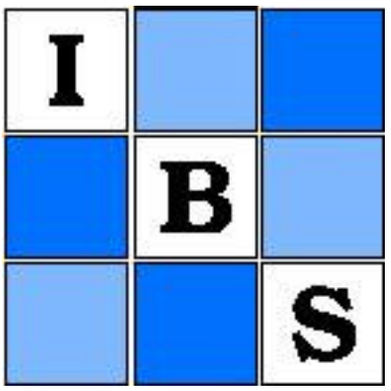

BRIEF OF PRESIDENTIAL ADDRESS:

THE 1954 TRIAL OF THE POLIOMYELITIS VACCINE IN THE UNITED STATES

WILLIAM G. COCHRAN

This trial represents an important application of biometrical principles in the struggle against disease. The experimental subjects were children in the first three grades or classes of school, of ages about 6–9 years. In terms of numbers of subjects the experiment may be the largest that has ever been conducted.

MAJOR DIFFICULTIES IN THE CONDUCT OF A TRIAL

(1) Poliomyelitis is a relatively rare disease. From past experience, the rate of paralytic polio in the study areas might be anticipated to be about 30 cases per 100,000 children aged 6–9 years. Given this attack rate, table I shows the probability of obtaining a statistically significant result (5% level) for various numbers of children and for various degrees of *true* effectiveness of the vaccine. With a vaccine that actually was 50% effective, about half a million children would be needed to make the risk of an inconclusive result small. Table II shows

TABLE I

Probability of obtaining a significant result (5% level)

No. of children in trial	True effectiveness of vaccine		
	50%	70%	90%
200,000	0.59	0.91	>0.99
400,000	0.88	>0.99	>0.99
600,000	0.97	>0.99	>0.99

TABLE II

Confidence limits for the true effectiveness

No. of children in trial	Observed effectiveness of vaccine		
	50%	70%	90%
200,000	2%—75%	34%—88%	68%—98%
400,000	20%—69%	48%—83%	77%—96%
600,000	27%—66%	53%—81%	80%—95%

the 95% confidence limits that would be obtained for the true effectiveness, if the observed effectiveness in the trial turned out to be 50%, or 70%, or 90%. Even with 600,000 children the true effectiveness can be none too well determined, except for a vaccine with an effectiveness up in the 90% range.

(2) The disease is difficult to diagnose: even in the paralytic form mistakes can be made. Some of the indefiniteness can be removed by adopting stringent criteria for the definition of a case. However, this device, if carried too far, may defeat its own ends by reducing the "accepted" cases to a very small number.

(3) The vaccination itself required 3 injections, the second given one week and the third 5 weeks after the first.

(4) The experiment subjects were children. Would parents give permission? Would physicians, health officers and medical societies give and encourage cooperation?

(5) Some biometricians have learned from bitter experience to take a pessimistic view of the prospects of success of *any* large trial with human subjects. Procedures that are essential for valid comparisons are apt to be cast aside as administratively impractical: instructions issued from a central office may be misread, misinterpreted or simply changed by persons a long way off; incomplete record forms and missing data flourish, and so on.

THE PLAN OF THE STUDY

The National Foundation for Infantile Paralysis invited the states individually to participate in the trial. If a state agreed, the vaccine was tested in all schools in certain counties within the state that had been selected by the Foundation. In order that the evaluation of the vaccine should be independent of the Foundation, the operation of the trial and the analysis of results were placed under the direction of Dr. Thomas Francis, with headquarters at the University of Michigan.

The plan announced by the Foundation was that the second-grade children in a participating school would receive the vaccine, while first and third grade children would remain unvaccinated to serve as controls.

