doi.org/10.26398/IJAS.0032-015

THE VISUAL ACUITY MEASUREMENT AND ITS MULTIPLE TESTING FOR PATIENTS IMPLANTED WITH INTRAOCULAR LENSES DURING THE CATARACT SURGERY

Norihiro Mita[1], Hiroshi Sasaki

Department of Ophthalmology, Kanazawa Medical University, Ishikawa-Uchinada, Japan

Tomofusa Yamauchi

Department of Ophthalmology, Tsukazaki Hospital, Himeji, Japan

Kazutaka Kani

Department of Orthoptics and Visual Science, Kyushu University of Health and Welfare, Nobeoka, Japan

Akio Tabuchi

Department of Orthoptics, Kawasaki University of Medical Welfare, Kurashiki, Japan

Heihachiro Hara[2]

Department of Health Informatics, Kawasaki University of Medical Welfare, Kurashiki, Japan

Abstract*: In the present article we propose a statistical approach on the visual acuity measurement and its multiple testing for patients implanted with intraocular lenses (IOLs) during the cataract surgery. Main themes of our research are (i) obtaining psychometric functions, psychophysical thresholds and their variances of individual patients, (ii) testing the difference of visual acuities between monofocal IOL group and multifocal IOL group by adopting the Cochran-Mantel-Haenszel (CMH) test statistic (based on psychometric functions), (iii) multiple comparisons of visual acuities for individual patients by adopting the delta test statistic (based on psychophysical thresholds and their variances) and the false discovery rate (FDR-BH: Benjamini and Hochberg (1995)). Our 5-meter visual acuity test based on psychometric functions shows that the difference between the monofocal IOL group (Tecnis® 1-piece ZCB00) and the multifocal IOL group (Tecnis® Multifocal +4.0 ZMB00) is not significant under the 5% significance level (p-value = 0.2801 by CMH test statistic). We also see that FDR-BH procedure produces high average statistical power in the case of large number of null hypotheses (FDR-BH: 0.604, Holm: 0.392, Bonferroni: 0.370 for the number of null hypotheses = 210). It is very effective on the problem of multiple comparisons of individual visual acuities.*

[1] Corresponding author[1]: Norihiro Mita, email: n-mita@kanazawa-med.ac.jp
[2] Corresponding author[2]: Heihachiro Hara, email: hhrhara@gmail.com

Keywords: *Visual acuity measurement, intraocular lenses (IOLs), monofocal IOL, multifocal IOL, false discovery rate (FDR).*

1. INTRODUCTION

In cataract surgery, the intraocular lenses (IOLs) are implanted inside the eyes to replace the eyes' natural lenses. Monofocal IOL is designed to provide clear vision at a single focal point. Multifocal IOL, however, can help seeing things at different distances. It is important to obtain visual acuities of individual patients for deciding ophthalmological treatments after surgery. Yamauchi et al. (2013), Gundersen and Potvin (2013), Cardona et al. (2018) studied the comparative visual performance with monofocal and multifocal IOLs. In Yamauchi et al. (2013) the visual acuity was measured by using the decimal visual acuity chart (not based on psychometric functions) and the measured decimal values were converted to the logarithm of the minimum angle of resolution (LogMAR) scale.

Main themes of our research are (i) obtaining psychometric functions, psychophysical thresholds and their variances of individual patients, (ii) testing the difference of visual acuities between monofocal IOL group and multifocal IOL group by adopting the Cochran-Mantel-Haenszel (CMH) test statistic (based on psychometric functions), (iii) multiple comparisons of visual acuities for individual patients by adopting the delta test statistic (based on psychophysical thresholds and their variances) and the false discovery rate (FDR-BH: Benjamini and Hochberg (1995)).

The automated measurements of the visual acuity were proposed by Bach (1996); Beck et al. (2003); Schulze-Bonsel et al. (2006). However, their methods cannot calculate the variance of the visual acuity in one measurement. The testing of thresholds estimated by constant stimuli method was developed by Nagai et al. (2006), and the measurement of visual acuities and variances for individual patients by using psychometric functions was proposed by Mita et al. (2010). It has been reported that the psychometric function method has higher repeatability than the conventional methods (Tokutake et al., 2014). Therefore the new method can catch very few changes of visual acuities and variances which cannot been found by the conventional measurements.

For multiple comparison problems, several statistical approaches were proposed. Ryan, Einot-Gabriel and Welsch proposed some stepwise procedures for multiple comparison problems (Einot and Gabriel (1975); Ryan (1960); Welsch (1977)). Félix and Menezes (2018) reported comparisons of ten corrections meth-

ods for t-test in multiple comparisons via Monte Carlo study. By adopting the bootstrap logistic regression, Mita et al. (2014) developed an algorithm of multiple comparisons with a control. Mita et al. (2017, 2020) proposed an algorithm of closed multiple test based on the step-down procedure with Ryan-Einot-Gabriel-Welsch (REGW) significance levels. The closed multiple testing allows the rejection of any elementary null hypothesis (H_{ij}^0) if all possible intersections of null hypotheses involving H_{ij}^0 can be rejected using valid local level test. It controls the familywise error rate (FWER) for all the hypotheses in the strong sense. The number of elements in the hierarchical family of null hypotheses for closed multiple testing is given by $M = \sum_{k=2}^{N} {}_N C_k$, where N is the number of all patients, ${}_N C_k$ is the notation of combination, and thus M becomes very large number as N increases.

