

# MEASUREMENTS IN A FIELD SURVEY OF TRACHOMA BEFORE AND AFTER TREATMENT\*

A. M. ADELSTEIN, M.D., D.P.H. (RAND), *Johannesburg*

This paper describes the results of a survey, conducted by 2 ophthalmic surgeons, of trachoma in African school children at 4 schools in the Potgietersrust district of Northern Transvaal. The primary objective was to obtain a measure of the effect of certain routine treatments. In doing so it was necessary to measure the inconsistency in a surgeon's diagnosis from time to time and the differences in diagnoses which 2 surgeons manifest when examining the same cases (i.e. the variation *within* and *between* surgeons).

The diagnosis of trachoma in mass surveys raises a number of problems, as is discussed by Scott†.<sup>1</sup> The signs are not specific and may be confused with those of other conditions. Varying degrees of activity and healing may be present. It is also known that diagnosis is subject to uncertainty in general.<sup>2,7</sup> This has been discussed and analysed in various medical fields, especially X-ray diagnosis.<sup>3-5,7</sup>

This type of field survey is in essence a rapid method under difficult conditions. Several hundred cases may be examined in an hour. Apart from the expected variation in judgment other errors, such as errors by recorders, may occur.

## *Method*

The 4 schools were visited twice, first on 24 and 25 September 1958 and then on 2 and 3 March 1959. After the first

visit the children of 3 of the schools were given routine treatment for trachoma while those of the 4th school received a placebo treatment. The identity of the placebo school was not known to the examining surgeons. Routine treatment was carried out by the teachers, who were instructed to apply the antibiotic ointment twice a day for the first 3 days of each month. All children, healthy and affected, were to be treated. The drugs used were ilotycin, mycitracin and achromycin, and the placebo was a bland ointment.

Each child was given a numbered card on which appeared his full name, age and class. The two ophthalmic surgeons working out of doors in clear daylight were separated by a distance sufficient to ensure no possibility of seeing or hearing each other's diagnosis. The surgeons diagnosed either 'normal' or 'trachoma' and each diagnosis was recorded by a clerk. Routine examination consisted of pulling down the lower lids and examining by naked eye the conjunctiva of the lower lid, the fornix and the cornea and then also everting the upper lid (this eversion was omitted in a small series to be described later).

## *Results*

The prevalence of trachoma is known to become less as the children get older. The percentage of trachoma cases according to the average of the 2 surgeons' readings were on the first visit before treatment.

\* A paper presented at the 42nd South African Medical Congress (M.A.S.A.), East London, C.P., September-October 1959.

† See page 441 of this issue.

Up to age 10 years	25.3% (576 children examined)
Between 11 and 14 years	13.8% (476 children examined)
Age 15 and over	11.7% (102 children examined)

## FIRST VISIT

*Variation within and between Surgeons*

A comparison of the findings of the two surgeons is shown in Table I. The results at each school appear in a column. For example at school I, 166 children were seen by both surgeons. They agreed that 115 were normal and that 26 had trachoma; 15 cases were diagnosed as trachoma by

TABLE I. EXAMINATIONS BEFORE TREATMENT AT 4 SCHOOLS BY 2 SURGEONS

	I	II	III	IV	Total
Agreed Normal .. ..	115	163	275	302	855
Agreed Trachoma .. ..	26	39	34	50	149
Disagreements:					
Surgeon A: Trachoma ..	15	31	23	21	90
Surgeon B: Trachoma ..	10	16	12	22	60
Total .. ..	<u>166</u>	<u>249</u>	<u>344</u>	<u>395</u>	<u>1,154</u>
Disagreements as a % of:					
Trachoma cases .. ..	49.0	54.6	50.7	46.2	50.2
Total cases .. ..	15.1	18.8	10.1	10.8	12.9
% Trachoma diagnosed					
{ Surgeon A .. ..	24.6	28.1	16.5	17.9	20.7
{ Surgeon B .. ..	21.6	22.0	13.3	18.2	18.1
$\chi^2$ between disagreements	1	4.7	3.4	0.2	6
p .. ..	.3	.05	.10	.90	.02

surgeon A and were called normal by B while 10 cases were called trachoma by B and normal by A. Thus A diagnosed 41 cases (26+15) and B 36 cases (26+10). Surgeon A diagnosed more cases in 3 of the schools. In the total of 1,154 children he diagnosed 239 cases (i.e. 149 agreed and 90 disagreed) while surgeon B diagnosed 209 cases (149+60), the percentage prevalence being 20.71% for surgeon A and 18.11% for surgeon B. A test of significance for this difference

