11}$  and  $f_{22}$ , respectively, it had minimal impact on narrowing of the prediction intervals (results available on request).

## Discussion

The effect of the presence of ocular co-morbidity has not previously been formally evaluated with respect to its impact on refractive outcomes, partly because of the difficulties in collecting a large enough sample to adequately power the statistical comparison. The adoption of electronic patient records by NHS trusts and the data collection for the National Ophthalmology Database has enabled the collection of 1,070,601 cataract operations, of which 490,989 cases could be analysed. This scale of data collection is unprecedented and has allowed identification of conditions which have contributed to refractive outcome following cataract surgery with a high level of precision. In this study we used a cross-classified multilevel model to take into account the multilevel structure of the data (i.e., operations nested within patients nested within surgeons and allowing a patient to receive two operations from different surgeons). Using statistical methods that ignore the correct multilevel structure (e.g., standard regression) will give standard errors that are too small (i.e., under-estimate the uncertainty), often leading to incorrect conclusions [7].

We have demonstrated how to fit a multivariate two-way cross classified model to refractive data. A multivariate approach was necessary because we were interested in the simultaneous relationships between a covariate and the three components of the spherocylindrical refractive outcome. The multilevel model appropriately accounted for dependencies within the data and enabled investigation of how each level (e.g. surgeon, patient, operation) in the data contributed to the variance of the outcome. A key strength of multilevel modelling is its ability to accommodate a wide variety of multilevel structures and extensions to multivariate settings [ref]. A downside of complex models, such as the multivariate cross classified model, is that they can be difficult to estimate using maximum likelihood methods [26]. However, with increasing computational power it is becoming easier to estimate such models using MCMC methods, which combined with diffuse priors can approximate maximum likelihood estimation. (23) Many different types of multilevel

models are implemented in several popular software packages (e.g., Stata [27], SAS [28] and R[29]), although some of the more complex models can require specialised software such as MLwiN (25).

With regards to the missing data, we conducted a complete case analysis. This is a valid approach provided the chance of being missing does not depend on the outcome given the covariates of the model (e.g. complete case analysis would be invalid if those with a poor outcome were more likely to have missing data than those with a good outcome). Although this is not something we can directly test because we do not have access to the missing outcome values, we have carefully considered the reasons for missing data in our choice of analysis.

Most of the missing refractive data was due to patients seen in services where EMR data entry occurs in the operating theatre but not in postoperative outpatient clinics and so the main reason for missing data was independent of the underlying outcome values. We acknowledge that for a small percentage of patients, the chance of being included in our complete case analysis may depend on their missing outcome data. It is unlikely, however, that our results would be significantly affected by these few patients. Note, we did not conduct multiple imputation because we did not have access to auxiliary data (i.e. additional data outside of our analysis that explain the reasons for the missing data or provide information about the missing values). In the absence of auxiliary data, multiply imputing the outcome will not reduce bias nor gain information about the missing data, and could introduce bias due to miss-specification of the imputation model, and lead to larger standard errors due to multiple imputation's inherent random process (30)

Treating the intended refractive outcome as a spherocylinder has been shown to improve the precision for detecting clinically significant refractive outliers.(18) Applying the same approach to patients with and without co-pathology or stated cataract surgical complications has identified important differences which were previously unrecognised using analyses based on the spherical equivalent or treating the components of the refractive error e.g., the sphere and cylinder as

individual independent terms. (18) The difference between the intended, expected or target refractive outcome and the actual post-operative refractive outcome, which here we have termed the 'error in the expected refractive outcome' (EERO) reflects a range of influences on outcome, not all of which will be measurable, as well as measurement error for influences which are measurable. We used the term Error of Expected Refractive Outcome (EERO) rather than prediction error, as this error may include errors of measurement whereas the term 'prediction error' assumes errors arising from the prediction model and not from measurement errors.

