THE DESIGN AND SIGNIFICANCE OF SYNERGIC ACTION TESTS

By K. MATHER

John Innes Horticultural Institution, Merton, London

(With 6 Figures in the Text)

1. The data

In a recent paper, published in this *Journal*, Henderson & Gorer (1940) have described experiments designed to test the possibility of synergic action of certain forms of treatment of *Vibrion septique* infection in mice. Each test mouse was injected intradermally with a suspension of the spores, the dose being constant and known to be sufficient to cause death. Actually, out of more than three hundred control mice infected in this way, one did survive, but there is a definite suspicion that it had inadvertently escaped injection. The infected mice were then treated in various ways. Some were given sulphapyridine, others an antitoxin and still others an antibacterial serum. Each of these treatments, if given in suitable dose and at a suitable time, was able to save the lives of some mice. The precise rate of survival could be adjusted by varying the dosage and time of treatment. In addition to these single treatments, other mice were given combinations of the treatments, taken two at a time. Thus some mice were given sulphapyridine together with antitoxin, some sulphapyridine and antibacterial serum and finally a third group were injected with both antitoxin and antibacterial serum. The results of the experiments are shown in Table 1 (which is a copy of Henderson & Gorer's Table V). Six experiments were performed in all. Not all the treatments and combinations of treatments were given in each experiment, but the whole plan was designed to equalize the numbers of mice subjected to each treatment or pair of treatments.

There are certain fairly obvious features of this table. In the first place Exp. 2 seems to be giving subnormal survival rates in all the treatments it contains. This particular group of mice may have been genetically of higher susceptibility to the attack of *V. septique* than were the mice of the other experiments. Secondly sulphapyridine seems to show strong synergic interaction with both of the serological treatments, whereas the last two seem to be without interaction. We can, however, only be sure of the heterogeneity between the mice of the different experiments, and of the different interactions between the treatments, if a suitable statistical analysis confirms these apparent properties of the data. Furthermore, a statistical analysis should

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

Table 1

Treatment		Experiment 1 Surviving	1 Died	2 Surviving	2 Died	3 Surviving	3 Died	4 Surviving	4 Died	5 Surviving	5 Died	6 Surviving	6 Died	Total Surviving	Total Died	Total
Sulphapyridine (*S*)	Obs.	2	28	0	30	—	—	5	25	—	—	—	—	7	83	90
	Exp.	*2·3333*	*27·6667*	*2·3333*	*27·6667*	—	—	*2·3333*	*27·6667*	—	—	—	—	—	—	—
Antitoxin (*A*)	Obs.	2	27	—	—	3	27	—	—	2	28	—	—	7	82	89
	Exp.	*2·2809*	*26·7191*	—	—	*2·3596*	*27·6404*	—	—	*2·3596*	*27·6404*	—	—	—	—	—
Antibacterial serum (*B*)	Obs.	—	—	1	29	3	27	—	—	—	—	3	27	7	83	90
	Exp.	—	—	*2·3333*	*27·6667*	*2·3333*	*27·6667*	—	—	—	—	*2·3333*	*27·6667*	—	—	—
S + *A*	Obs.	26	4	—	—	—	—	28	1	25	5	—	—	79	10	89
	Exp.	*26·6292*	*3·3708*	—	—	—	—	*25·7416*	*3·2584*	*26·6292*	*3·3708*	—	—	—	—	—
S + *B*	Obs.	—	—	12	18	—	—	26	4	—	—	27	3	65	25	90
	Exp.	—	—	*21·6667*	*8·3333*	—	—	*21·6667*	*8·3333*	—	—	*21·6667*	*8·3333*	—	—	—
A + *B*	Obs.	—	—	—	—	6	24	—	—	2	28	6	24	14	76	90
	Exp.	—	—	—	—	*4·6667*	*25·3333*	—	—	*4·6667*	*25·3333*	*4·6667*	*25·3333*	—	—	—
Total	Obs.	30	59	13	77	12	78	59	30	29	61	36	54	179	359	538
	Exp.	*31·2434*	*57·7566*	*26·3333*	*63·6667*	*9·3596*	*80·6404*	*49·7416*	*39·2584*	*33·6555*	*56·3445*	*28·6667*	*61·3333*	—	—	—
Contribution to χ^2		**0·0763**		**9·5433**		**0·8314**		**3·9021**		**1·0287**		**2·7547**		**18·1365**		

Obs. = number observed. Exp. = number expected.

provide information as to the best way of designing such experiments, i.e. how to lay out the experiment in order to obtain the maximum of useful information per mouse used.

2. HETEROGENEITY OF THE EXPERIMENTS

Each of the six experiments comprised a different set of treatments, and in consequence the number of mice surviving in the experiments will be different, as the treatments and combinations of treatments appear to have very different effects on the survival rate. This means that a direct test of heterogeneity of the six experimental totals is useless, as they will merely reflect the treatment differences. The simplest way out of the difficulty would be to test each of the eighteen double entries in the body of the table separately, but this is unfortunately impossible as the expectations are often so small that χ^2 cannot be used. An adaptation of this method may, however, be used. The survival and death rates expected for each of the eighteen treatment tests are obtained. The expectations for the experiment totals are then found by summing the relevant values, and these may then be used as the basis of a χ^2 test.