TABLE II. EXAMINATIONS AFTER TREATMENT AT 4 SCHOOLS BY 2 SURGEONS

	I	II	III	IV	Total
Agreed normal .. ..	79	117	202	239	637
Agreed trachoma .. ..	13	6	18	17	54
Disagreements:					
Surgeon A: Trachoma ..	17	10	23	23	73
Surgeon B: Trachoma ..	0	1	4	1	6
Total .. ..	<u>109</u>	<u>134</u>	<u>247</u>	<u>280</u>	<u>770</u>
Disagreement as % of:					
Trachoma cases .. ..	56.6	64.7	60.0	58.5	59.4
Total cases .. ..	15.5	8.2	10.9	8.5	10.2
% Trachoma diagnosed:					
Surgeon A .. ..	27.5	11.9	16.5	14.2	16.4
Surgeon B .. ..	11.9	5.2	8.9	6.4	7.7
$\chi$ between disagreements	15.0	5.8	12.0	18.3	56.82
p .. ..	.001	.02	.001	.001	.001

(which allows for the correlation in that the same cases are examined by both surgeons) may be performed.<sup>6</sup> If there were no difference in the surgeons' results, the disagreements (90, 60) should differ by no more than chance.  $\chi^2 = 6$  (1 d.f.) This gives a p value of less than .02 which may be considered significant. However, the differences are significant at only one of the individual schools.

The amount of disagreement may be related to the total children examined or to the trachoma cases diagnosed since there are a large majority of obviously normal children. Thus the disagreements in the 4 schools would be 150 (90+60) out of 299 cases diagnosed trachoma by both surgeons, i.e. 50%. If the disagreements are related to the total children, there are 13%.

Such a level of disagreement is not unexpected, because as pointed out by Garland<sup>7</sup> it is found in a variety of clinical procedures when tested even under the best conditions. Garland states that it is probably 'quite universal'. The difference in the rates diagnosed by the two surgeons is only 2.6% in the total cases, which is surprisingly small when the difficult conditions are considered. Among the diagnoses of each doctor there will be a certain number of false negatives as well as false positives and the resulting total is nearer to the true total.

*Variation within surgeons.* At one school 134 cases were examined twice by both surgeons. After being examined the 134 children formed a second queue and were re-examined immediately. A slight modification in examination technique was introduced. Surgeon B on his first round of examination did not evert all the upper lids, but only those which he

TABLE III. VARIATION WITHIN AND BETWEEN SURGEONS IN 134 CASES SEEN TWICE BY BOTH SURGEONS

		(a) Within Surgeons	
		Surgeon A	Surgeon B
Agreed Normal on both readings .. ..		98	94
Agreed Trachoma on both readings .. ..		29	21
Disagreements	{		
Trachoma 1st reading .. ..		4	7
Trachoma 2nd reading .. ..		3	12
Total .. ..		<u>134</u>	<u>134</u>
% Trachoma diagnosed at:			
1st reading .. ..		24.63	20.89
2nd reading .. ..		23.88	24.63
Disagreement as a % of Trachoma .. ..		19.4	47.5
$\chi^2$ between disagreements .. ..		0	.84
p .. ..		1	.6
		(b) Between Surgeons	
		1st Examination	2nd Examination
Agreed Normal by both surgeons .. ..		96	95
Agreed Trachoma .. ..		23	26
Disagreements	{		
Surgeon A: Trachoma .. ..		10	6
Surgeon B: Trachoma .. ..		5	7
Total .. ..		<u>134</u>	<u>134</u>
% Trachoma diagnosed	{		
Surgeon A .. ..		24.63	23.88
Surgeon B .. ..		20.89	24.63
Disagreement as a % of Trachoma .. ..		39.4	33.3
$\chi^2$ between disagreements .. ..		1	0
p .. ..		.5	1

decided merited such examination. On his second round he everted all the lids. Surgeon A everted the upper lid in all cases.

Table III shows the findings. Surgeon A has less variation—7 cases compared to 19 cases which surgeon B diagnosed differently in the two examinations. This may be accounted for by the fact that surgeon B did not evert all the upper lids on his first examination.

The percentage trachoma diagnosed by surgeon A was 24.6 on the first round and 23.8 on the second round while B diagnosed 20.8% and 24.6%. In each surgeon the difference between the first and second round is not significant.

#### SECOND VISIT

At this visit, after treatment, all the children who had previously been examined and were present were again examined and the results recorded on the relevant cards, the surgeons not knowing the previous results. Although some children who were not present at the first examination were also examined, there is a group of cases in which every individual has been examined before and after treatment and which allows the assessment of the effect of treatment without regard to age (for it is known that the prevalence varies with age). The results of the second examination are shown in Table II, in the same manner as Table I.

#### Variation within and between Surgeons

The inconsistency of diagnosis between the surgeons has now increased and surgeon B was now diagnosing less trachoma than surgeon A at every school. Surgeon A diagnosed 16.49% at the 4 schools while B diagnosed 7.79%. At each school the difference between them was significant at the 1% level.