This plan is subject to a number of potential biases. It requires the assumption that the attack rate among second-grade children is the same as the average attack rate amongst first and third grade children. Secondly, not all parents of second-grade children would allow their children to be vaccinated. Actually, 69% of them gave permission. Thus the plan compares a selected 69% of the second-grade children with the other two grades. There are epidemiological grounds for

arguing that this selection biases the results against the vaccine. Further, in any suspected case of the disease, it would be easy to discover whether a child had been vaccinated. This fact could create an unintentional bias in diagnosis by the local physician and could affect the completeness of reporting, as well as the precautions taken by parents for their children in the event of an epidemic.

It might be argued that the cumulative effect of these sources of biases was bound to be small and that results could not be seriously distorted if the vaccine was potent. But this assertion cannot be proved and with this method there must remain an element of doubt.

This plan was followed in 33 states, with 222,000 second-grade children vaccinated and 725,000 controls from the first and third grades.

A number of states adopted a different plan. Participating children in the 3 grades were divided at random into two groups. One group received three shots of the vaccine: the other received three shots of an inert fluid made up to have the same appearance as the shots of vaccine. The two treatments were distinguished by code numbers accessible only to those in charge of the study.

This plan raised more administrative difficulties than the first plan, but was free from the sources of biases that have been mentioned with respect to the first plan. All diagnoses, reporting and classification of cases, and all except the final stages of the analysis were done in ignorance of whether the child had received vaccine or placebo.

This plan was adopted in 11 states. Each treatment (vaccine or placebo) was represented by some 201,000 children. It is highly encouraging to biometricians that state officers and epidemiologists in these states expressed their preference for this plan, despite its many difficulties of execution.

Space permits mention of only a few aspects of the operation of the experiment. Collection of data was a formidable task, involving large numbers of letters, telegrams, telephone calls, regional and local conferences and special visits by members of the evaluation team to local areas. These efforts produced a high degree of completeness: missing data were of negligible importance.

Diagnoses were obtained in the following manner. When a suspected case appeared, a clinical history, including spinal fluid examination and blood and stool specimens, was made by the local physician on a standard form. A muscle examination was conducted by a physical therapist 10–20 days after onset, and a further examination 50–70 days after onset: each muscle report was reviewed by a local physician experienced in the clinical aspects of polio.

On the basis of these local records, a team of experts recruited by

the evaluation center at Michigan classified each case into one of the categories: (1) not polio (2) suspect (3) non-paralytic polio and (4) paralytic polio. The paralytic cases were further classified as to type and severity of paralysis. All these diagnoses were made by criteria that had been thrashed out and written down in advance by the team.

For record keeping and statistical analysis at the evaluation center itself, a small team of persons familiar with the handling and processing of large masses of data was obtained on leave of absence from the Bureau of the Census.

SOME RESULTS

Results were analysed and presented separately for the two plans. Areas covered by the original plan were called *observed* areas, while those that participated in the second plan were called *placebo* areas.

Table III shows the numbers of cases and the case rates per 100,000 children in the two areas. Incidentally, the paralytic case rates among non-vaccinated children were 43 in the placebo area and 44 in the

TABLE III

Cases and case rates per 100,000 children

Areas	No. of children in study	Polio cases			
		Paralytic		Non-paralytic	
		No.	Rate	No.	Rate
Placebo					
Vaccinated	200,745	33	16	24	12
Placebo	201,229	115	57	27	13
Not inoculated	338,778	121	36	36	11
Observed					
Vaccinated	221,998	38	17	18	8
Controls	725,173	330	46	61	8
2nd Grade not inoculated	123,605	43	35	11	9

observed area. Both rates were substantially above the anticipated rate of 30 which I used in discussing the needed sample size, so that the study had good fortune in not taking place during a year of unduly low incidence. The cases included in the results were all those that occurred between two weeks after the third injection and December 31, 1954.

In the placebo areas, paralytic case rates were 16 for vaccinated children and 57 for unvaccinated children. This gives an estimated effectiveness of 72%. In the observed areas the corresponding rates were 17 and 46, with an indicated effectiveness of 64%.