Although the procedures of controlling FWER have been commonly used in multiple testing, Benjamini and Hochberg (1995) proposed a new point of view on the problem of multiplicity: the procedures of controlling the false discovery rate (FDR-BH) which is the expected proportion of errors among the rejected null hypotheses. In many applications the control of the FDR is the desirable control against errors originating from multiplicity. The power of the FDR controling method is uniformly larger than that of the other methods, and the loss of powers, as the number of null hypotheses increases, is relatively small for the FDR controlling method (Benjamini and Hochberg (1995)). The null hypotheses in FDR testings are H_{ij}^0 $(1 \leq i < j \leq N)$. Therefore the number of FDR testings is given by $M = {}_N C_2$, and it is much smaller than that of closed multiple testing. Further Benjamini et al. (2006) studied two-stage procedure (FDR-TST). We applied the basic FDR-BH procedure to our visual acuity problem of multiple comparisons for many individual patients implanted with IOLs.

In the present article, (i) psychometric functions, psychophysical thresholds and their variances are obtained by adopting a guassing rate (Section 2), (ii) two different types of test statistics are computed based on psychometric functions and psychophysical thresholds (Section 3), (iii) multiple testings are performed by adopting FDR-BH procedure and also by adopting FWER procedures (Holm, Bonferroni) for reference (Sections 4, 5).

2. MODELLING OF THE VISUAL ACUITY FOR INDIVIDUAL PATIENT

We consider the logistic regression with a constant guessing rate (Mita et al. (2010, 2017); Tokutake et al. (2014)).

Let $\pi(x_t; \alpha, \beta)$ be the probability that a patient answers the visual acuity test

t $(t = 1, \cdots, T)$ correctly:

$$\pi(x_t; \alpha, \beta) = (1 - c_0) \cdot \frac{\exp(\alpha + \beta x_t)}{1 + \exp(\alpha + \beta x_t)} + c_0 \quad (t = 1, \cdots, T), \tag{1}$$

where x_t is the visual target of test t (in our measurement, x_t is the logarithmic visual acuity: LogVA), α and β are intercept and slope parameters, respectively, and c_0 $(0 \leq c_0 < 1)$ is a guassing rate constant which depends on the method of measurement.

We obtain optimum parameters $\hat{\alpha}$ and $\hat{\beta}$ for α and β, respectively, by adopting the classical maximum likelihood estimates based on Fisher score method. Then we can obtain the psychometric function $\varphi(x)$ which is a mathematical function from a stimulus level x to a response level π such that:

$$\varphi(x) = \pi(x; \hat{\alpha}, \hat{\beta}) \quad (-\infty < x < +\infty). \tag{2}$$

The psychophysical threshold ξ is the stimulus intensity that is detected exactly at the central value of $\varphi(x)$ $(-\infty < x < +\infty)$. We see that

$$\lim_{x \to -\infty} \varphi(x) = 1 \quad \text{and} \quad \lim_{x \to +\infty} \varphi(x) = c_0.$$

Thus we define the psychophysical threshold ξ of the patient with guessing rate c_0 such that

$$\xi = \varphi^{-1}\left(\frac{1 + c_0}{2}\right). \tag{3}$$

The threshold ξ and its variance $\mathrm{var}(\xi)$ are obtained by

$$\xi = -\frac{\hat{\alpha}}{\hat{\beta}}, \quad \mathrm{var}(\xi) = \xi^2 \left(\frac{\mathrm{var}(\hat{\alpha})}{\hat{\alpha}^2} - 2\frac{\mathrm{cov}(\hat{\alpha}, \hat{\beta})}{\hat{\alpha}\hat{\beta}} + \frac{\mathrm{var}(\hat{\beta})}{\hat{\beta}^2} \right), \tag{4}$$

where $\mathrm{var}(\hat{\alpha})$ and $\mathrm{var}(\hat{\beta})$ are variances of $\hat{\alpha}$ and $\hat{\beta}$, respectively, and $\mathrm{cov}(\hat{\alpha}, \hat{\beta})$ is the covariance of $\hat{\alpha}$ and $\hat{\beta}$.

3. TEST STATISTICS

3.1. TEST STATISTIC BASED ON PSYCHOMETRIC FUNCTIONS

For the test of patient groups, we can use the Cochran-Mantel-Haenszel (CMH) test statistic based on psychometric functions (Landis et al. (1978); Penfield (2001); Somes (1986)).

Let r $(r = 1, \cdots, R)$ be strata of observations and let $u^r_{g\ell}$ $(r = 1, \cdots, R; g = 1, \cdots, G; \ell = 1, \cdots, L)$ be the number of observations at r-th stratum of the $G \times L$ contingency table. For applying the CMH test statistic to the visual acuity results, we choose indices r, g, ℓ in the contingency table such that
(i) strata $r\,(r = 1, \cdots, R)$: visual acuity strata $(x^1, \cdots, x^R)$,
(ii) rows $g\,(g = 1, \cdots, G)$: patient groups $(\Gamma_1, \cdots, \Gamma_G)$,
(iii) columns $\ell\,(\ell = 1, \cdots, L)$: responses of visual tests ($\ell = 1$ for "correct" and $\ell = 2$ for "incorrect").

Then we fix the number of columns at $L = 2$ and choose the number of observations $u^r_{g\ell}$ $(r = 1, \cdots, R; g = 1, \cdots, G; \ell = 1, 2)$ such that

$$u^r_{g1} = \sum_{i \in \Gamma_g} \varphi_i(x^r), \quad u^r_{g2} = N^r_g - u^r_{g1} \quad (r = 1, \cdots, R; g = 1, \cdots, G), \tag{5}$$

where N^r_g is the number of patients in group Γ_g, and $\varphi_i(x^r)$ is the value of psychometric function of patient i at the visual acuity stratum x^r.

The null hypothesis is described such that there is no association between the row (patient group) and the column (response of visual test).

