THE BRITISH JOURNAL

OF

OPHTHALMOLOGY

FEBRUARY, 1941

COMMUNICATIONS

THE CHI-SQUARE TEST OF SIGNIFICANCE APPLIED TO A SERIES OF INTRA-CAPSULAR CATARACT EXTRACTIONS

BY

W. J. B. RIDDELL

GLASGOW

THE object of this paper is to extract information from a small series of intra-capsular cataract operations carried out upon hospital and private patients. The majority of ophthalmic surgeons in this country have to form their views either from a limited experience or else depend for their judgments upon authority at second hand. As clinical opinion is frequently based upon small numbers it seems desirable that an attempt should be made to weigh the probabilities. It is not necessary to know anything about the inner workings of a watch in order to tell the time of day and similarly we may employ complex statistical tests to control our clinical enthusiasms without pretending to be statisticians.

The intra-capsular operation was considered to be a suitable one in 88 out of 122 consecutive patients. The final decision as to the type of operation was not made until the section was completed. All the patients were in danger of having the operation performed and the reasons for rejection were those commonly described by other writers, inadequate control of the eye, insufficient working room for the intra-capsular forceps, too small a section, insufficient

dilatation of the pupil and so on. Any patient into whose eye intra-capsular forceps were inserted is classified as an intra-capsular extraction. This definition is used because it is in the manoeuvring of the forceps that the principal danger of the operation lies. The method employed was essentially that described by Sinclair (1932), and his single action forceps were used in most of the cases. A pair modified to my own design were found to have no advantages. The large majority of the operations were done without a skilled assistant and a conjunctival suture was used only on a few occasions.

There were 49 men and 73 women in the series, and it may be asked if this ratio does not show an excess of women. The mean age of the patients of both sexes was practically 65 years, with a standard deviation of 8 years. From the 1931 census figures for the West of Scotland the expected proportion in the age.groups concerned was estimated to be 57·54 men and 64·46 women. From these figures it can be calculated whether the observed ratio differs from the expected one to a greater degree than might arise by chance. One method of doing this is shown in Table I.

TABLE I

Observed and Expected Sex Incidence

	Expected (m)	Observed (x)	Difference (x-m)	$\Sigma \frac{(x-m)^2}{m}$
Men ...	57·54	49	−8·54	1·2675
Women ...	64·46	73	+8·54	1·1278
Total ...	122	122		$\chi^2 =$ 2·3953

The sum of the right hand column gives the value of a statistic known as chi-square (χ^2). As there is only one possible way of distributing the data there is only one degree of freedom (n). The probability (P) is obtained from a table showing the distribution of χ^2 for one degree of freedom and it is found to lie between 0·10 and 0·20. This test of significance has been elaborated for small samples by Professor R. A. Fisher although the distribution of χ^2 was established by Karl Pearson in 1900 (Fisher, 1938). The exact relationship between the χ^2 value and the probability (P) is an extremely complex one, but Fisher's table is suitable for the type of problem which arises in clinical work and is readily available (Mainland, 1938, Hill, 1939). The level of 0·05 or 5 per cent. is the usual dividing line which is made between events which are likely to have arisen by chance and those which are unlikely. In this case the result is above this conventional level and we may say that the observed figures do not show a significant excess of women. At the same time the probability is sufficiently low to arouse doubt, and on finer grouping of the material it was found that one hospital at one

period did in fact show a figure of 0·05 compared with 0·30 to 0·40 in the others. With the combined figures the higher probabilities were reduced and the figure 0·10 emerged.