Thus 90 mice in all were tested with sulphapyridine by itself. Of these 90, 7 survived. Then by simple proportion we expect $\frac{7\times30}{90}=2{\cdot}\dot{3}$ mice to have survived in each of the three groups of 30, of which the total is composed. These expectations are entered below the corresponding observed numbers in Table 1. Where antitoxin was used 7 survived out of 89, and so in Exp. 1, which included 29 of the 89, we should expect $\frac{7\times29}{89}$ or 2·2809 mice to survive. In Exps. 3 and 5, which each included 30 out of the 89, $\frac{7\times30}{89}$ or 2·3596 survivors would be expected. Where sulphapyridine and antitoxin were given together a total of 89 mice showed 79 survivors. These would be expected to be distributed in the proportions $\frac{79\times30}{89}$, $\frac{79\times29}{89}$ and $\frac{79\times30}{89}$ among Exps. 1, 4 and 5. Thus 26·6292, 25·7416 and 26·6292 would be expected to live, respectively. So the three treatments of Exp. 1 should give 2·$\dot{3}$, 2·2809 and 26·6292 survivors respectively, i.e. out of the 89 mice in this experiment 31·2434 are expected to live and 57·7566 to die. The expectations for the remaining five experimental totals are arrived at in the same way. Having obtained these expectations, a χ^2 for agreement of the observed values may be calculated, using the formula

$$\chi^2=S\,(a^2/m)-n,$$

where a is the number observed in any class, m the corresponding expectation, S indicates summation over both classes in each experiment, one of living and

one of dead mice, and n is the total number of mice tested. The $a^2/m - r$ values are shown in the bottom line of the table. Summing these values, we find

$$\chi^2 = 18{\cdot}1365.$$

This χ^2 has five degrees of freedom, as of the six, one from each experiment, one is taken up by the necessity of the experiments adding up to the grand total.

Now a χ^2 of 18·1365 for five degrees of freedom has a probability of less than 1%, so we may consider that heterogeneity of the experiments is significantly demonstrated.

A closer inspection of the data shows that the main contribution to χ^2 is made by Exp. 2. Now we can test the heterogeneity of the remaining five experiments (1 and 3–6) in the same way as we tested all six, except, of course, that the table of values will have Exp. 2 omitted. This is done in Table 2. The expectations are derived and χ^2 calculated just as before. We find $\chi^2 = 2{\cdot}3980$, for four degrees of freedom, and the probability lies between 70 and 50%. Thus there is no sign of heterogeneity when Exp. 2 is omitted. We must suppose that the mice used in this experiment were in some way, genetically perhaps, different in their reactions to *Vibrion septique* from those of the other five experiments. We must treat the results of Exp. 2 separately from those of the other five. If they were not separated the test of significance as developed below would be rendered invalid, since such heterogeneous data, when summed, give variances larger than those expected from the binomial distribution. Exp. 2 is, however, not to be wholly rejected. It can and does provide useful information when considered apart from the others. The remaining five may be treated together, and so in what follows below the values in the rightmost margin of Table 2 will be taken as true estimates of the effects of the six different treatments.

3. The test of significance for synergic action

If we consider the joint action of two treatments, such as sulphapyridine and antitoxin, we may set out their effects in a 2×2 table. Let sulphapyridine itself fail to save p_S of the mice to which it is given. There will then be $1 - p_S$ survivors. Similarly, antitoxin will save $1 - p_A$ mice and fail to save p_A of mice. Then clearly if both treatments are given together and provided that their action is independent, the first, sulphapyridine, will save $1 - p_S$ mice, and of the remaining p_S which would normally die, $1 - p_A$ would be saved by the antitoxin. Thus a proportion $p_S p_A$ would be expected to die, and $1 - p_S p_A$ to live (see Table 3).

Our problem is then to find out whether the survival rate after the joint application of two drugs conforms with this expectation. In order to derive the necessary test, consider the general case of three trials, in the first of which the mice received treatment S, in the second treatment A and in the

Table 2

	Experiment												
	1		3		4		5		6		Total		
Treatment	Surviving	Died	Surviving	Died	Surviving	Died	Surviving	Died	Surviving	Died	Surviving	Died	Total
S	2	28	—	—	5	25	—	—	—	—	7	53	60
A	2	27	3	27	—	—	2	28	—	—	7	82	89
B	—	—	3	27	—	—	—	—	3	27	6	54	60
S + *A*	26	4	—	—	28	1	25	5	—	—	79	10	89
S + *B*	—	—	—	—	26	4	—	—	27	3	53	7	60
A + *B*	—	—	6	24	—	—	2	28	6	24	14	76	90
Total Obs.	30	59	12	78	59	30	29	61	36	54	166	282	448
Exp.	*32·4101*	*56·5899*	*10·0263*	*79·9737*	*55·7416*	*33·2584*	*33·6555*	*56·3445*	*34·1667*	*55·8333*	—	—	—
Contribution to χ^2	**0·2818**		**0·4372**		**0·5097**		**1·0287**		**0·1406**		**2·3980**		

Table 3

Proportion of individuals surviving after treatment *S*	Proportion of individuals dying after treatment *S*	Total
Proportion of individuals surviving after treatment A		
Survive $(1-p_S)(1-p_A)$ by virtue of *A* and *S*	Survive $p_S(1-p_A)$ by virtue of *A*	$1-p_A$
Proportion of individuals dying after treatment A		
Survive $p_A(1-p_S)$ by virtue of *S*	Die $p_A p_S$	p_A
Total $1-p_S$	p_S	1

third both treatments combined, denoted by C. Let the numbers of mice tested by N_S, N_A and N_C respectively, the proportions surviving be q_S, q_A and q_C and the proportions dying p_S, p_A and p_C. Then

$$p_S+q_S=p_A+q_A=p_C+q_C=1.$$

The number of living and dead mice expected in the first trial will be $q_S N_S$ and $p_S N_S$ and so on. Let the numbers observed in these classes be Y_S, X_S, etc., as set out in Table 4.

