J. Hyg., Camb. (1971), **69**, 717
Printed in Great Britain

Criteria for measuring the efficacy of trachoma vaccines in baboons

By L. H. COLLIER
Medical Research Council Trachoma Unit, Lister Institute of Preventive Medicine, London, S.W.1
and ELAINE LIGHTMAN
Medical Research Council Statistical Research and Services Unit, University College Hospital Medical School, London, W.C.1

(*Received* 23 *June* 1971)

SUMMARY

Trachoma vaccines are usually assayed by testing their ability to protect monkeys or baboons against subsequent challenge of the conjunctiva with a pathogenic strain of trachoma/inclusion conjunctivitis (TRIC) agent. In such experiments the course of infection in vaccinated baboons was compared in terms of arbitrary scores assigned to a range of clinical signs, and of counts of TRIC inclusions in conjunctival scrapings. Analysis of many such scores indicated that after a large challenge dose of strain MRC-4*s*, the scores for signs of inflammation reached their maximum earlier than the follicle score; the inflammation score was closely related to the number of inclusions, whereas the follicle score was not. With this system, the optimum periods for eliciting differences between vaccinated and control measures varied according to the sign used; it was later for follicles than for inflammation or inclusions. For assessing the influence of vaccination, the mean of the inflammation scores read weekly for the first 3 weeks after challenge and the mean inclusion score over the same period were equally satisfactory, and either was rather better than the mean of three follicle scores taken over the period 3–6 weeks.

For assessing the influence of vaccines or therapeutic agents on experimental trachoma it is important to determine which signs discriminate best between treated and control animals, and the optimum times for measuring them.

INTRODUCTION

The potency of trachoma vaccines is usually assessed in terms of their ability to protect against infection with a pathogenic trachoma or inclusion conjunctivitis (TRIC) agent. To do this, the challenge dose is inoculated into the conjunctivae of vaccinated monkeys or baboons, and the course of infection is compared with that in an unvaccinated control group. It is obviously desirable, first, to be able to express the severity of infection quantitatively, and second, to determine which features are most useful for estimating the influence of immunization.

In man, the main conjunctival and corneal signs of trachoma were long ago assigned scores for severity, first as 0, +, ++, +++, and later as numbers (W.H.O. Expert Committee on Trachoma, 1962). In 1966 a W.H.O. Scientific Group on Trachoma Research recommended a much more comprehensive system primarily for use in field studies. Soon after their start in the early 1960s it became apparent that a scoring system was also needed for vaccination experiments in non-human primates; but because corneal lesions and scarring do not occur in simians except in special circumstances, it did not have to be as elaborate as for man. In 1960, Grayston and co-workers assessed the potency of vaccines in monkeys solely on the presence or absence of conjunctival follicles. Collier (1961) working with baboons assigned separate scores to signs of inflammation (comprising conjunctival hyperaemia, oedema and infiltration) and to conjunctival follicles. Dawson, Jawetz, Thygeson & Hanna (1961) recorded the intensity of hyperaemia, discharge, infiltrate and conjunctival follicles; on the basis of specificity and reproducibility of assessment they considered that follicles were the most suitable lesions for scoring, and used such scores to compare the infections induced by different TRIC agents and to assess the influence of vaccination. Mordhorst (1967) later extended this system to include six more physical signs, the proportion of inflammatory cells in conjunctival scrapings and the number of inclusion bodies. Wang (1967) assigned numerical scores to a range of physical signs in monkeys and used them to compare the infections induced by different TRIC agents and to assess the influence of vaccines and penicillin treatment. Collier & Blyth (1966*a*) were the first to employ statistical techniques; they described a scoring system for infection in baboons, and used the values in analyses of variance to determine the significance of differences in the responses to challenge of vaccinated and control animals. The desirability of introducing statistical methods was emphasized by the wide ranges of the 95% confidence limits on scores for individual vaccinated animals.