Although the population-average (i.e., fixed) effects were small, it is of interest that both age (per decade) and particularly sex were associated with category differences in EERO. Savini et al., reported improved refractive outcomes when taking into account gender and race but did not quantify the effect of gender (31). We found that males tended to have a more hyperopic EERO than females by  $+0.12/+0.05 \times 91$ . Females tend to have smaller eyes (axial length and corneal diameter) than males and possibly these differences influenced the prediction accuracy of the IOL formulae used. A number of co-pathologies were associated with small but statistically significant increases in the EERO such as diabetic retinopathy, pseudoexfoliation or phacodonesis, previous vitrectomy or trabeculectomy surgery, high myopia, glaucoma or patients who had posterior capsular rupture. Diabetic retinopathy, pseudoexfoliation or phacodonesis were associated with hypermetropic astigmatic shift in EERO compared to eyes with no co-pathology or complication, whilst posterior capsule rupture, glaucoma, high myopia, previous trabeculectomy and vitrectomy were associated with myopic and astigmatic shift in EERO. Specifically, posterior capsule rupture was associated with an increased myopic and astigmatic error of  $-0.36/+0.03 \times 73$ , consistent with a more anteriorly placed IOL (32). Some of these increases in EERO could be expected. For example, patients with pseudoexfoliation have been reported to have deeper post-operative anterior chamber depths (33), conversely patients with myopia have been reported to have negative prediction errors (5). The possible reasons behind the hyperopic shift in patients with diabetic retinopathy are less clear.

Although some differences in biometric parameters such as shallower anterior chamber depths and thicker lenses have been reported in patients with Type 1 Diabetes, these changes have not been observed in Type 2 diabetic patients, which form the majority of patients (34). It is interesting to speculate that the difference may be due to a biometric bias as diabetic patients with higher axial length/corneal radius ratios appear to have a lower risk of developing diabetic retinopathy (35) so the relative hyperopic shift in patients with diabetic retinopathy may be simply due to an overrepresentation of eyes with shorter axial lengths which is associated with a hyperopic error (5). The effect of gender on refractive outcome may be related to the association of longer axial length and flatter corneas in men (36).

Uveitis/synechiae, other retinal pathology and brunescens/white mature cataract however had no significant effect on the EERO. Combinations of co-pathology which might co-exist (tables 3 and 4) were associated with a correspondingly small increase in the EERO. Whilst the contribution of these co-pathologies either in isolation or in combination on the EERO are small, they do contribute to lowering the precision of IOL formulae and could be important in refining future IOL power formulae.

For the individual patient, the EERO is more apparent when 95% prediction intervals (table 5) are considered. For example, for a patient without co-pathology or known complication, the 95% prediction intervals for the EERO are  $-2.82+1.46 \times 37$  to  $+1.23+1.38 \times 146$  for a woman aged 77 years and  $-2.65+1.45 \times 38$  to  $+1.41+1.35 \times 145$  for a man aged 77 years. These intervals are wide and in addition, it would be expected that 5% (24 500 cataract operations from this dataset) would fall outside these prediction intervals. In comparison to other comorbidities such as glaucoma, diabetic retinopathy, pseudoexfoliation, high myopia, previous trabeculectomy or vitrectomy, the effect of a posterior capsular rupture during surgery can be seen in the widening of the 95% prediction intervals for the EERO to  $-3.02+1.45 \times 39$  to  $1.04+1.33 \times 145$  for a 77 year old male and  $-3.18+1.47 \times 38$

to  $+0.90+1.34 \times 10^{-6}$  for a 77 year old female. It is important to note that the predicted outcome is the deviation from the expected refractive outcome and not the deviation from emmetropia. The width of the intervals for the population means reflect the many variables which determine the refractive outcome in cataract surgery. In order to address this, we considered the random effects (unexplained variance) at 3 levels: surgeon, patient and eye. The random effects at the surgeon level were minimal which indicates that there is little variation in EERO between surgeons and that most surgeons are undertaking similar surgery, that is, approximately 4% of the unexplained variance was due to unobserved differences between surgeons. In contrast, around 23% and 73% of the unexplained variance was attributed to unobserved influences at the patient and eye level, respectively. For example, a patient's height would be a patient-level variable that is unlikely to change between operations and potentially might affect outcome. Examples of eye level factors that may affect outcome, include eye features measured during biometry (as well as those not routinely measured such as posterior corneal curvature) or conditions that change between operations.