Considering the results of the two visits, it appears that when a surgeon re-examines cases which he has previously examined on the same day he has a good level of consistency for survey purposes. But over a period of time, in this case 6 months, he may change his level of diagnostic criteria.

On the first visit, the disagreements among the first 50 cases examined were re-examined by the surgeons and criteria of diagnosis established. This was not done during the second examination, when there was no discussion of criteria. It would appear that the surgeons can be reasonably consistent within themselves at a particular time, and if they have clearly laid down and have understood each other about their criteria then they will also be reasonably consistent with each other. The variation in diagnosis arises largely from those cases which are healing or have healed but still show some signs of the disease. This suggests that in such field surveys it would be of advantage if surgeons compared and agreed on some cases before beginning a survey.

#### THE EFFECT OF TREATMENT

Because of the change in consistency between the surgeons after the first visit, some conflict of opinion has appeared.

The results of each surgeon will first be discussed separately, and then use will be made of a measure of agreement between them.

Table IV shows the results of the examinations of those children who were present at both the first and second examinations, i.e. before and after treatment. The results of each surgeon at each school are shown in a separate column, both examinations, before and after treatment, being shown together. For instance, surgeon A at school I, which was the control (placebo) school, examined 109 children on both occasions. Of these 72 were diagnosed as normal on both occasions, 17 had trachoma on both occasions, 7 had trachoma on the first occasion and were normal on the second, and 13 were normal on the first examination and were diagnosed as trachoma on the second occasion. Thus there were 24 cases of trachoma diagnosed on the first visit and 30 cases on the second visit.

Surgeon A at the control school found an increase in prevalence of trachoma but it was not significant. At the first treated school (II) he found a significant reduction of trachoma, viz. from 30.6 to 11.9%. At the other two treated schools he found reductions (from 18.2% to 15.9% and from 16.7% to 14.2%) but these would not be classed as statistically significant.

The findings of surgeon B are somewhat different. Significant reductions were found at all the schools. However, the values of  $\chi^2$  for testing these differences are clearly much greater at the three treated schools than at the untreated school. It appears that this surgeon may have altered his level of demarcation at his second visit, but that the treated schools have a more highly significant diminution, indicating the effect of treatment.

#### The Combined Result of the two Surgeons

As there is doubt when diagnosing the lesser degrees of the clinical manifestations of the disease it is postulated that a firmer basis for accuracy of comparison may be given by those cases which were agreed on by both surgeons. The agreed trachoma cases will be less than the 'true' number, but will contain a minimum of false positive cases. Thus at the control school one case agreed as normal on the first visit was agreed on as having trachoma on the second visit, and contrariwise one agreed on as trachoma on the first visit was agreed as normal on the second visit. A test is done to determine the significance of the change of the agreed normals which become agreed trachoma as compared with the opposite agreed trachoma which become agreed normals after treatment. Table V summarizes the agreed cases which changed.

TABLE IV. COMPARISON AT 4 SCHOOLS OF SAME CHILDREN BEFORE AND AFTER TREATMENT

	Surgeon A				Surgeon B			
	I (Control)	II	III	IV	I (Control)	II	III	IV
Normal on both examinations .. ..	72	91	188	220	85	101	202	228
Trachoma at both examinations .. ..	17	14	29	27	12	5	17	15
Trachoma at first and normal at second .. ..	7	27	18	20	11	26	23	34
Normal at first and trachoma at second .. ..	13	2	12	13	1	2	5	3
Total cases .. ..	109	134	247	280	109	134	247	280
% Trachoma at 1st examination .. ..	22.02	30.60	18.28	16.78	21.1	23.13	16.19	17.5
% Trachoma at 2nd examination .. ..	27.52	11.94	15.95	14.28	11.92	5.22	8.90	6.42
$\chi^2$ between difference .. ..	1.8	21.55	1.2	1.48	6.75	20.57	11.57	25.97
p .. ..	.2	.001	.3	.3	.01	.001	.001	.001

TABLE V. CASES OF AGREED DIAGNOSIS BY TWO SURGEONS WHICH CHANGED FROM AGREED NORMAL TO AGREED TRACHOMA AND VICE VERSA (4 SCHOOLS)

	I (Control)	II	III	IV
Changed from:				
Agreed Normal to Agreed Trachoma .. ..	1	0	1	3
Agreed Trachoma to Agreed Normal .. ..	1	12	8	11
$\chi^2$ .. ..	0	10	4	3.5
p .. ..	1	.005	.05	.10

In all the treated schools, more agreed trachoma cases become agreed normals than the reverse position. The levels of statistical significance vary at approximately .005, .05 and .10. On the other hand the control school has no difference, one case changing in either direction.