Table 4

Treatment		S	A	$S+A$
Dying	Obs.	X_S	X_A	X_C
	Exp.	$N_S p_S$	$N_A p_A$	$N_C p_C\,[=N_C r p_S p_A]$
Surviving	Obs.	Y_S	Y_A	Y_C
	Exp.	$N_S q_S$	$N_A q_A$	$N_C q_C\,[=N_C(1-r p_S p_A)]$
	Total	N_S	N_A	N_C

Now let us suppose that the two treatments may interact when given jointly, i.e. p_C may not be equal to $p_S p_A$. We may measure their interaction by the use of a coefficient r, such that $p_C=r p_S p_A$. On the simple hypothesis of no interaction, which it is desired to test, this quantity r has the value 1. Then if we can find the deviation of r from 1 (i.e. $1-r$) and its expected variance (V_r), $\dfrac{(1-r)^2}{V_r}$ will be distributed as a χ^2 for one degree of freedom and will be a test of the hypothesis that the treatments do not interact. A significant value of χ^2 will indicate synergic action.

For this purpose it is necessary to estimate p_S, p_A and r. The log likelihood expression will be

$$L=X_S \log p_S+Y_S \log(1-p_S)+X_A \log p_A+Y_A \log(1-p_A) \\ +X_C \log r p_S p_A+Y_C \log(1-r p_S p_A).$$

Partial differentiation with respect to p_S, p_A and r gives

$$\frac{\partial L}{\partial p_S}=\frac{X_S}{p_S}-\frac{Y_S}{1-p_S}+\frac{X_C r p_A}{r p_S p_A}-\frac{Y_C r p_A}{1-r p_S p_A}=0,$$

$$\frac{\partial L}{\partial p_A}=\frac{X_A}{p_A}-\frac{Y_A}{1-p_A}+\frac{X_C r p_S}{r p_S p_A}-\frac{Y r p_S}{1-r p_S p_A}=0,$$

$$\frac{\partial L}{\partial r}=\frac{X_C p_A p_S}{r p_A p_S}-\frac{Y_C p_S p_A}{1-r p_S p_A}=0,$$

from which the maximum likelihood estimates are

$$p_S=\frac{X_S}{X_S+Y_S}=\frac{X_S}{N_S},$$

$$p_A=\frac{X_A}{X_A+Y_A}=\frac{X_A}{N_A},$$

$$r=\frac{X_C}{p_S p_A (X_C+Y_C)}=\frac{X_C N_S N_A}{N_C X_S X_A} \quad \left(=\frac{p_C}{p_A p_S}\right).$$

These results are not surprising as they give a perfect fit.

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

Then

$$1-r=1-\frac{p_C}{p_S p_A}=1-\frac{X_C N_S N_A}{N_C X_S X_A}=\frac{1}{N_C X_S X_A}(N_C X_S X_A-X_C N_S N_A).$$

In order to obtain the expected variance of r, (V_r) we note that

$$r=\frac{p_C}{p_S p_A}.$$

Then $\log r=\log p_C-\log p_S-\log p_A$.

Now in the two-class classification $Np : N(1-p)$

$$V_p=\frac{p(1-p)}{N},$$

and

$$V_{\log p}=V_p\left(\frac{d\log p}{d_p}\right)^2=\frac{1}{p^2}V_p.$$

Then

$$V_{\log p_S}=\frac{1}{p_S{}^2}V_{p_S}=\frac{1-p_S}{p_S N_S},$$

and similarly for the rest.

But on our hypothesis of no interaction $p_C=p_S p_A$. Hence

$$V_{\log p_C}=\frac{1-p_C}{p_C N_C}=\frac{1-p_S p_A}{p_A p_S N_C}.$$

Now these quantities p_C, p_S and p_A are estimated from independent sets of data and so

$$\begin{aligned}V_{\log r}&=V_{\log p_C}+V_{\log p_S}+V_{\log p_A}\\&=\frac{1-p_A p_S}{p_A p_S N_C}+\frac{1-p_S}{p_S N_S}+\frac{1-p_A}{p_A N_A}\\&=\frac{1}{p_A p_S}\left[\frac{1-p_A p_S}{N_C}+\frac{p_A(1-p_S)}{N_S}+\frac{p_S(1-p_A)}{N_A}\right].\end{aligned}$$

Now

$$V_r=V_{\log r}\left(\frac{dr}{d\log r}\right)^2=V_{\log r}r^2.$$

But r is 1 on our hypothesis, and so

$$\begin{aligned}V_r=V_{\log r}=&\frac{1}{p_S p_A N_C N_S N_A}\\&[N_S N_A(1-p_S p_A)+N_C N_A p_A(1-p_S)+N_C N_S p_S(1-p_A)].\end{aligned}$$

Substituting the observed values for p_S and p_A in this expression we get

$$V_r=\frac{N_S N_A}{N_S N_A N_C X_S X_A}\left[N_A N_S\left(1-\frac{X_A X_S}{N_A N_S}\right)+N_C N_A\frac{X_A Y_S}{N_A N_S}+N_C N_S\frac{X_S Y_A}{N_S N_A}\right]$$

$$=\frac{1}{N_C N_A N_S X_S X_A}[N_S N_A(N_S N_A-X_S X_A)+N_C N_A X_A Y_S+N_C N_S X_S Y_A].$$

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

Then

$$\chi^2=\frac{(1-r)^2}{V_r}$$
$$=\frac{(N_C X_S X_A - X_C N_S N_A)^2 X_S X_A N_C N_S N_A}{N_C{}^2 X_A{}^2 X_S{}^2 [N_S N_A (N_S N_A - X_S X_A) + N_C N_A X_A Y_S + N_C N_S X_S Y_A]}$$
$$=\frac{(N_C X_S X_A - X_C N_S N_A)^2 N_S N_A}{X_S X_A N_C [N_S N_A (N_S N_A - X_S X_A) + N_C N_A X_A Y_S + N_C N_S X_S Y_A]}.$$

This formula has been arrived at by a method dependent on estimation of r. It is, however, of general validity as it could have been obtained from any assumption as to how r should be measured. We might have defined r_1 as $p_C - p_S p_A$. The same formula for χ^2 would have been obtained by putting $\chi^2 = r_1{}^2/V_{r_1}$, the expected value of r_1 being 0. The formula is not dependent upon any assumption as to the type of interaction. It tests the hypothesis of no interaction.