The Cochran-Mantel-Haenszel (CMH) test statistic Ψ for the set of G groups is given by Landis et al. (1978); Mita et al. (2020),

$$\Psi = \boldsymbol{u}^T \boldsymbol{W}^{-1} \boldsymbol{u}, \tag{6}$$

where $\boldsymbol{u}$ is the vector of the number of observations and $\boldsymbol{W}$ is the variance-covariance matrix which can be determined by using the number of observations themselves (without using variances). This CMH test statistic Ψ has an asymptotic χ^2-distribution with $(G-1)$ degrees of freedom under the null hypothesis (Agresti, 2002).

3.2. TEST STATISTIC BASED ON PSYCHOPHYSICAL THRESHOLDS

We apply the delta test statistic to the problem of comparisons of visual acuities of individual patients (Mita et al., 2017; Nagai et al., 2006).

We shall test the null hypothesis for the thresholds of n patients:

$$H^0_{1\ldots n}: \quad \xi_1 = \cdots = \xi_n \quad (2 \le n \le N), \tag{7}$$

where N is the number of all patients, n is the number of patients which are considered in the null hypotheses. If we define the following notations:

$$u_j = \xi_1 - \xi_{j+1} \quad (j = 1, \cdots, n-1), \tag{8}$$

then the null hypothesis can be rewritten such that $u_1 = \cdots = u_{n-1} = 0$.

Since the vector $\boldsymbol{u} = (u_j)^T$ has the normal distribution in multiple variables in our visual acuity testing, the delta test statistic Δ for the set of n patients defined by

$$\Delta = \boldsymbol{u}^T \boldsymbol{W}^{-1} \boldsymbol{u} \tag{9}$$

has an asymptotic χ^2-distribution with $(n-1)$ degrees of freedom under the null hypothesis, where $\boldsymbol{W}$ is the variance-covariance matrix which is determined by using the variances of thresholds var(ξ).

In the cases of FDR-BH, Holm and Bonferroni procedures (refer to Sections 5.2, 5.3), we test the following null hypotheses for all pairs of patients:

$$H_{ij}^0 : \quad \xi_i = \xi_j \quad (1 \leq i < j \leq N). \tag{10}$$

In these cases, the number of patients in considered is $n = 2$, and the delta test statistic Δ_{ij} is reduced to the following simple form:

$$\Delta_{ij} = \frac{(\xi_i - \xi_j)^2}{\mathrm{var}(\xi_i) + \mathrm{var}(\xi_j)}. \tag{11}$$

4. FALSE DISCOVERY RATE

Let Q be the proportion of the rejected null hypotheses which are erroneously rejected:

$$Q = \frac{V}{V+S}, \tag{12}$$

where V is the number of true null hypotheses which are declared significant, and S is the number of non-true null hypotheses which are declared significant. Benjamini and Hochberg (1995) defined the false discovery rate (FDR) $E(Q)$ to be the expectation of Q.

The false discovery rate controlling procedure (FDR-BH) at q^* (q^* is usually chosen as $q^* = 0.05$) is stated as follows.

Consider testing null hypotheses H_k^0 $(k = 1, \cdots, M)$ based on the corresponding p-values p_k $(k = 1, \cdots, M)$, where p-values are assumed ordered such that $p_1 \leq \cdots \leq p_M$. Define the following multiple testing procedure: let K be the largest k satisfying

$$p_k \leq \frac{k}{M} q^* \equiv q_k ; \tag{13}$$

then reject all H_k^0 $(k = 1, \cdots, K)$.

5. APPLICATION TO THE VISUAL ACUITY TESTING

5.1. 2-GROUP TEST BETWEEN MONOFOCAL IOL AND MULTIFOCAL IOL

Since we adopt the Landolt-C of four different orientations in our measurement, the guessing rate c_0 is chosen as $c_0 = 0.25$, where the Landolt-C consists of a ring that has a gap. The gap can be at various positions; usually left, right, bottom and top. The explanatory variable x in our measurement is the logarithmic visual acuity (LogVA). We shall test the difference of 5-meter visual acuity between 2 groups:

Γ_1: patients implanted with monofocal IOL (Tecnis® 1-piece ZCB00),

Γ_2: patients implanted with multifocal IOL (Tecnis® Multifocal +4.0 ZMB00),

where Tecnis® is a trade mark of Precisionlens Company, USA. Data are taken from patients underwent cataract surgery at the Cataract Unit of the Department of Ophthalmology, Tsukazaki Hospital, Himeji, Japan.

The number of Landolt-C targets T_g $(g = 1,2)$ and the number of patients N_g in groups Γ_g are chosen such that

$$T_1 = 7,\ T_2 = 7\,;\ \ N_1 = 21,\ N_2 = 12.$$

For adopting the constant stimuli method, we choose the number of trials v_{it} for patient i to the test t satisfying such that

$$v_{it} = 0 \text{ or } 20\ \ (i = 1,\cdots,N_g; t = 1,\cdots,T_g),\quad \sum_{t=1}^{T_g} v_{it} = 120 \qquad (i = 1\cdots,N_g).$$

For each patient i $(i = 1,\cdots,N_g)$ in group Γ_g $(g = 1,2)$, we obtain optimum values of parameters $\hat{\alpha}_i$, $\hat{\beta}_i$, and psychophysical thresholds ξ_i and their variances $\mathrm{var}(\xi_i)$ by adopting the Fisher score method. The psychometric functions $\varphi_i(x)$ $(i = 1,\cdots,N_g)$ for individual patients in groups Γ_g $(g = 1,2)$ are shown in Figures 1 (Γ_1), 2 (Γ_2).

The mean values of logistic regression results for each group Γ_g $(g = 1,2)$ are shown in Table 1 (intercepts $\overline{\alpha}_g$, slopes $\overline{\beta}_g$ and psychophysical thresholds $\overline{\xi}_g$).