From what has already been said, there is obviously much variation in the criteria employed somewhat arbitrarily by different workers for measuring the severity of infection. Those used so far fall into three categories: clinical signs, appearance in conjunctival scrapings of inflammatory cells, and inclusion bodies. Of these, only the presence of inclusions is pathognomonic, and it therefore serves as a base line to which other variables can be related. With regard to the clinical appearances, Collier (1967) pointed out that because follicular hyperplasia is the dominant lesion it has been used by some as the sole index of severity of infection; but since these lymphoid follicles are probably part of the antibody-forming apparatus they could be regarded as secondary lesions that may not provide such a direct indication of severity as do the signs of inflammation – lid oedema, discharge, conjunctival infiltration and hyperaemia – which reach their maximum earlier. This supposition was borne out by the results of examining 41 unvaccinated baboons 4 times at weekly intervals after conjunctival inoculation with the MRC-4*s* strain of TRIC agent. There was a highly significant positive correlation between the numbers of inclusion bodies and the cumulative scores for signs of inflammation, but none between the inclusion counts and scores for follicles. The

present paper describes the further analysis of these and other data, and in particular the value of various criteria for assessing the degree of immunity induced by trachoma vaccines.

MATERIALS AND METHODS

Immunization experiments in baboons

Part of the data to be analysed are derived from five experiments in which groups of young baboons (*Papio cynocephalus*) were immunized parenterally with live TRIC agents. The infection induced by inoculating a large dose of TRIC agent into one eye was compared with that in non-immunized control animals. The challenge strain was originally known as LB4; its full designation in the 'Montreal system' (Gear, Gordon, Jones & Bell, 1963) is TRIC/2/GB/MRC-4/ON. In these experiments, only the 'slow-killing' parent strain MRC-4*s* (Reeve & Taverne, 1963, 1967) was used for challenge.

These five immunization experiments have been published with the following reference numbers: no. 2 (Collier & Blyth, 1966*a*); nos. 8 and 10 (Collier & Blyth, 1966*b*); nos. I and II (Collier & Smith, 1967). In all of them, the vaccines used exerted a greater or less measure of protection as tested by analysis of variance and by determining whether the scores for individual vaccinated animals fell within the confidence limits computed for the experiment in question. In the present paper the data for unvaccinated animals are supplemented with the results of challenging control baboons with MRC-4*s* in five other experiments in which the vaccines failed to immunize.

The methods of inoculating and examining baboons and of staining inclusions by the iodine method were described by Collier (1961) and Collier & Blyth (1966*a*). Each animal was examined on the day of challenge (day 0) and in most instances on days 7, 14, 21, 28 and 42 thereafter. All examinations were made by one person (L.H.C) without reference to previous findings.

In some experiments the intervals between examinations were somewhat different, so that each of the five examinations undergone by every animal is referred to in terms of a range of days after challenge, e.g. 4–8, 9–15.

Scoring system

The inflammation score is the sum of the scores for (*a*) external oedema of the lids, (*b*) purulent discharge, (*c*) conjunctival hyperaemia and (*d*) conjunctival infiltration (recognized by loss of transparency, oedema and thickening). Each of these signs was scored 1, 2 or 3 according to severity, 0 if absent. Hyperaemia and infiltration were scored separately for the upper and lower lids.

The follicle score is the sum of those for the upper and lower lids using the same scale as for inflammation. The small superficial translucent follicles characteristic of 'non-specific folliculosis' were sometimes present in normal animals and were ignored.

Occasionally one or another lesion was present in the inoculated eye on the day of challenge, or appeared in the control eye during the course of observation.

Rules for correcting scores appropriately were based on two assumptions: (*a*) that physical signs in the control eye appearing or increasing at any time after the day of inoculation (day 0) were caused by inadvertent contamination at the time of inoculation or by subsequent cross-infection from the inoculated eye. From this it follows that scores for the control eye were subtracted from the test eye scores only when the lesion was present on day 0. (*b*) That day 0 scores would remain stationary over the comparatively short period of observation.