The formula may be applied to the data of Table 2. The data concerning sulphapyridine and antitoxin are set out in Table 5 in a way which shows how this formula may be used.

Table 5

Treatment	*S*	*A*	*S* + *A*
Dying	53	82	10
Surviving	7	7	79
Total	60	89	89

We then find that

$$\chi^2=\frac{(89\times 53\times 82 - 10\times 89\times 60)^2\, 60\times 89}{53\times 82\times 89\,[60\times 89\,(60\times 89 - 53\times 82) + 89\times 89\times 82\times 7 + 89\times 60\times 53\times 7]}$$
$$=129{\cdot}623.$$

This χ^2 has one degree of freedom and so its probability is extremely small. We may consider that our hypothesis of no interaction is disproved. Synergic action of sulphapyridine and antitoxin is clearly shown.

In Table 6 the data for the interaction of sulphapyridine and antibacterial serum are shown.

Table 6

Treatment	*S*	*B*	*S* + *B*
Dying	53	54	7
Surviving	7	6	53
Total	60	60	60

$\chi^2 = 87{\cdot}181$ for 1 degree of freedom, P is very small; so there is clear evidence of synergic action of the two treatments.

Table 7 completes the series by giving the joint data on antitoxin and antibacterial serum.

Table 7

Treatment	*A*	*B*	*A* + *B*
Dying	82	54	76
Surviving	7	6	14
Total	89	60	90

$\chi^2 = 0{\cdot}0662$ for 1 degree of freedom, P is about 80%. Hence there is no evidence of interaction between the antitoxin and antibacterial serum.

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

4. PLANNING EXPERIMENTS FOR DETECTION OF SYNERGIC ACTION

Having derived the appropriate method of testing for synergic action, it is possible to reach certain conclusions as to how the experimental procedure may be designed so as to give the best chance of detecting such interaction.

Let us first consider the question of how many mice to give to each test. Every such experiment consists of three tests, one for each of the single treatments and one for their joint action. It seems not unreasonable to expect that, where a limited number of mice are available, more can profitably be assigned to the joint test than to either of the single ones. That this is in fact true can be shown by a consideration of the χ^2 formula. We can indeed go further and calculate the ratio of N_C to N_S and N_A, which will give the best results.

For this purpose, however, a more convenient form of the χ^2 expression is necessary. Instead of expressing the formula solely in terms of the observed quantities it may be written as

$$\chi^2=\frac{(1-r^2)\,p_S p_A}{\left[\frac{1-p_S p_A}{N_C}+\frac{p_A q_S}{N_S}+\frac{p_S q_A}{N_A}\right]}.$$

This form is easily derived from the intermediate stages of the working in the previous section, if q_S is written for $1-p_S$, etc.

Now the totals N_S, N_A and N_C occur only in the denominator of the fraction and so the problem of the most efficient distribution of mice among the three tests resolves itself into that of finding how N_S, N_A and N_C should be related in order to make the denominator a minimum.

Let us write

$$N_A=l_A N_S, \quad N_C=l_C N_S.$$

Then the ratio $N_C/N_S=l_C$ and $N_A/N_S=l_A$. The total number of mice used is $N_S(1+l_A+l_C)=N_T$. The denominator of the χ^2 fraction then becomes

$$\frac{1}{N_T}\left[\frac{(1-p_S p_A)(l_C+l_A+1)}{l_C}+\frac{p_A q_S (l_C+l_A+1)}{1}+\frac{p_S q_A (l_C+l_A+1)}{l_A}\right],$$

and this is to be minimized by adjustment of l_C and l_A. Partial differentiation gives

$$\frac{\partial}{\partial l_C}=\frac{1}{N_T}\left[\frac{1}{l_C^2}\{l_C(1-p_S p_A)-(1-p_S p_A)(l_C+l_A+1)\}+p_A q_S+\frac{p_S q_A}{l_A}\right]=0, \quad (1)$$

$$\frac{\partial}{\partial l_A}=\frac{1}{N_T}\left[\frac{1-p_S p_A}{l_C}+p_A q_S+\frac{1}{l_A^2}\{l_A p_S q_A-p_S q_A(l_C+l_A+1)\}\right]=0. \quad (2)$$

Then by subtraction, (1)−(2)

$$\frac{1}{N_T}\left[\frac{1-p_S p_A}{l_C^2}(l_C-l_C-l_A-1-l_C)-\frac{p_S q_A}{l_A^2}(l_A-l_C-l_A-1-l_A)\right]=0$$

or

$$\frac{1-p_S p_A}{l_C^2}=\frac{p_S q_A}{l_A^2}.$$

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

Substituting in equation (2)

$$\frac{1}{N_T}\left[\frac{1-p_S p_A}{l_C}+p_A q_S-\frac{1-p_S p_A}{l_C^2}\ (l_C+1)\right]=0,$$

i.e.
$$\frac{1}{N_T}\left[p_A q_S-\frac{1-p_S p_A}{l_C^2}\right]=0$$

and
$$\frac{1-p_S p_A}{l_C^2}=p_A q_S=\frac{p_S q_A}{l_A^2},$$

or
$$N_C=N_S\sqrt{\frac{1-p_S p_A}{p_A q_S}} \quad \text{and} \quad N_A=N_S\sqrt{\frac{p_S q_A}{p_A q_S}},$$

when $p_S=p_A$, l_A reduces to 1 and $N_A=N_S$. We may note that these optimum values are not dependent on any assumptions about the nature of r.