It may be asked if the ratio of intra-capsular operations to other methods in hospital patients is no different from that in private patients. Such a proposition is called a null hypothesis and the object of the χ^2 test is to disprove it. The observed and expected ratios are shown in Table II, and it is quite obvious that the hypothesis is not disproved. The expected figures are calculated from the marginal totals. Only one cell requires to be filled by this

TABLE II

Distribution of hospital and private patients relative to the type of operation

	Observed				Expected		
	Intra-capsular	Other methods			Intra-capsular	Other methods	
Hospital	70	28	98	Hospital	70·69	27·31	98
Private	18	6	24	Private	17·31	6·69	24
	88	34	122		88	34	122

$\chi^2 = 0{\cdot}1223$ (Yates Correction 0·0091)
$n = 1$ $P = 0{\cdot}90$

means because the other three can be found by subtraction. There is therefore only one degree of freedom. In a 2×2 table such as this, where the expected number in any cell is five or less, a slightly different method of calculation is advisable, known as Yates' correction for continuity (Fisher, 1938). The χ^2 value by the ordinary method is 0·1223, and with Yates' correction it is 0·0091. This makes it highly probable that the hypothesis is true.

By constructing similar tables from the data it can be shown that there was no significant difference in the treatment between right and left eyes ($\chi^2=0.1284$ and $P=0{\cdot}70$) and that the sexes were treated alike with regard to the type of operation performed ($\chi^2=1{\cdot}3352$ and $P=0{\cdot}25$).

The vision recorded in all the intra-capsular patients is shown in Table III, and by arranging the data in age groups a form of correlation table is made. The visual acuity was entered in the cell appropriate to the last completed test line, thus 6/12 partly was classified as 6/18. The null hypothesis that there is no difference in the visual result between patients 64 years and under and those 65 years and over can be tested by re-grouping the material. This must be done because the test for significance becomes excessively complicated if any of the cells have small expectations. The

TABLE III

Visual results and age distribution

Vision	6/6	6/9	6/12	6/18	6/24	6/36	6/60	Bad	Total
AGE : 59 and under ...	17	3	1	1	2	0	0	0	24
60-64	10	0	2	1	0	0	1	1	15
65-69	10	7	5	1	1	0	0	1	25
70 and over ...	1	10	4	+	2	0	1	2	24
TOTAL	38	20	12	7	5	0	2	4	88

TABLE IV

Grouped visual results and age

Observed

	6/6	6/9	6/12 & 6/18	6/24 & worse	
64 and under	27	3	5	4	39
65 and over	11	17	14	7	49
	38	20	19	11	88

Expected

	6/6	6/9	6/12 & 6/18	6/24 & worse	
64 and under	16·84	8·86	8·42	4·88	39
65 and over	21·16	11·14	10·58	6·12	49
	38	20	19	11	88

Contributions to χ^2

64 and under	6·0498	3·8757	1·3891	0·1585
65 and over	4·8784	3·0825	1·1055	0·1265

$\chi^2 = 20{\cdot}666$ $\quad n = 3$ $\quad P =$ less than 0·01

observed and expected figures with the contributions made to the total χ^2 from each cell is shown in Table IV. In this test three expected figures have to be calculated before the remaining cells can be filled in by subtraction from the marginal totals. There are therefore three degrees of freedom. With the grouping employed the χ^2 value is very high and the probability (P) that the visual result is unaffected by the age of the patient is extremely unlikely. This highly significant result from the test must be interpreted in

the light of the clinical facts and it raises the question whether we are justified in expecting our elderly patients to reach a standard of 6/6. The highest contributions to the total χ^2 are made by the 6/6 and 6/9 cells; there are too many younger patients with 6/6 and too few with 6/9. If these two cells are combined in both age groups the number of degrees of freedom becomes 2, $\chi^2=4{\cdot}0714$ and P. rises to 0·10 which is above the conventional level. This does not prove that no real difference is present with the data arranged in this way, but the null hypothesis is not disproved.

The null hypothesis that there is no difference between the five year age group distribution and vision of 6/9 or better and 6/12 or worse can be tested. The distribution is shown in Table V. The χ^2 value is 7·600, n=3 and P=0·05. This is at the level of significance and it is open to the observer to make what interpretation he chooses. I am indebted to Professor Fisher for the remainder of this paragraph. He pointed out to me that confidence in the reality of the effect of age is strengthened by the fact that the results in the four age groups fall out consistently in order of age, the contrast being greatest between the oldest and the youngest groups. One way of allowing this fact to weigh with the test is to consider only the 2×2 table supplied by the extreme classes in Table V, the youngest 20 and 4 and the oldest 11 and 13. Here with only one degree of freedom $\chi^2=7{\cdot}38$ showing that nearly all the heterogeneity observed among the four age groups is due to the contrast between the youngest and the oldest. With Yates' correction $\chi^2=5{\cdot}8$, so that P is well below 2 per cent.