The 2-group test of difference between Γ_1 and Γ_2 is performed by adopting Cochran-Mantel-Haenszel (CMH) test statistic. The contingency table for computation of CMH test statistic is constructed by choosing strata $x^r = -0.2 + 0.05(r-1)$; $(r = 1,\cdots,9)$. The contingency table is shown in Table 2, where data are presented only for strata $r = 1,5,9$. The results of 2-group test are shown in Table 3.

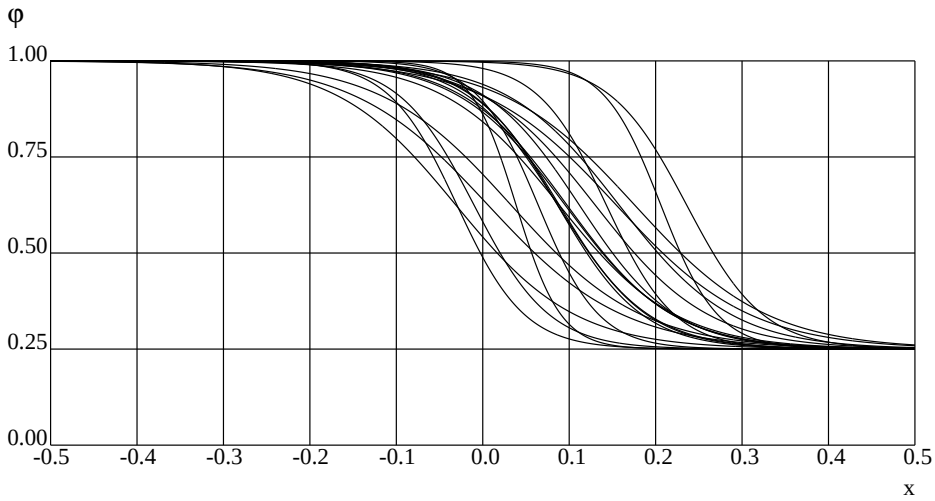

Fig. 1: Psychometric functions $\varphi_i(x)$ (i = 1,…,…,21) in Γ_1 (monofocal IOL).

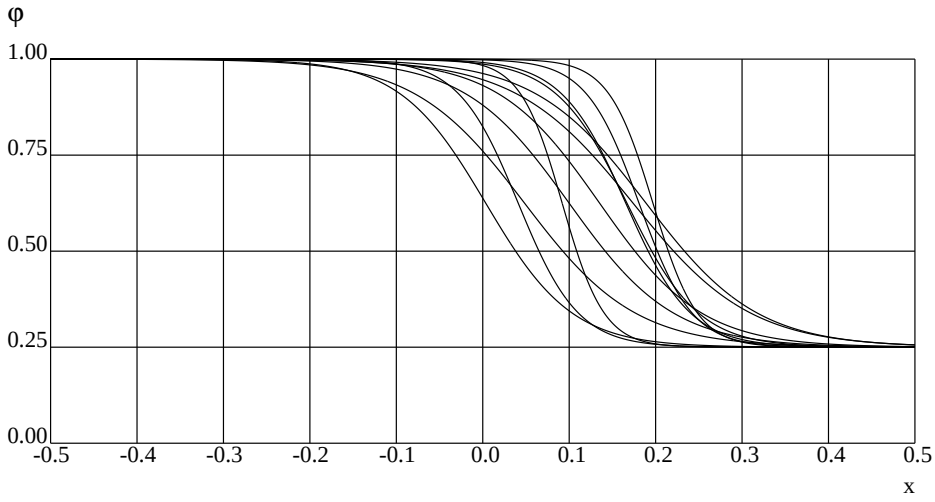

Fig. 2: Psychometric functions $\varphi_i(x)$ (i = 1,…,…,12) in Γ_2 (multifocal IOL).

Tab. 1: Mean values of logistic regression results

groups	patients N_g	intercepts $\overline{\alpha}_g$	slopes $\overline{\beta}_g$	thresholds $\overline{\xi}_g$	variances $\overline{\text{var}}(\xi_g)$	SD $\overline{\text{sd}}(\xi_g)$
$g = 1(\Gamma_1)$	21	1.80192	-19.7598	0.091760	0.000391788	0.0197936
$g = 2(\Gamma_2)$	12	3.10836	-24.3544	0.124381	0.000285412	0.0168941

Tab. 2: The 2×2 contingency table based on psychometric functions

strata	groups	$\ell=1$ (correct) u^r_{g1}	$\ell=2$ (incorrect) u^r_{g2}	total
$x^1=-0.2$	$g=1\,(\Gamma_1)$	20.7833	0.2167	$n^1_1=21.000$
	$g=2\,(\Gamma_2)$	11.9592	0.0408	$n^1_2=12.000$
	total	$m^1_1=32.7426$	$m^1_2=0.2574$	$t^1=33.000$
		.....		
$x^5=0.0$	$g=1\,(\Gamma_1)$	17.5265	3.4735	$n^5_1=21.000$
	$g=2\,(\Gamma_2)$	10.9021	1.0979	$n^5_2=12.000$
	total	$m^5_1=28.4286$	$m^5_2=4.5714$	$t^5=33.000$
		.....		
$x^9=0.2$	$g=1\,(\Gamma_1)$	8.2394	12.7606	$n^9_1=21.000$
	$g=2\,(\Gamma_2)$	5.1059	6.8941	$n^9_2=12.000$
	total	$m^9_1=13.3453$	$m^9_2=19.6547$	$t^9=33.000$

Tab. 3: Two-group test (Γ_1: monofocal IOL and Γ_2: multifocal IOL)

	monofocal N_1	multifocal N_2	strata R	dof	CMH statistic Ψ	p-value
H^0_{12}	21	12	9	1	1.1665	0.2801

5.2. MULTIPLE TEST OF INDIVIDUAL PATIENTS BY FDR-BH PROCEDURE

Hereafter to save the pages of the article, we show only the results of Γ_1 (monofocal IOL): number of patients $N=21$, number of null hypotheses $M={}_NC_2=210$.