The effect of varying l_C round its optimum value of $\sqrt{\frac{1-p_S p_A}{p_A q_S}}$ is shown in Fig. 1. We may call $\frac{1}{V_r}$ the amount of information concerning r and write it I_r (cf. Fisher, 1938). Then, in the present case,

$$I_r=\frac{p_A p_S}{\left[\frac{1-p_A p_S}{N_C}+\frac{p_A q_S}{N_S}+\frac{p_S q_A}{N_A}\right]}$$

To make the representation simpler, put $p_A=p_S=p$ and let N_A take its optimum value of N_S. Then

$$I_r=\frac{p^2 N_T}{(2+l_C)\left[\frac{1-p^2}{l_C}+2p\,(1-p)\right]}$$

We may write
$$i_r=\frac{I_r}{N_T}=\frac{p^2}{(2+l_C)\left[\frac{1-p^2}{l_C}+2p\,(1-p)\right]}.$$

i_r may then be plotted against l_C when p is fixed. This relation is shown in Fig. 1 for $p=0{\cdot}9$ and $p=0{\cdot}95$. The choice of these two values of p for representation will become clear from what follows. It will be seen that the amount of information i_r is maximized at $l_C=1{\cdot}4530$ when $p=0{\cdot}90$ and at $l_C=1{\cdot}4340$ when $p=0{\cdot}95$, as our formula $l_C=\sqrt{\frac{1-p^2}{p\,(1-p)}}=\sqrt{\frac{1+p}{p}}$ would lead us to expect. Thus for a good experiment, with these values of p_A and p_S, N_C will be almost half as large again as N_S and N_A.

The next point in the planning of such experiments is that of the choice of the most suitable values of p_A and p_S. Since the doses of sulphapyridine, etc., may be varied, p_A and p_S can be controlled by the conditions of experiment. It is unfortunately not possible to give an explicit answer to this question. Putting N_C and N_A at their optimum values,

$$i_r=\frac{p_A p_S}{\left[\sqrt{\{(1-p_A p_S)\,p_A q_S\}}+p_A q_S+\sqrt{\{p_A p_S q_A q_S\}}\right]\left[1+\sqrt{\frac{1-p_S p_A}{p_A q_S}}+\sqrt{\frac{p_S q_A}{p_A q_S}}\right]}.$$

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

This expression becomes infinitely large when $p_S = 1$. Clearly, however, in such a case the experiment cannot be carried out as all the mice would die. We may, however, note that for i_r to be large p_S and p_A should be kept as large as is compatible with the experimental technique. The effect of doing

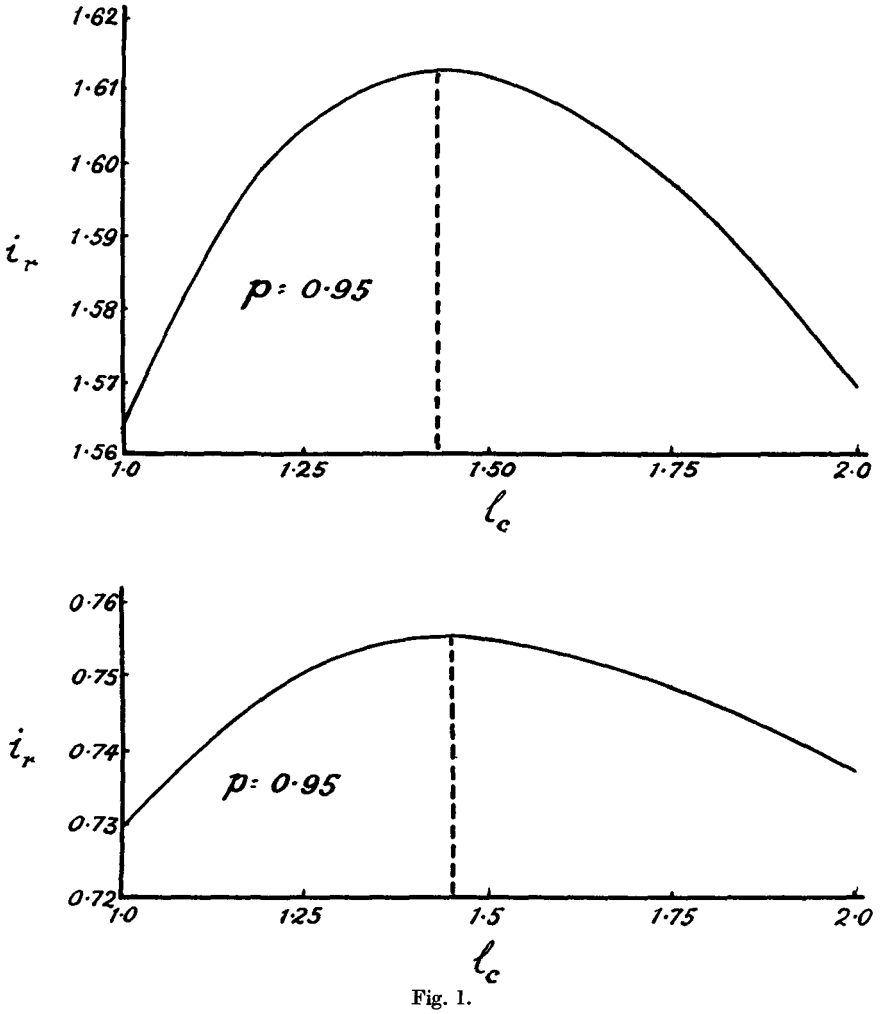

Fig. 1.

this is shown in Fig. 2 for which purpose p_A has been made equal to p_S. i_r then becomes

$$\frac{p^2}{(1-p)\,(2p+\sqrt{[p\,(1+p)]})\left(2+\sqrt{\frac{1+p}{p}}\right)},$$

and is plotted against p over the range $p=0$ to $p=1$. The value of keeping p_A and p_S high is well shown by the sharp rise in the amount of information i_r as $p_S\,(=p_A)$ approaches 1 (Fig. 2). Taking into account the necessity for having a sufficiently large number of mice in all the classes in the experiment,

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

so that a χ^2 test may legitimately be used, it would appear that p_S and p_A should be kept somewhere about 0·95 and in any case not allowed to drop below 0·90. Henderson and Gorer's experiments fulfil this condition very well, as the death rate after single treatments was about 93% in all cases.