The rules are:

(1) If a score for a given physical sign was recorded for the control eye on day 0, this and all subsequent control scores (if any) were subtracted from the corresponding score for the inoculated eye.

(2) If the control eye score was greater than the inoculated eye score, the latter was assigned zero value (because a negative value means 'better than normal', which is unrealistic).

(3) All control eye scores not covered by (1) were ignored.

(4) If a physical sign was present in the inoculated eye on day 0, the score was subtracted from all subsequent corresponding scores for that sign, subject to rule (2).

The inclusion score was derived from the total number of inclusions in scrapings from both upper and lower lids of the inoculated eye. Because the distribution of the numbers of inclusions was decidedly skew a transformation of the number was used as the score, namely $\log_{10}$ (number of inclusions + 1). To avoid counting large numbers of inclusions, slides containing more than 30 were recorded simply as '> 30', but for statistical analysis a value of 32 was arbitrarily assigned to such readings.

RESULTS

Time course of infection in control and vaccinated baboons

Fig. 1 gives the mean values of the scores at various intervals after challenge with TRIC agent. In the 58 unvaccinated control animals the scores for inflammation and inclusions both rose rapidly, remained high for the first 3 weeks, and then diminished. By contrast, the mean follicle score rose steadily to a maximum at 24–34 days, and remained high until at least the 6th–7th week. In vaccinated animals, the inflammation and follicle scores followed the same trends as in the controls, but with lower values. The scores for inclusions behaved rather differently however; the highest score in vaccinated animals was attained within the first week, and after the 2nd week it declined to a very low value. In other words, vaccination appeared to change the shape of the inclusion score curve from concave downwards to concave upwards, whereas for the other two variables the shape remained much the same.

These findings suggest that the best indicators of the effect of vaccination would be the average score for inflammation over the period 4–22 days, or for follicles over the period 17–49 days, since the differences between the scores for vaccinated and control animals were greatest at these times. Although the shape of the inclusion curve in vaccinated animals differed from that in the controls, the average