Let ξ_i $(i=1,\cdots,N)$ be psychophysical thresholds of patients i in the group Γ_1. We assume that

$$\xi_1\le\xi_2\le\cdots\le\xi_N.$$

We shall test the following null hypotheses for all pair of patients:

null hypotheses for the pair of patients (i,j) $H^0_{ij}:\ \xi_i=\xi_j\quad(1\le i<j\le N)$.

Let p_{ij} $(1\le i<j\le N)$ be p-values obtained by delta test statistic for the null hypotheses H^0_{ij}. The p-values p_{ij} are shown in Figure 3, where p-values are presented by adopting the following symbols:

$$[*]\le 10^{-5}<[5]\le 10^{-4}<[4]\le 10^{-3}<[3]\le 10^{-2}<[2]\le 0.03<[1]\le 0.05<[0].$$

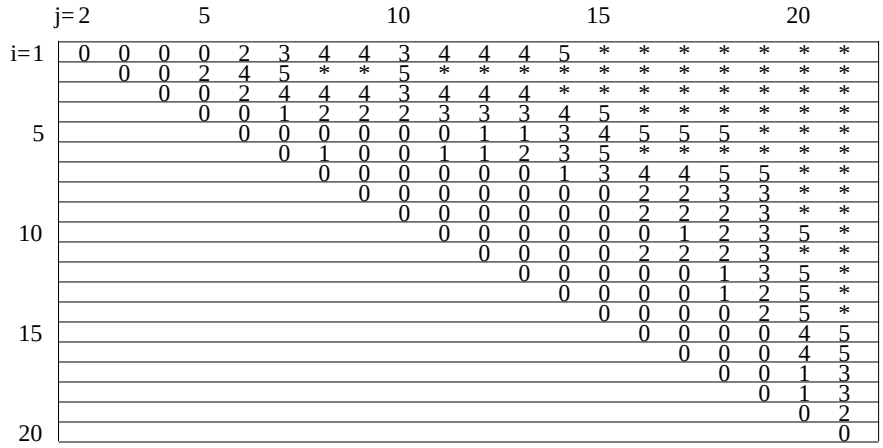

Fig. 3: The *p*-values p_{ij} ($1 \le i < j \le 21$) in Γ_1 (monofocal IOL). Symbols: [*] $\le 10^{-5}$ < [5] $\le 10^{-4}$ < [4] $\le 10^{-3}$ < [3] $\le 10^{-2}$ < [2] $\le 0{:}03$ < [1] $\le 0{:}05$ < [0]

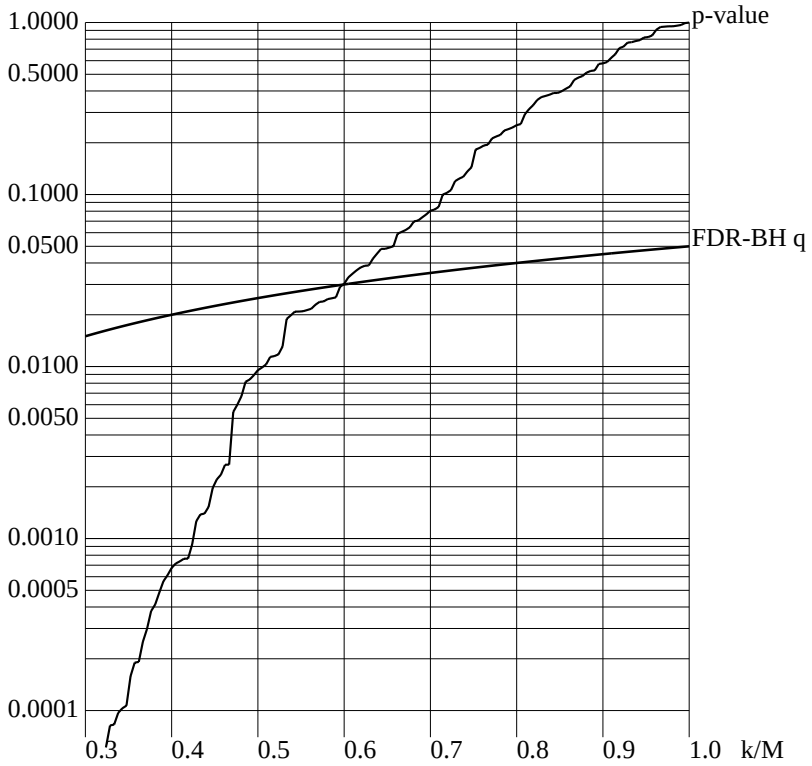

Fig. 4: FDR-BH q_k and *p*-values p_k ($k = 1, \ldots, M$) in Γ_1 (monofocal IOL).