For non-paralytic cases, the rates were practically the same in vaccinated and control groups in both the placebo and observed areas. Although this result is somewhat unexpected, at least to a layman, it need not give concern from a public health point of view, since non-paralytic polio is not a major hazard like the paralytic form of the disease.

Table III also carries two lines marked "Not inoculated." In placebo areas this line refers to children in all three grades whose parents did not give permission to participate, plus a small number of children who received only one or two shots of placebo. In observed areas this group comprizes second-grade children whose parents did not request participation. In both areas the "not inoculated" group showed lower paralytic rates than the corresponding controls (36 against 57 and 35 against 46).

A difference in this direction had been anticipated on epidemiological grounds. Children of parents who withheld permission might be expected to be of a somewhat lower economic level than participating children, and to have acquired a greater degree of natural protection against polio through a previous subclinical attack of the disease. This type of selective bias has no effect on the results in the placebo areas, in which the comparison between vaccine and placebo was made entirely from participating children. In the observed areas, the bias would tend to reduce the apparent effectiveness of the vaccine. The fact that the vaccine showed lower effectiveness in the observed than in the placebo areas (64% against 72%) is in line with this explanation. A special sample survey that was made of participating and non-participating parents also tended to confirm the presence of a difference in economic level.

Table IV shows the estimated effectiveness of the vaccine as obtained from two more stringent criteria of classification. The main points to note are that the more severe criteria bring about some increase in the estimated effectiveness, and that the effectiveness figures run consistently about 10% lower in the observed than in the placebo areas.

The problem of making tests of significance and constructing confidence limits requires some consideration. One approach is to assume that the number of cases under a specific treatment in a school will follow a Poisson distribution. The total number of cases over all schools will then also follow a Poisson distribution, and the tests and

TABLE IV

Results given by more severe diagnostic criteria

Diagnosis	Placebo areas			95% limits for effectiveness
	No. of cases		Estimated effectiveness %	
	Vac.	Control		
Paralytic	33	115	72	57—81
Lab. confirmed	10	68	85	71—93
Positive virus obtained	15	70	80	62—89
	Observed areas			
Paralytic	38	330	62	47—74
Lab. confirmed	16	198	74	56—86
Positive virus obtained	20	210	69	50—82

limits can be constructed from Poisson theory. A more conservative approach, which avoids the Poisson assumption, is to regard the county as the basic sampling unit. The tests and limits are made by "continuous variable" theory, using the interaction with counties as the measure of error.

By either approach there is no doubt of the statistical significance of the beneficial effect of vaccine on paralytic cases. Confidence limits obtained by the Poisson approach appear in table IV, and serve to indicate the realm of uncertainty in our information as to the real effectiveness of the vaccine. The corresponding limits as obtained from the continuous variable approach would be somewhat wider.

Much credit is due to all who cooperated in this trial, and particularly to Dr. Francis and his staff, for the high standards maintained throughout the operation, despite the huge numbers of children to be processed. Among the many factors that contributed to give a fully valid comparison in the placebo areas, some of the most important were: (1) Randomization of children between vaccine and placebo (2) Keeping those concerned with case finding, diagnosis and classification in ignorance as to the treatment given to any child (3) Adoption of detailed criteria for the final diagnosis and classification and (4) Willingness to take endless pains to secure completeness and uniformity in reporting.

The question of the safety of the vaccine when given to such large numbers was of great concern. Special reports on all deaths of children,

from whatever cause, records of unusual reactions following shots, and studies of absenteeism from schools following shots were made. None of these indicated any basis for apprehension about the safety of the vaccine in this trial.

No discussion has been given here of a large volume of laboratory work designed to test the lots of vaccine, to study the rises in antibody levels following vaccination and to attempt to identify the virus from any case.

The Summary Report issued by the Vaccine Evaluation Center, University of Michigan, from which the data presented here were taken, should be consulted for a much more adequate account of the trial.