The results of this study suggest that factors such as co-morbidity have a small but statistically significant impact on the EERO. For example, a male with diabetes is likely to have on average at least a 0.3D difference in outcome than a female with glaucoma. To our knowledge, current IOL formulae do not consider the effect of co-morbidities and sex on the predicted refractive outcome. These effects, particularly when multiple co-morbidities are present, contribute significantly to the error of the expected refractive outcome.

Factoring in the variables identified in this study which contribute to both the explained (e.g., comorbidity) and unexplained variance (e.g., biometry and IOL prediction formulae) of the expected refractive outcome, could improve the precision of formulae that predict refractive outcomes. In particular, consideration of these effects in IOL prediction models will improve precision and potentially lead to more accurate and precise refractive outcomes.

### **Acknowledgements/Disclosure:**

#### a. Funding/Support

Rachael Hughes was supported by the NIHR Programme for Applied Research grant RP-PG-0611-20013, and a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant Number 215408/Z/19/Z).

John Sparrow was supported by the NIHR Programme for Applied Research grant RP-PG-0611-20013.

#### b. Financial disclosures: No financial disclosures

#### c. Other acknowledgements: No conflicting relationship exists for any author

### **References**

1. Shekhawat NS, Stock MV, Baze EF, Daly MK, Vollman DE, Lawrence MG, Chomsky AS. Impact of First Eye versus Second Eye Cataract Surgery on Visual Function and Quality of Life. *Ophthalmology* 2017; 124(10): 1496-1503
2. Kane JX, Van Heerden A, Atik A, Petsoglou C. Accuracy of 3 new methods for intraocular lens power selection. *J Cataract Refract Surg* 2017; 43 (3): 333-339
3. Cooke DL, Cooke TL. Comparison of 9 intraocular lens power calculation formulas. *J Cataract Refract Surg* 2016; 42(8): 1157-1164.
4. Huang J, McAlinden C, Huang Y, Wen D, Savini G, Tu R and Wang Q. Meta-analysis of optical low-coherence reflectometry versus partial coherence interferometry biometry. *Scientific Reports* 2017; 7: 43414
5. Aristodemou P, Knox Cartwright NE, Sparrow JM, Johnston RL. Formula choice: Hoffer Q, Holladay 1, or SRK/T and refractive outcomes in 8108 eyes after cataract surgery with biometry by partial coherence interferometry, *J Cataract Refract Surg* 2011; 37(1): 63-71.
6. Atchison DA, Mathur A. Visual acuity with astigmatic blur. *Optom Vis Sci* 2011; 88(7): 798-805.

7. Guo H, Atchison DA. Subjective blur limits for cylinder. *Optom Vis Sci* 2010; 87(8): 549-559.
8. Wilkins MR, Allan B, Rubin G, Moorfields IOL Study Group. Spectacle use after routine cataract surgery. *Br J Ophthalmol* 2009; 93(10): 1307-1312.
9. Harris WF, Rubin A. Closed Surfaces of Constant Visual Acuity in Symmetric Dioptric Power Space. *Optom Vis Sci* 2001; 10: 744-753.
10. Chen SI, Hove M, McCloskey CL, Kaye SB. The Effect of Monocularly and Binocularly Induced Astigmatic Blur on Depth Discrimination Is Orientation Dependent. *Optom Vis Sci* 2005; 82(2): 101-113.
11. Kobashi H, Kamiya K, Shimizu K, Kawamorita T, Uozato H. Effect of axis orientation on visual performance in astigmatic eyes. *J Cataract Refract Surg* 2012; 38(8): 1352-1359.
12. Kaye SB. Approximating lens power. *Optom Vis Sci* 2009; 86(4): 382-94.
13. Harris WF. Invariance of ophthalmic properties under spherocylindrical transposition. *Optom Vis Sci* 1997; 74: 459-462.
14. Harris WF. Statistical inference on mean dioptric power: hypothesis testing and confidence regions. *J Ophthal Physiol Optics* 1990; 10: 363-372.
15. Kaye SB, Harris WF. Analysing refractive data. *J Cataract Refr Surg* 2002; 28: 2109-2116.
16. Kaye SB. Actual and intended refraction after cataract surgery. *J Cataract and Refract Surg* 2003; 29: 2189-2194.
17. Hoffer KJ, Aramberri J, Haigis W, Olsen T, Savini G, Shammas H.J, Bentow S. Protocols for studies of intraocular lens formula accuracy. *American journal of ophthalmology* 2015; 160(3): 403-405.
18. Aristodemou P, Sparrow JM, Kaye SB. Evaluating refractive outcomes after cataract surgery. *Ophthalmology* 2019; 126(1): 13-18.
19. Sparrow JM, Taylor H, Qureshi K, Smith R, Birnie K, Johnston RL. The Cataract National Dataset Electronic Multi-centre Audit of 55,567 Operations: Risk Indicators for Monocular Visual Acuity Outcomes. *Eye* 2012; 26(6): 821-826.
20. Narendran N, Jaycock P, Johnston RL, Taylor H, Adams M, Tole DM, Asaria RH, Galloway P, Sparrow JM and the UK EPR user group. The Cataract National Dataset electronic multicentre audit of 55,567 operations: risk stratification for posterior capsule rupture and vitreous loss. *Eye* 2009; 23(1): 31-37.
21. Goldstein H. *Multilevel Statistical Models*, Fourth ed., Chichester: John Wiley & Sons inc, 2010.