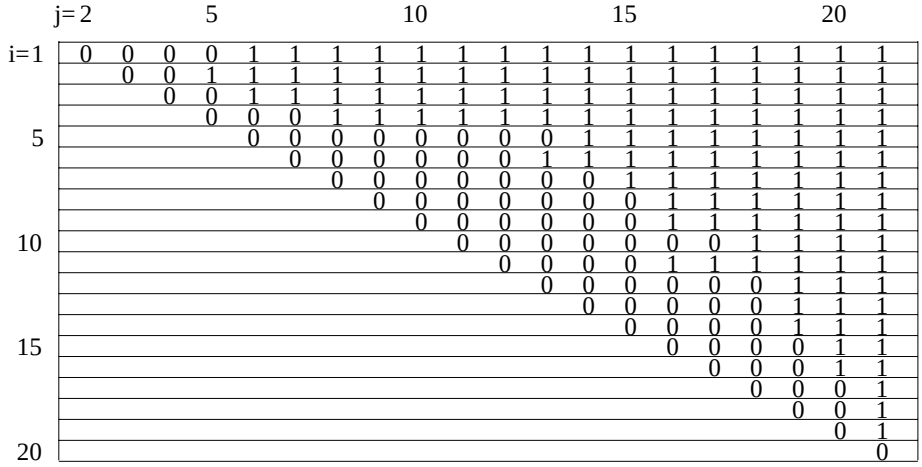

Fig. 5: The significance results of FDR-BH multiple test for null hypotheses H_{ij}^0 ($1 \le i < j \le 21$) in Γ_1 (monofocal IOL). Symbols: [0] means not-significant, [1] means significant.

Let M be the number of null hypotheses (the number of patient pairs (i, j) $(1 \le i < j \le N)$) in the group Γ_1 , and let p_k $(k = 1, \cdots, M)$ be the ordered p-values of p_{ij} such that

$$p_1 \le p_2 \le \cdots \le p_M.$$

FDR-BH q_k $(k = 1, \cdots, M)$ and p-values p_k are shown in Figure 4 (Γ_1), where we adopt FDR at $q^* = 0.05$. The significance results of multiple test for null hypotheses H_{ij}^0 $(1 \le i < j \le N)$ are shown in Figure 5, where [0] means not-significant and [1] means significant.

5.3. MULTIPLE TEST OF INDIVIDUAL PATIENTS BY HOLM AND BONFERRONI PROCEDURES (FWER)

In the Holm procedure (Holm , 1979), the significance level at kth step is chosen such that

$$\alpha_k = \frac{\alpha_0}{M - k + 1} \qquad (k = 1, \cdots, M), \tag{14}$$

where α_0 is the type I familywise error rate (FWER).

In the Bonferroni procedure, the significance level at kth step is chosen such that

$$\alpha_k = \frac{\alpha_0}{M} \qquad (k = 1, \cdots, M), \tag{15}$$

where α_0 is the type I familywise error rate (FWER).

Holm and Bonferroni α_k $(k = 1, \cdots, M)$ and p-values p_k are shown in Figures 6 (Holm; Γ_1), 7 (Bonferroni; Γ_1), where we adopt FWER at $\alpha_0 = 0.05$.

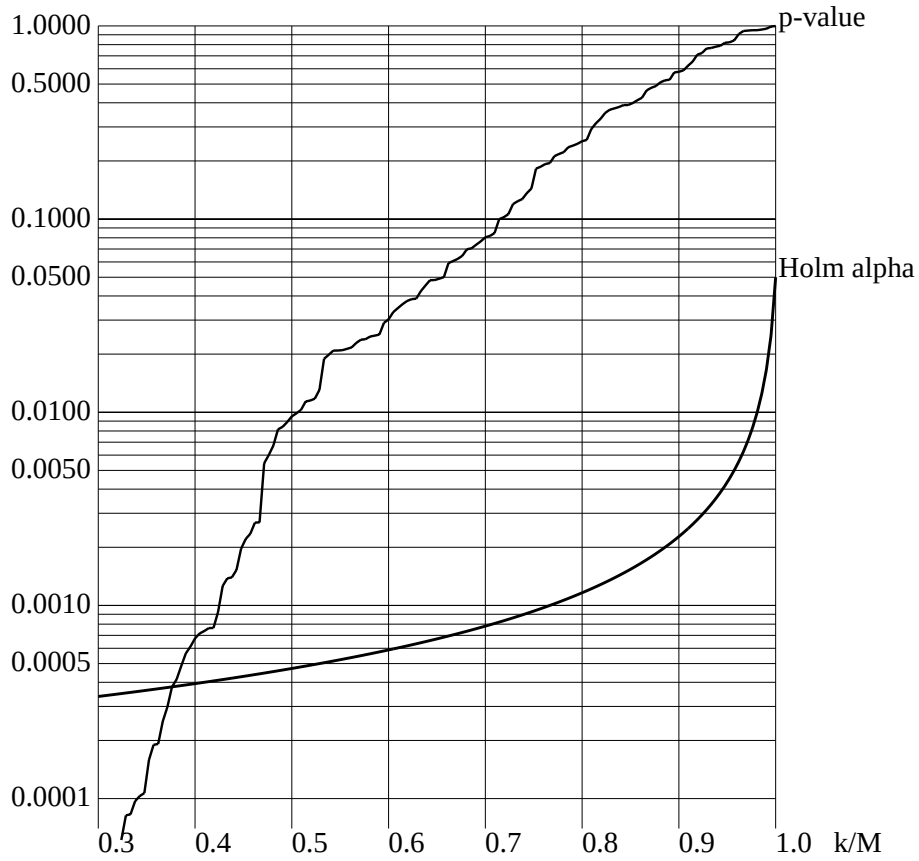

Fig. 6: Holm α_k and *p*-values p_k (*k* = 1, …, *M*) in Γ_1 (monofocal IOL).