TABLE V

Grouped visual results and age

	Observed		Expected		Contribution to χ^2	
	6/6-6/9	6/12 & worse	6/6-6/9	6/12 & worse		
59 and under	20	4	15·82	8·18	1·105	2·135
60-64 ...	10	5	9·88	5·12	0·001	0·003
65-69 ...	17	8	16·48	8·52	0·016	0·031
70 and over	11	13	15·82	8·18	1·468	2·840
Total ...	58	30	58	30		

$\chi^2 = 7{\cdot}600 \quad n = 3 \quad P = 0{\cdot}05$

Considered in conjunction with Table IV, I am of the opinion that a real difference is present. If the data are arranged in age groups as in Table V with the 6/6 patients tested against those

with 6/9 or worse, a result similar to that in Table IV is obtained with $\chi^2=27{\cdot}259$, n=3 and P very low, disproving the null hypothesis.

In 52 of the 88 patients the lens was removed completely intact. In 36 the capsule was ruptured either deliberately or by the forceps, the lens parting company involuntarily. The series was carried out with certain basic prejudices against the procedure and at a time when my operative work was limited. Feelings akin to those of a small boy stealing apples from an orchard tended to make delivery more rapid than was desirable and led to frequent rupture of the capsule at the final stage. This happened evenly throughout the series and in all the age groups (Table VI, A and B). In the latter group $\chi^2=2{\cdot}2439$, n=3 and P=0·50 and if the two younger and two older age groups are combined into a 2×2 table it becomes highly improbable that there is any difference between the ratios, the χ^2 value is lowered to 0·00392 and P=0·95. This technical failure appears to affect the visual result (Table VII) and by the test a significant difference is shown to be present ($\chi^2=8{\cdot}9765$, n=3 and P=0·02). This supports those who stress the superior visual results in successful intra-capsular extraction. It would appear that if the operation is confined to patients under 65 the better results are attributable to age rather than to the technique.

For purposes of comparison a series of fifty cataract extractions was compiled where the object was to carry out a simple extraction with a peripheral iridectomy followed by an early capsulotomy. The mean age of the series was 66·8 with a standard deviation of 7·5 years. The 34 patients of the original series could not be used as

TABLE VI

Intact and ruptured capsule

A
Serial distribution

Serial No.	Intact	Ruptured	
1—22 ...	14	8	
23—44 ...	13	9	
45—66 ...	13	9	
67—88 ...	12	10	
Totals ...	52	36	88

χ^2 value very low

B
Age distribution

	Intact	Ruptured	
Under 59	12	12	24
60—64 ...	11	4	15
65—69 ...	14	11	25
Over 70 ...	15	9	24
	52	36	88

$\chi^2 = 2{\cdot}2439$
n = 3
P = 0·50

TABLE VII

Effect of rupture of the capsule upon visual acuity

Observed

	6/6	6/9	6/12	6/18 & worse	
Intact	27	10	9	6	52
Ruptured ...	11	10	3	12	36
Totals	38	20	12	18	88

Expected

	6/6	6/9	6/12	6/18 & worse	
Intact	22·5	11·8	7·1	10·6	52
Ruptured ...	15·5	8·2	4·9	7·4	36
Total	38	20	12	18	88

$\chi^2 = 8{\cdot}9765$ n = 3 P = 0·02

a basis of comparison because they were subjected to obvious selection. The observed differences and the χ^2 values are shown in Table VIII. The successful intra-capsular cases are just significantly better than the extra-capsular. There is no significant

TABLE VIII

Intra-capsular extraction compared with simple extraction with peripheral iridectomy

Lens delivered intact

	6/6	6/9	6/12	6/18 & worse	
Intra-capsular ...	27	10	9	6	52
Extra-capsular ...	14	16	7	13	50
Total	41	26	16	19	102

$\chi^2 = 8{\cdot}3014$ n = 3 P = under 0·05

Rupture of lens capsule

	6/6	6/9	6/12	6/18 & worse	
Intra-capsular ...	11	10	3	12	36
Extra-capsular ...	14	16	7	13	50
Total	25	26	10	25	86

$\chi^2 = 1{\cdot}1637$ n = 3 P = 0·65

difference between the ruptured capsule group and the extra-capsular operation. There is no reason to suppose that the two series did not come from a universe represented by the marginal totals. This conclusion is of some importance because it shows that if the operation was not technically successful the patients in my hands did not have a result any worse than those not exposed to the risks of intra-capsular extraction.