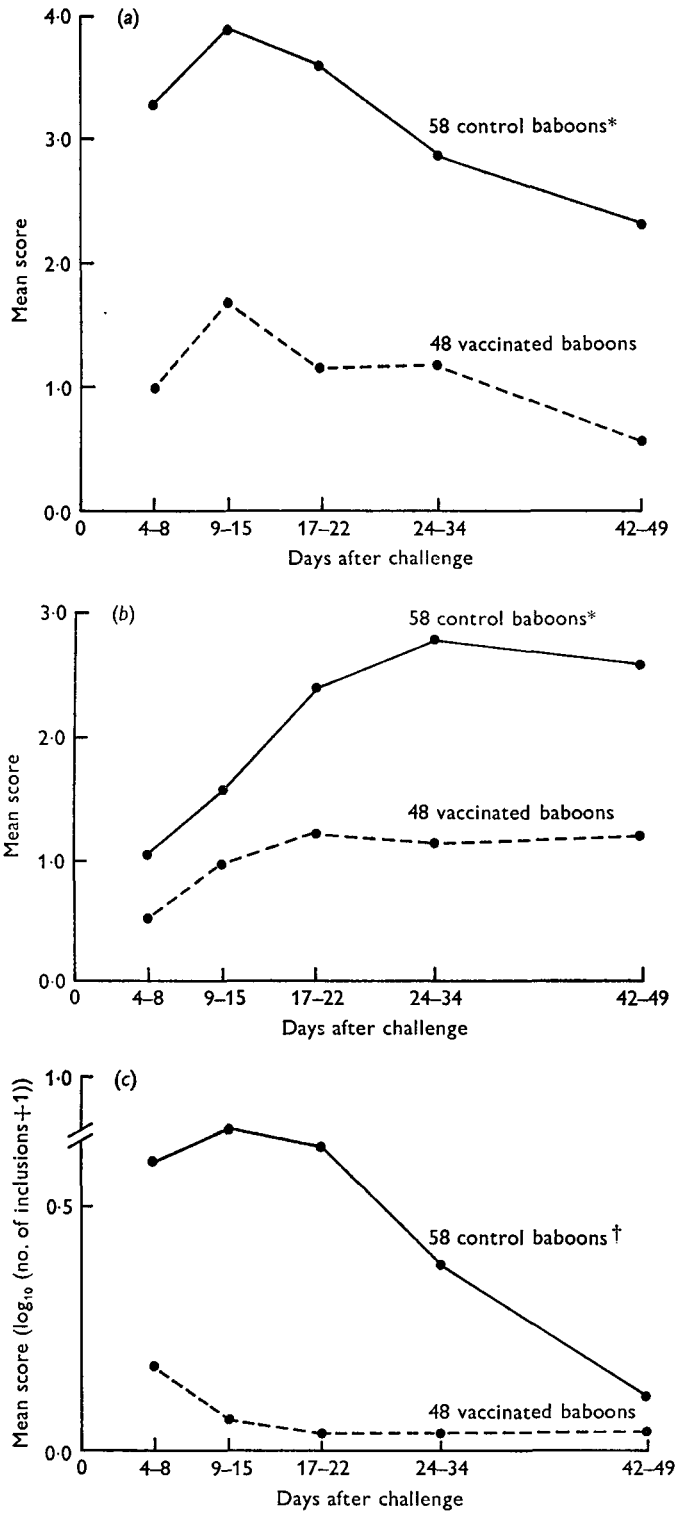

Fig. 1. Mean values of scores at various intervals after challenge of baboons with TRIC agent. The abscissae are marked at the midpoints of the periods indicated. (*a*) Scores for inflammation. (*b*) Scores for follicles. (*c*) Scores for inclusions. * Value at 42–49 days based on 45 animals only. † Value at 4–8 days based on 57 animals and at 42–49 days on 39 animals only.

score for inclusions over the period 4–22 days appeared to be the most sensitive indicator of the effect of vaccination.

Interrelationships of the variables in unvaccinated animals

Before comparing inflammation, follicles and inclusions in terms of the effects of vaccination upon them, their interrelationships in all 58 unvaccinated animals

Table 1. *Correlations between scores for inflammation (mean of readings at 4–22 days) follicles (mean of readings at 17–49 days) and inclusions (mean of readings at 4–22 days)* for 58 *unvaccinated baboons*

Variables correlated	Correlation coefficient	Probability (P)
Inflammation × follicles	0·14	> 0·100
Inflammation × inclusions	0·67	< 0·001
Follicles × inclusions	0·19	> 0·100

Table 2. *Values of Student's* t *for differences in mean scores between vaccinated and unvaccinated baboons in individual experiments*

Experiment no.	No. of baboons		Criterion for assessing efficacy of vaccination		
	Vaccinated	Control	Inflammation score (mean from 4–22 days)	Follicle score (mean from 17–49 days)	Inclusion score (mean from 4–22 days)
2	12	6	4·23***	3·12**	4·04**
8	6	6	3·23**	0·69	2·66*
10	5	5	2·63*	0·62	2·08
I	11	6	3·21**	1·22	2·89*
II	14	6	3·90**	2·76*	6·25***
All experiments	48	29	7·39***	3·95***	7·64***

* Significant at 5% level; ** 1% level; *** 0·1% level.

were examined. Table 1 shows that there was a highly significant correlation coefficient of 0·67 between the inflammation score and the inclusion score, which is perhaps not surprising since they behaved similarly with respect to time. There was, however, little or no correlation between the scores for inflammation and follicles nor between those for inclusions and follicles.