22. Long WF. A matrix formalism for decentration problems. *Am J Optom Physiol Optics* 1976; 53: 27-33.
23. Browne WJ. MCMC Estimation in MLwiN, V3.03. Centre for Multilevel Modelling, University of Bristol, 2019. [Online]. Available: <http://www.bristol.ac.uk/cmm/media/software/mlwin/downloads/manuals/3-03/mcmc-web.pdf>.
24. Leckie G, Charlton C. runmlwin - A Program to Run the MLwiN Multilevel Modelling Software from within Stata. *Journal of Statistical Software* 2013; 52(11): 1-40.
25. Rasbash J, Charlton C, Browne WJ, Cameron B. MLwiN Version 3.4. Centre for Multilevel Modelling, University of Bristol., 2009. [Online]. Available: <http://www.bristol.ac.uk/cmm/software/mlwin/>.
26. Harvey Goldstein (2010). *Multilevel Statistical Models*, 4th Edition. John Wiley & Sons, Chichester, West Sussex, UK.
27. StataCorp. 2019. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC.
28. SAS/STAT Software : Usage and Reference, Version 15.2. Cary, NC :SAS Institute, 2020.
29. R: A Language and Environment for Statistical Computing. R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2020.
30. Hughes RA, Heron J, Sterne JAC, Tilling K. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. *Int J Epidemiol.* 2019;48(4):1294-1304.
31. Savini G, Hoffer KJ. Clinical Results of the Hoffer H-5 Formula in 2707 Eyes: First 5th-generation Formula Based on Gender and Race. *Int Ophthalmol Clin* 2017; 57(4): 213-219.
32. Knox Cartwright NE, Aristodemou P, Sparrow JM, Johnston RL. Adjustment of intraocular lens power for sulcus implantation. *J Cataract Refract Surg* 2011; 37(4): 798-800
33. Gür Güngör S, Akman A, Asena L, Aksoy M, Sarıgül Sezenöz A. Changes in Anterior Chamber Depth after Phacoemulsification in Pseudoexfoliative Eyes and their Effect on Accuracy of Intraocular Lens Power Calculation. *Turk J Ophthalmol* 2016; 46(6): 255-258.

34. Wiemer NG, Dubbelman M, Kostense PJ, Ringens PJ, Polak BC. The influence of diabetes mellitus type 1 and 2 on the thickness, shape, and equivalent refractive index of the human crystalline lens. *Ophthalmology* 2008; 115(10): 1679-1686.
  
35. He J, Xu X, Zhu J, Zhu B, Zhang B, Lu L, He X, Bai X, Xu X, Zou H. Lens Power, Axial Length-to-Corneal Radius Ratio, and Association with Diabetic Retinopathy in the Adult Population with Type 2 Diabetes. *Ophthalmology* 2017; 124(3): 326–335.
  
36. Ferreira TB, Hoffer KJ, Ribeiro F, Ribeiro P, O'Neill JG. Ocular biometric measurements in cataract surgery candidates in Portugal. *PLoS One*. 2017;12(10):e0184837