A matter of interest arises concerning operation upon the second eye. On eight occasions in this series one eye was done by the intra-capsular method and the other with a simple extraction with peripheral iridectomy in view. These cases are shown in Table IX. It is obvious that by this method it is possible to stabilise many

TABLE IX

Sex, age, and visual result in patients operated upon by both methods

No.	Sex	Age	Intra-capsular operation		Extra-capsular operation	
1.	M	52	R.E.	6/6	L.E.	6/6
2.	F	63	L.E.	6/6	R.E.	6/6
3.	M	64	L.E.	6/6	R.E.	6/6
4.	M	60	R.E.	6/6	L.E.	6/9
5.	F	60	R.E.	6/6	L.E.	6/9
6.	M	62	L.E.	6/9	R.E.	6/9
7.	F	65	R.E.	6 6	L.E.	6/12*
8.	F	62	L.E.	6/6	R.E.	6/18*

(*Patients 7 and 8 declined to have a capsulotomy performed.)

random factors which may affect the ultimate result and it appears to be the most likely way of forming an opinion as to the ultimate results over a period of years. It will be observed that two patients were unwilling to submit to a capsulotomy. They were quite happy with reversible glasses which they seldom altered. It is in this matter of prognosis after the operation that our knowledge appears to be most defective. Personally I would prefer to have an intra-capsular extraction performed upon one eye and an extra-capsular performed upon the other.

An account of adventures with individual cases is deliberately omitted because if complications had led to frequent disaster the results would have appeared in the tables. This has not occurred. In 2 out of the 88 patients I regret that the procedure was attempted. The operation is not one which should be attempted without meticulous attention to the detail described by such masters of the technique as Dr. A. H. H. Sinclair.

Summary

In a series of 122 senile cataract extractions the intra-capsular method was considered to be suitable in 88. The results have been treated statistically by the χ^2 test of significance. The sex ratio observed, the treatment of hospital and private patients, the treatment of right and left eyes and of the sexes have all been considered. A better visual result was found where the lens was removed intact within its capsule than in those cases where the capsule ruptured. Rupture of the capsule occurred evenly throughout the series and in all the age groups. Evidence is presented to show that the better visual results tend to occur below rather than over the age of 65, and that this is not necessarily attributable to the type of operation. Comparison is made with fifty patients in whom the operation intended was a simple extraction with a peripheral iridectomy followed by early capsulotomy. The intact intra-capsular cases gave the better visual result and those which ruptured did not differ significantly from the control group. In eight patients the two eyes were submitted to the different operations.

I have great pleasure in acknowledging to Dr. A. H. H. Sinclair the great help and encouragement which he has always given to me and I am indebted to Dr. John Gilchrist for permission to operate upon many cases primarily under his care.

The design of this paper is intended to show an application to clinical problems of the methods so ably advocated by Professor R. A. Fisher. Although he is not responsible for the accuracy of the calculations I am obviously deeply in his debt both for the method of approach and for his kindly criticism.

REFERENCES

1. FISHER, R. A.—The Design of Experiments. Second Edition, Edinburgh, 1937.
2. ——— Statistical Methods for Research Workers. Seventh Edition, Edinburgh, 1938.
3. HILL, A. BRADFORD.—Principles of Medical Statistics, Second Edition, London, 1939.
4. MAINLAND, DONALD.—The Treatment of Clinical and Laboratory Data, Edinburgh, 1938.
5. SINCLAIR, A. H. H.—*Trans. Ophthal. Soc. U.K.*, Vol. LII, 1932.