Criteria for assessing the effects of vaccination

Table 2 gives the values of Student's t for the differences in the 'best indicator' scores between vaccinated animals and their controls in the separate experiments and in all combined. In general, whenever one of the measures indicates a significant difference, the others also do. The values of t are smallest for the follicle scores, but there is little to choose between the inflammation and the inclusion scores. The values of t for all experiments combined show the same pattern.

For measuring the influence of vaccination it seems that the mean of the in-

flammation scores read at weekly intervals for the first 3 weeks after challenge with MRC-4*s*, and the mean inclusion score over the same period, are equally good; either is preferable to the best measure based on follicles, namely the mean of three follicle scores taken over the period 3–6 weeks.

DISCUSSION

The analysis reported here confirms a previous observation that the number of inclusions is closely related to the intensity of inflammation induced in the baboon eye by TRIC agent, but not to the degree of follicular hyperplasia, which attains its maximum later (Collier, 1967). It is therefore not surprising that modification of the clinical signs associated with inflammation is a rather better index of the efficacy of a vaccine than prevention of follicle formation. As we have seen, however, some workers attach more importance to follicles than to other lesions; for example, the W.H.O. Scientific Group (1966) recommended giving extra weighting to the scores for mature follicles in man. Again, Wang's system for scoring TRIC infections in monkeys is so constructed that follicles are likely to be given a greater weight than other lesions; furthermore, unless some follicles are present no score at all is given to signs of inflammation, apparently because they are 'not as specific as trachomatous follicles'. In considering whether a given feature of the disease should be scored, and if so, how it should be weighted, the purpose of the study must be considered; a system that is suitable for an epidemiological survey in man may be much less so for vaccine assays in monkeys. In this connexion it is worth recalling the distinction between the *relative intensity* of trachoma (the degree of activity in an individual case at a given time) and the *relative gravity* (the degree of disabling complications or of active lesions that will lead to such sequelae if untreated) (World Health Organization, 1962); these indices may have little in common in man (Assaad & Maxwell-Lyons, 1967) and even less in simians, in whom sequelae are very rare. For assessing relative gravity in man it may be quite justifiable to weight the score for follicles, because their extent may determine the degree of subsequent cicatrization (World Health Organization, 1962). By contrast, vaccine experiments in monkeys involve the assessment of something corresponding more closely to relative intensity, to which follicular lesions are perhaps less relevant.

Most observers, ourselves included, have relied on the clinical appearances and on the presence of specific inclusion bodies to delineate the course of TRIC infection in the eye. Mordhorst (1967) also assigned a numerical score to the proportion of inflammatory cells in conjunctival scrapings; we have not used this variable, but our own observations on large numbers of slides suggest that it may well be a very useful guide to the severity of infection, especially in the early stages. By analogy with the distinction between inflammatory and follicular lesions, it is likely that the proportion of polymorphonuclear leucocytes would be a more useful index than that of mononuclear cells, which tend to appear late.

One must decide not only what to measure but when to measure it. As a rule, scores are taken during the period when the infection is most severe. Dawson,

Mordhorst & Thygeson (1962) examined their monkeys 2 or 3 times a week, and summated the 3 highest scores in both eyes for each clinical sign recorded during the first 3 weeks after inoculation. Collier & Blyth (1966*a*) examined baboons 1, 2, 3, 4 and 6 weeks after inoculation. The cumulative score for the first 4 weeks only was used for interpreting the results, since by the 6th week spontaneous regression of physical signs had diminished the differences between control and vaccinated animals. Wang (1967) employed an average obtained by adding the scores at weekly examinations and dividing the sum by the number of weeks of observation, sometimes as many as 25. Our findings indicate that the optimum periods for taking scores may differ with the variable being measured. The addition of scores taken much outside these periods will obviously impair the sensitivity of discrimination between treated and control animals; and it follows that scores for the various signs of infection should not be combined unless their time courses are similar.