It should be noted, however, that whereas both the test of significance of the previous section, and the optimum values of l_C and l_A can be arrived at without making any assumptions as to how the treatments interact, i.e. as

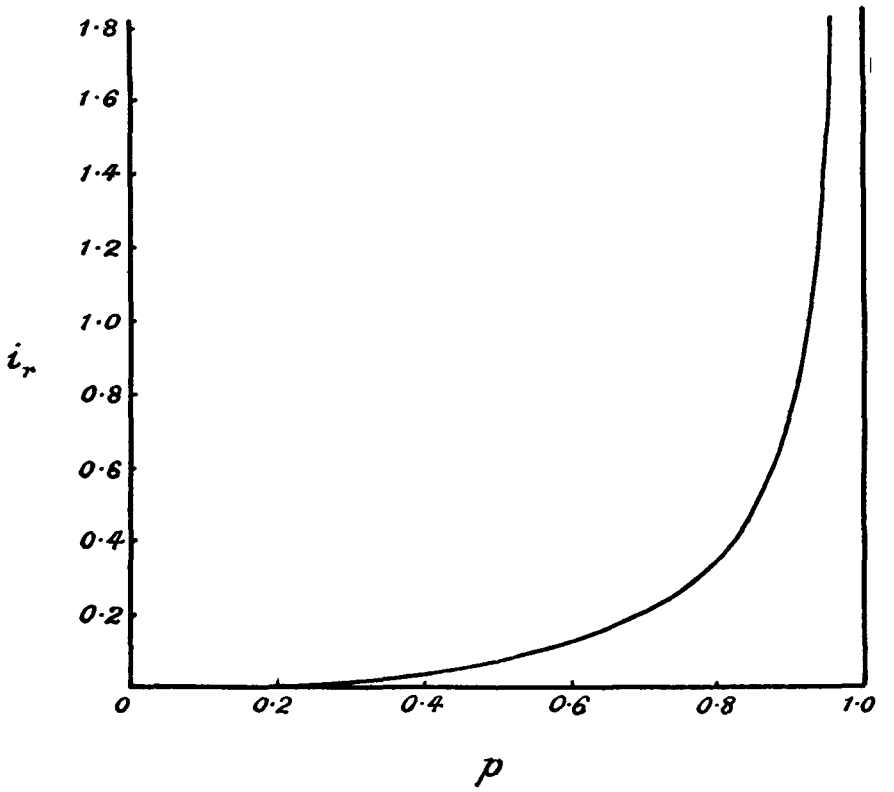

Fig. 2.

to the nature of r, the above determination of the desirable values of p_A and p_S is dependent on the assumption that $p_C = r p_A p_S$. If for example p_C is defined as $(p_S p_A - r)$, i.e. the interaction is of an intrinsically different kind, the optimum values of p_S and p_A would not be as above. They would in fact be low values below 0·1, rather than high values above 0·9. This is because with such an interaction i_r has a different formula, viz.

$$\frac{1}{p^2(1-p)(2p+\sqrt{[p(1+p)]})\left(2+\sqrt{\frac{1+p}{p}}\right)},$$

where $p_S = p_A$. The above calculation was given, however, as it seems much more likely that r has the nature of a coefficient rather than a constant dif-

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

ference. This point could be tested by measuring r over a range of values of p_S and p_A. If r is as assumed in the above calculations, $\frac{p_C}{p_S p_A}$ will be constant. If this is not found to be the case a new consideration of the p_S and p_A optima would become necessary.

Finally we may consider the question of how many mice should be used in experiments of this kind, in order to obtain a reasonably sure test of synergic action. Let $p_S = p_A = p$ for simplicity.

Now
$$\chi^2 = \frac{(p_C - p^2)^2 N_T}{p^2 (1-p) (2p + \sqrt{[p (1+p)]}) \left(2 + \sqrt{\frac{1+p}{p,}}\right)},$$

where l_A and l_C have taken their optimum values. Then the problem is to find out what value of N_T, the total number of mice, will detect with reasonable certainty a given difference between p_C and p^2. For a given value of p this is the same as asking what value of N_T will detect a given value of $r-1$ with reasonable certainty. The simple difference of p_C and p^2 is preferable as it is more easily determined for the data in a graphical test of significance, as shown below. In order to solve this problem we must determine the relation between $d\ (=p_C - p^2)$ and N_T. Let us first consider detection of deviations significant at the 5% level of probability. For one degree of freedom the 5% value of χ^2 is 3·841. Therefore the minimum value of d which can be found as significant at this level is given by

$$3{\cdot}841 = \frac{d^2 N_T}{p^2 (1-p) (2p + \sqrt{[p (1+p)]}) \left(2 + \sqrt{\frac{1+p}{p}}\right)}$$

or
$$d_{5\%} = \sqrt{\frac{3{\cdot}841\, p^2 (1-p) (2p + \sqrt{[p (1+p)]}) \left(2 + \sqrt{\frac{1+p}{p}}\right)}{N_T}}.$$

As the 1% χ^2 is 6·635 we get the corresponding expression

$$d_{1\%} = \sqrt{\frac{6{\cdot}635\, p^2 (1-p) (2p + \sqrt{[p (1+p)]}) \left(2 + \sqrt{\frac{1+p}{p}}\right)}{N_T}}.$$

The relation between d and N_T can then be calculated for any given value of p. Thus when $p = 0{\cdot}90$

$$d_{5\%} = \sqrt{\frac{3{\cdot}841 \times 0{\cdot}81 \times 0{\cdot}1 \times 3{\cdot}108 \times 3{\cdot}453}{N_T}} = \frac{1{\cdot}8272}{\sqrt{N_T}}$$

and
$$d_{1\%} = \sqrt{\frac{6{\cdot}635 \times 0{\cdot}81 \times 0{\cdot}1 \times 3{\cdot}108 \times 3{\cdot}453}{N_T}} = \frac{2{\cdot}3615}{\sqrt{N_T}}.$$

These relations are plotted in Figs. 3 and 4 for the three cases of $p = 0{\cdot}95$, 0·90 and 0·80. It will be seen that when p increases, smaller values of d become significant; as would be expected from some of the foregoing calculations.