5.4. STATISTICAL POWERS

Let χ_k^2 be the critical value for q_k (or α_k in the case of FWER) such that

$$\Pr[X^2 > \chi_k^2] = q_k \quad (k = 1, \cdots, M), \tag{16}$$

where $\Pr[X^2 > \chi_k^2]$ is the upper probability of the chi-squared distribution X^2 with 1 degree of freedom. Then, the statistical power η_k $(k = 1, \cdots, M)$ can be obtained by

$$\eta_k = \Pr[\tilde{X}^2(\theta) > \chi_k^2] \quad (k = 1, \cdots, M), \tag{17}$$

where $\Pr[\tilde{X}^2(\theta) > \chi_k^2]$ is the upper probability of the non-central chi-squared distribution $\tilde{X}^2(\theta)$ with 1 degree of freedom, and the non-centrality parameter θ is chosen as $\theta = \Delta_k$ (Δ_k is the delta test statistic which is referred to *p*-value p_k).

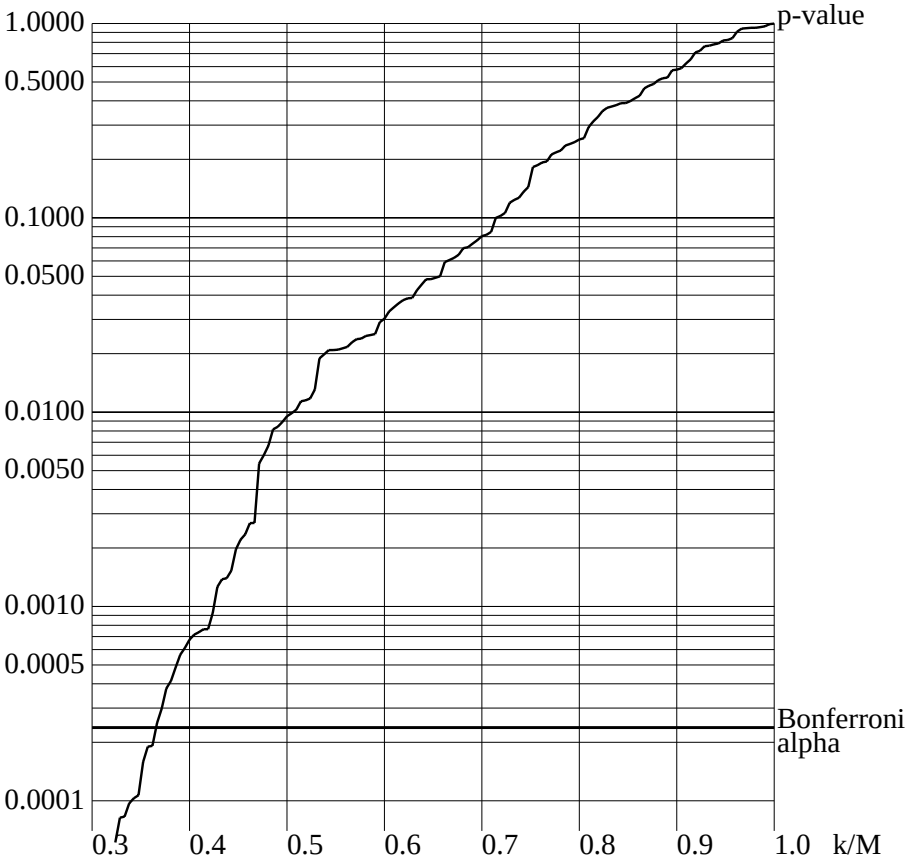

Fig. 7: Bonferroni α_k and *p*-values p_k (k = 1,…,M) in Γ_1 (monofocal IOL).

The average statistical power $\overline{\eta}$ is given by

$$\overline{\eta} = \frac{1}{M}\sum_{k=1}^{M}\eta_k. \tag{18}$$

Table 4 shows the average statistical powers of multiple tests by FDR-BH, Holm and Bonferroni procedures.

Tab. 4: Average statistical powers hby $\overline{\eta}$ FDR-BH, Holm and Bonferroni procedures

groups	hypotheses M	FDR-BH $\overline{\eta}$ $(q* = 0.05)$	Holm $\overline{\eta}$ $(\alpha_0 = 0.05)$	Bonferroni $\overline{\eta}$ $(\alpha_0 = 0.05)$
Γ_1	${}_{21}C_2 = 210$	0.604	0.392	0.370
Γ_2	${}_{12}C_2 = 66$	0.599	0.461	0.434

6. CONCLUSION

Novelty of the method in the context of visual acuity: One of the advantages of our method based on psychometric functions is that we can obtain not only the visual acuities (= psychophysical thresholds) but also their variances of individual patients. Thus we can assess changes in the visual acuity in an individual. Another characteristic of our method is that we adopt the delta test statistic (based on psychophysical thresholds and their variances) and the FDR-BH procedure for multiple testing in visual acuity problem. Therefore we can obtain the multiple testing for large number of patients implanted with intraocular lenses during cataract surgeries. For the test of patient groups, we can use the Cochran-Mantel-Haenszel (CMH) test statistic (based on psychometric functions) which requires only the numbers of observations (does not require variances). As we adopt two test statistics (delta and CMH) in our system, we can compute wide range of test problems in visual acuity field.

Monofocal IOL and multifocal IOL: Our 5-meter visual acuity test based on psychometric functions shows that the difference between the monofocal IOL group (Tecnis® 1-piece ZCB00) and the multifocal IOL group (Tecnis® Multifocal +4.0 ZMB00) is not significant under the 5% significance level (p-value $= 0.2801$ by CMH statistic χ^2-test). Yamauchi et al. (2013) also measured visual acuity by using the decimal visual acuity chart (not based on psychometric functions) and the measured decimal values were converted to the logarithm of the minimum angle of resolution (LogMAR) scale. Yamauchi's result of uncorrected 5.0 meter visual acuity (UDVA) shows that the difference between monofocal and multifocal IOLs is not significant under the 5% significance level (p-value $= 0.4976$ by Student's t-test).