Strictly speaking, our findings and the conclusions from them apply only to baboons inoculated with large doses of strain MRC-4*s*; alteration of the experimental conditions, particularly by diminishing the dose of TRIC agent, might well result in a different pattern of infection. Nevertheless, other experiments not reported here suggest that the relationships between inflammatory lesions, follicles and inclusions are likely to be generally true of infections induced by other pathogenic TRIC agents, at least in baboons. We hope that this study has illustrated the importance of defining the course of infection in control animals; of being selective in choosing criteria for assaying the influence of vaccines and therapeutic drugs; and of using statistical techniques for these purposes.

We are highly indebted to Dr I. Sutherland for his advice on the methods of statistical analysis used in this paper.

REFERENCES

Assaad, F. A. & Maxwell-Lyons, F. (1967). Application of clinical scoring systems to trachoma research. *American Journal of Ophthalmology* **63**, 1327.

Collier, L. H. (1961). Experiments with trachoma vaccines: experimental system using inclusion blennorrhoea virus. *Lancet* i, 795.

Collier, L. H. (1967). The immunopathology of trachoma: some facts and fancies. *Archiv für die gesamte Virusforschung* **22**, 280.

Collier, L. H. & Blyth, W. A. (1966*a*). Immunogenicity of experimental trachoma vaccines in baboons. I. Experimental methods, and preliminary tests with vaccines prepared in chick embryos and in HeLa cells. *Journal of Hygiene* **64**, 513.

Collier, L. H. & Blyth, W. A. (1966*b*). Immunogenicity of experimental trachoma vaccines in baboons. II. Experiments with adjuvants, and tests of cross-protection. *Journal of Hygiene* **64**, 529.

Collier, L. H. & Smith, A. (1967). Dissemination and immunogenicity of live TRIC agent in baboons after parenteral injection. *American Journal of Ophthalmology* **63**, 1589.

Dawson, C., Jawetz, E., Thygeson, P. & Hanna, L. (1961). Trachoma viruses isolated in the United States. 4. Infectivity and immunogenicity for monkeys. *Proceedings of the Society for Experimental Biology and Medicine* **106**, 898.

Dawson, C. R., Mordhorst, C. H. & Thygeson, P. (1962). Infection of rhesus and cynomolgus monkeys with egg-grown viruses of trachoma and inclusion conjunctivitis. *Annals of the New York Academy of Sciences* **98**, 167.

GEAR, J. H. S., GORDON, F. B., JONES, B. R. & BELL, S. D. (1963). Nomenclature of isolates of virus from trachoma and inclusion blennorrhoea. *Nature, London* **197**, 26.

GRAYSTON, J. T., WANG, S. P., WOOLRIDGE, R. L., YANG, Y. F. & JOHNSTON, P. B. (1960). Trachoma: studies of etiology, laboratory diagnosis, and prevention. *Journal of the American Medical Association* **172**, 1577.

MORDHORST, C. H. (1967). Experimental infections and immunogenicity of TRIC agents in monkeys. *American Journal of Ophthalmology* **63**, 1603.

REEVE, P. & TAVERNE, J. (1963). Observations on the growth of trachoma and inclusion blennorrhoea viruses in embryonate eggs. *Journal of Hygiene* **61**, 67.

REEVE, P. & TAVERNE, J. (1967). The significance of strain differences in the behaviour of TRIC agents in the chick embryo. *American Journal of Ophthalmology* **63**, 1162.

WANG, S. P. (1967). Clinical evaluation of monkey infection with TRIC agents: a numerical scoring system of disease severity. *American Journal of Ophthalmology* **63**, 1321.

WORLD HEALTH ORGANIZATION (1962). Expert Committee on Trachoma. Third Report. *Technical Report Series, World Health Organization*, no. 234.

WORLD HEALTH ORGANIZATION (1966). Fourth Scientific Group on Trachoma Research. *Technical Report Series, World Health Organization*, no. 330.