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

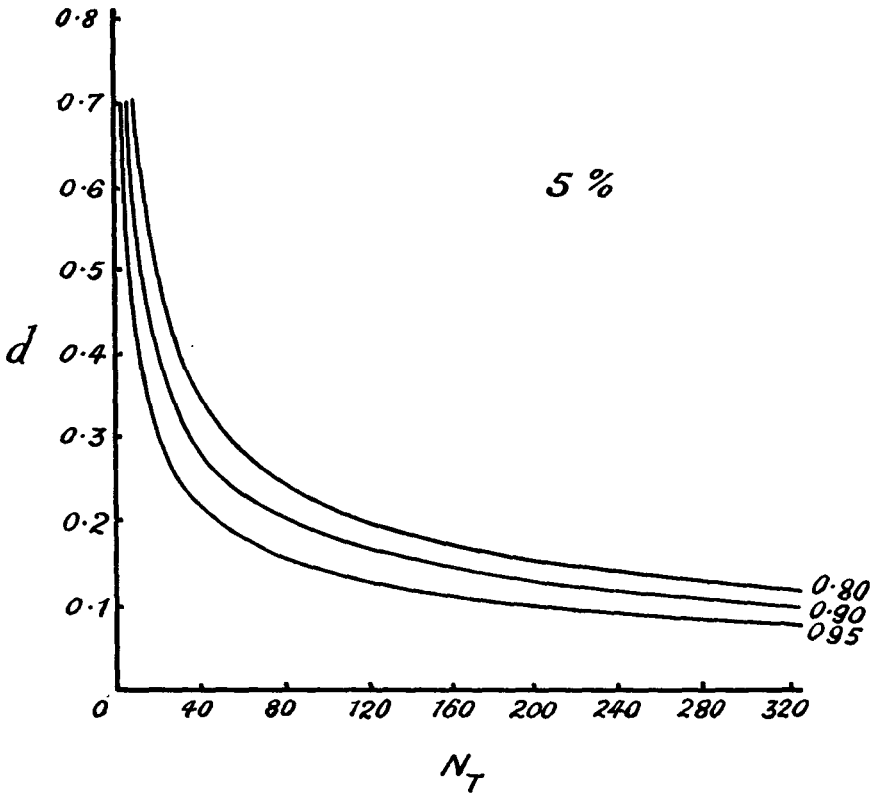

Fig. 3.

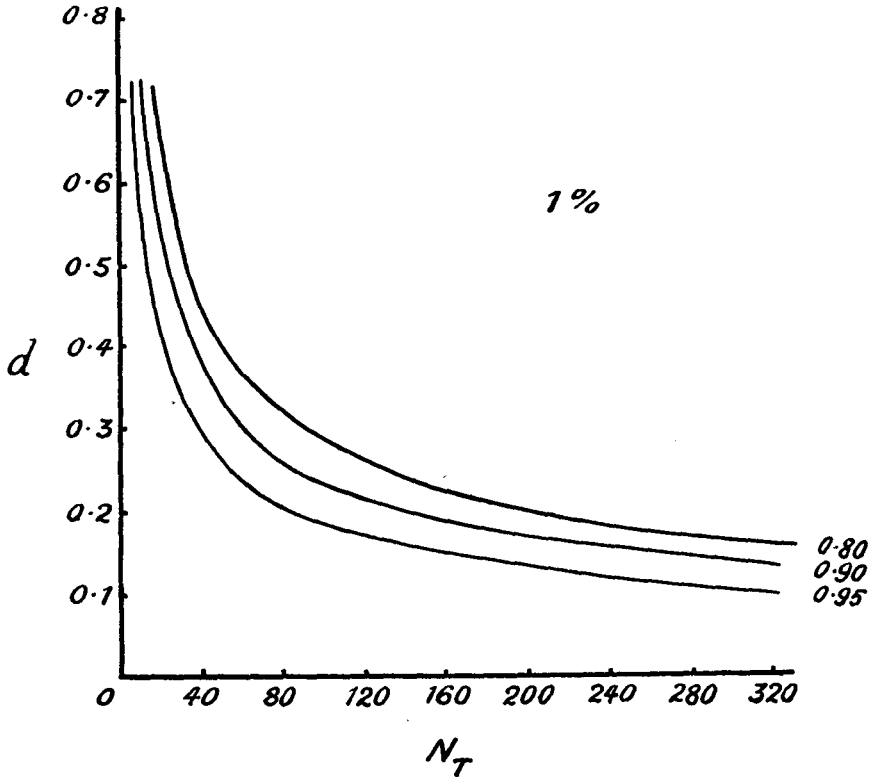

Fig. 4.

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

Such diagrams as Figs. 3 and 4 can be usefully employed in graphical tests of significance of actual experimental data. Thus, if an experiment is performed with $p_S = p_A = 0{\cdot}9$ and $N_S = N_A = \frac{N_C}{1{\cdot}5}$, the minimum deviations significant at the 5% and 1% levels will be shown sufficiently well by the $p = 0{\cdot}9$ lines of Figs. 3 and 4. Then if with a total N_T of 200 mice p_C is found to be 0·50, $d = 0{\cdot}81 - 0{\cdot}50 = 0{\cdot}31$ and this is highly significant as it lies above the $d_{1\%}$ line of Fig. 4. A d value of 0·2 with 60 mice is not significant as it lies below the 5% line of Fig. 3, but it is clear, that with twice as many mice this deviation would be significant.