Multiple comparisons of the visual acuities of individual patients: The result of multiple testing is valuable for doctors to make a suitable plan of ophthalmological treatment for each patient after the cataract surgery. We see that FDR-BH procedure produces high average statistical power in the case of large number of null hypotheses (FDR-BH: 0.604, Holm: 0.392, Bonferroni: 0.370 for the number of null hypotheses $= 210$ in patients with Monofocal IOLs). It is very effective on the problem of multiple comparisons of individual visual acuities.

The potential of applications of this method: In the conventional visual acuity measurement, a specially trained inspector shows a visual chart (Snellen, Landolt-C or LogMAR chart) to an individual patient and checks whether the patient can recognize letters or symbols correctly. It takes a lot of time and it puts stress on patients. As we developed our new visual acuity system on a handy personal computer, all necessary steps in actual optometry site including automatic mea-

surement, data collection, logistic regression, computing test statistics and multiple testing are easily available without help of specially trained inspector. We believe that our new system works satisfactorily, however we must take more data from wide range of patients before this system becomes standard in real optometric site.

ACKNOWLEDGEMENT

The authors are grateful to the editor and anonymous referees for their valuable comments and suggestions.

REFERENCES

Agresti, A. (2002). *Categorical Data Analysis, Second Edition.* Wiley.

Bach, M. (1996). The Freiburg visual acuity test – automatic measurement of visual acuity. In *Optometry and Vision Science*. 73: 49-53.

Beck, R.W., Mokexi, P.S., Turpin, A.H., Ferris 3rd, F.L., SanGiovanni, J.P., Johnson, C.A., Birch, E.E., Chandler, D.L., Cox, T.A., Blair, R.C. and Kraker, R.T. (2003). A computerized method of visual acuity testing: Adaptation of the early treatment of diabetic retinopathy study testing protocol. In *American Journal of Ophthalmology*. 135: 194-205.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. In *Journal of the Royal Statistical Society. Series B (Methodological)*. 57: 289-300.

Benjamini, Y., Krieger, A.M. and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. In *Biometrika*. 93: 491-507.

Cardona, G., Vega, F., Gil, M.A., Varón, C., Buil, J.A. and Millán, M.S. (2018). Visual acuity and image quality in 5 diffractive intraocular lenses. In *European Journal of Ophthalmology*. 28: 36-41.

Einot, I. and Gabriel, K.R. (1975). A study of the powers of several methods of multiple comparisons. In *Journal of the American Statistical Association*. 70: 574-583.

Félix, V.B. and Menezes, A.F.B. (2018). Comparisons of ten corrections methods for t-test in multiple comparisons via Monte Carlo study. In *Electronic Journal of Applied Statistical Analysis*. 11: 74-91.

Gundersen, K.G. and Potvin, R. (2013). Comparative visual performance with monofocal and multifocal intraocular lenses. In *Clinical Ophthalmology*. 7: 1979-1985.

Holm, S. (1979). A simple sequantially rejective multiple test procedures. In *Scandinavian Journal of Statistics*. 6: 65-70.

Landis, J.R., Heyman, E.R. and Koch, G.G. (1978). Average partial association in three-way contingency tables: a review and discussion of alternative tests. In *International Statistical Review*. 46: 237-254.

Mita, N., Hara, H., Kani, K. and Tabuchi, A. (2010). Use of statistical analysis for visual acuity measurement (in Japanese). In *Japanese Journal of Visual Science*. 31: 19-25.

Mita, N., Sasaki, H., Kani, K., Tabuchi, A. and Hara, H. (2014). Multiple comparisons for a psychophysical test in bootstrap logistic regression. In *Journal of Algorithms and Computational Technology*. 8: 339-355.

Mita, N., Sasaki, H., Kani, K., Tabuchi, A. and Hara, H. (2017). Multiple comparisons in two-factor testing of psychophysical thresholds by the delta test method. In *Communications in Statistics – Simulation and Computation*. 46: 669-678.

Mita, N., Sasaki, H., Kani, K., Tabuchi, A. and Hara, H. (2020). A statistical modelling of the visual acuity measurement and the multiple test procedure. In *Electronic Journal of Applied Statistical Analysis*. 13: 1-15.

Nagai, T., Hoshino, T. and Uchikawa, K. (2006). Statistical significance testing of thresholds estimated by constant stimuli method (in Japanese). In *Vision*, 18: 113-123.

Penfield, R.D. (2001). Assesing differential item functioning among multiple groups: A comparison of three Mantel-Haenszel procedures. In *Applied Measurement in Education*, 14: 235-259.

Ryan, T.A. (1960). Significance tests for multiple comparison of proportions, variances and other statistics. In *Psychological Bulletin*. 57: 318-328.

Schulze-Bonsel, K., Feltgen, N., Burau, H., Hansen, L. and Bach, M. (2006). Visual acuities “hand motion” and “counting fingers” can be quatified with the Freiburg visual acuity test. In *Investigative Ophthalmology and Visual Science*. 47: 1236-1240.

Somes, G.W. (1986). The generalized Mantel-Haenszel Statistic. In *Amercan Statistician*. 40: 106-108.

Tokutake, T., Yamauchi, T., Mita, N., Kawamoto, K., Tabuchi, H., Kani, K., Hara, H., Tabuchi, A. and Sasaki, H. (2014). Repeatability of visual acuity testing using a psychometric function. In *Japanese Journal of Ophthalmology*. 58: 381-387.

Welsch, R.E. (1977). Stepwise multiple comparison procedures. In *Journal of the American Statistical Association*. 72: 566-575.

Yamauchi, T., Tabuchi, H., Takase, K., Ohsugi, H., Ohara, Z. and Kiuchi, Y. (2013). Comparison of visual performance of multifocal intraocular lenses with same material monofocal intraocular lenses. In *Plos One*, 8: issue 6, e68236.