Summarizing the evidence for the effectiveness of the treatment: Dr. A. found an increased prevalence in the control school (not significant). He found a significant decrease in one school and decreases in both other schools although not statistically significant. Dr. B. found a statistically significant decrease at each of the schools, including the untreated control school. The levels of significance at the treated schools are considerably higher than at the untreated school.

The combined agreed effect shows no significant change at the control school and decreases at the treated schools with levels of probability of .005, .05 and .10.

Taken all together the evidence is in favour of the effectiveness of the treatment, particularly at school II but also at the other two schools.

#### The Factors in Treatment

A different drug was used at each of the three treated schools (ilotycin, mycitracin and achromycin at schools II, III and IV, respectively). The effectiveness of the drugs depends partly on the efficiency and enthusiasm of the school staff. It was observed by the examiners even before the results were assessed that school II (using ilotycin) had a better atmosphere of keenness than the other schools. The principal of school III acknowledged a break in treatment. Thus it would not be fair to compare the drugs except to note that ilotycin was particularly effective possibly because it was used more efficiently.

#### The Natural Course of the Disease

A measure of the natural history of trachoma may be derived from the cases in which the diagnosis was agreed on by the surgeons. In the control school there were 14 cases of agreed trachoma on the first visit which were re-examined on the second visit, 6 months later. Of these, 10 had remained as agreed trachoma, 3 were disagreed on and one had become an agreed normal case. This would suggest that the disease is stable and tends to be chronic in an individual. Only one case of the 76 agreed normals on the first visit became agreed trachoma on the second visit, suggesting a minimum of new infection at school ages. In the 3 treated schools a similar situation is found. Only 4 new cases of agreed trachoma were diagnosed on the second visit among the 493 agreed normal cases seen on the first visit. This low infection rate would, however, be expected as a result of the treatment. There were 32 agreed trachomas out of the 87 agreed trachomas of the first visit.

#### DISCUSSION

The method used was designed to obtain measures of variation of diagnosis within and between surgeons and also to obtain measures of the effect of treatment. In the event, the variation in one of the surgeons after a period of time during which treatment was applied proved to be large, and this factor interfered with the second aim, namely, the measurement of treatment effect. The surgeons knew that 3 of the 4 schools had received antibiotic treatment that was generally expected to be efficient. It is probable that a subjective element may have intruded because of the expected effect of the drug. Such a subjective change can occur in a routine survey and should be guarded against in survey assessments of the effect of treatment because, as this investigation shows, it may be substantial.

#### SUMMARY

The results of a field survey by two ophthalmic surgeons of trachoma at 4 schools are analysed according to (1) prevalence; (2) the variation in diagnosis of one individual surgeon at different times and the variation between different surgeons; and (3) the effect of treatment with 3 certain antibiotics.

The mean prevalence of trachoma as indicated by the two surgeon's diagnoses before treatment was 25% up to age 10 years, 13% at age 11-14 years, and 11% at age 15-18 years.

After the surgeons had consulted on a standard, the disagreement was nevertheless substantial, being 13% of all the children examined. The difference in prevalence found by the surgeons was 2.6% at the first examination. In a series of 134 cases examined twice in the same afternoon by both surgeons, there was close agreement both within and between surgeons. Six months later and without another consultation about standards, the difference in prevalence diagnosed by them was 8.7% and was highly significant.

Three of the schools had received antibiotic treatment and the 4th a placebo treatment. One surgeon found significant reductions at all the schools, including the control school, presumably because he had changed his approach on account of a subjective expectation of an effect of treatment. He found more highly significant changes at the treated schools than at the control school. The other surgeon found no significant change at the control school, a highly significant decrease at one treated school and decreases at the other treated schools which were not significant. A combined measurement of agreed cases is in favour of the effectiveness of the treatment, particularly at one of the schools.

There is evidence that, in the absence of treatment, trachoma is a chronic disease and that there is a minimum of re-infection at school ages, during which the prevalence gradually diminishes as the children get older.

This work was done on behalf of the National Council for the Blind as part of a series of investigations which were arranged by Dr. J. G. Scott. I thank the surgeons who took part for their collaboration.

#### REFERENCES

1. Scott, J. G. (1959): *S. Afr. Med. J.*, **441**.
2. Editorial (1947): *Brit. Med. J.*, **1**, 215.
3. Yerushalmy, J. (1953): *Dis. Chest*, **24**, 133.
4. Fletcher, C. M. and Oldham, P. D. (1949): *Brit. J. Industr. Med.*, **6**, 168.
5. Birkelo, C. C. et al (1947): *J. Amer. Med. Assoc.*, **133**, 359.
6. McNemar, O. (1949): *Psychological Statistics*, p. 204. New York: John Wiley & Sons.
7. Garland, L. H. (1959): *Amer. J. Roentgenol.*, **82**, 1.