Fig. 5 shows similar sets of curves for the cases of Henderson and Gorer's combination of sulphapyridine (S) with antibacterial serum (B) (Table 6). In this case, however, the 1% and 5% lines are shown on the same graph.

The relation of d to N_T for this experiment is calculated quite simply:

$$\chi^2 = \begin{Bmatrix} 3{\cdot}841 \\ 6{\cdot}635 \end{Bmatrix} = \frac{(p_C - p_A p_S)^2}{p_A p_S \left(\frac{1 - p_A p_S}{N_C} + \frac{p_A q_S}{N_S} + \frac{p_S q_A}{N_A} \right)}.$$

In the case of sulphapyridine and antibacterial serum (Table 6)

$$p_S = \tfrac{53}{60} = 0{\cdot}883 \quad \text{and} \quad p_B = \tfrac{54}{60} = 0{\cdot}90. \quad N_C = N_S = N_B.$$

(We may note that though p_S and p_A were fixed at useful values, the number of mice allocated to testing the joint treatment could have been increased with profit.)

$$\chi^2 = \begin{Bmatrix} 3{\cdot}841 \\ 6{\cdot}635 \end{Bmatrix} = \frac{(p_C - 0{\cdot}7950)\, N_T}{0{\cdot}7950\ (0{\cdot}2050 + 0{\cdot}1050 + 0{\cdot}0883) \times 3{\cdot}0}$$

and

$$d_{5\%} = \sqrt{\frac{3{\cdot}841 \times 0{\cdot}7950 \times 0{\cdot}3993 \times 3{\cdot}0}{N_T}} = \frac{1{\cdot}9126}{\sqrt{N_T}},$$

$$d_{1\%} = \sqrt{\frac{6{\cdot}635 \times 0{\cdot}7950 \times 0{\cdot}3993 \times 3}{N_T}} = \frac{2{\cdot}5137}{\sqrt{N_T}}.$$

The curves relating to d and N_T, as derived from these formulae, are given in Fig. 5. In actual practice when $N_T = 180$, p_C was found to be 0·111 and so

$$(p_B p_S - p_C) = (0{\cdot}7450 - 0{\cdot}1167) = 0{\cdot}6283,$$

which on reference to Fig. 5 is clearly very significant.

Turning to the case of antitoxin (A) and antibacterial serum (B) (Table 7)

$$p_A = \tfrac{82}{89} = 0{\cdot}92135, \; p_B = \tfrac{54}{60} = 0{\cdot}90, \; N_B = \tfrac{60}{89} N_A, \; N_C = \tfrac{90}{89} N_A.$$

Then

$$\chi^2 = \begin{Bmatrix} 3{\cdot}841 \\ 6{\cdot}635 \end{Bmatrix} = \frac{(p_C - 0{\cdot}82922)^2\, N_T}{239 \times 0{\cdot}82922\ (0{\cdot}001898 + 0{\cdot}000795 + 0{\cdot}001536)}$$

and so

$$d_{5\%} = p_C - 0{\cdot}8292 = \sqrt{\frac{917{\cdot}999 \times 0{\cdot}8292 \times 0{\cdot}004229}{N_T}} = \frac{1{\cdot}79419}{\sqrt{N_T}},$$

$$d_{1\%} = p_C - 0{\cdot}8292 = \sqrt{\frac{1585{\cdot}77 \times 0{\cdot}8292 \times 0{\cdot}004229}{N_T}} = \frac{2{\cdot}3581}{\sqrt{N_T}}.$$

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

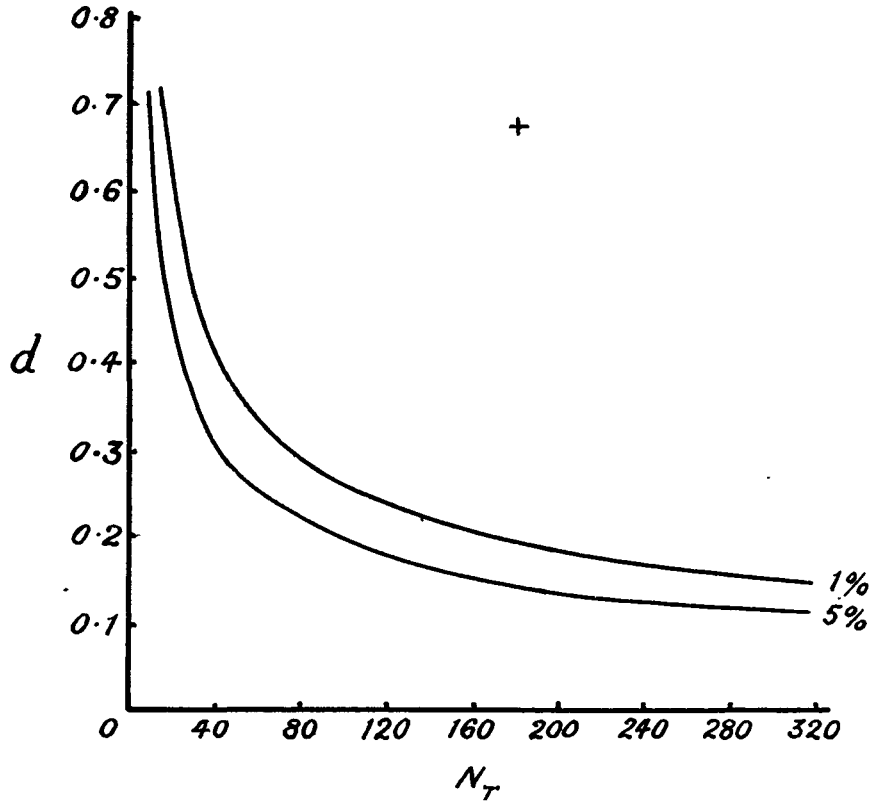

Fig. 5.

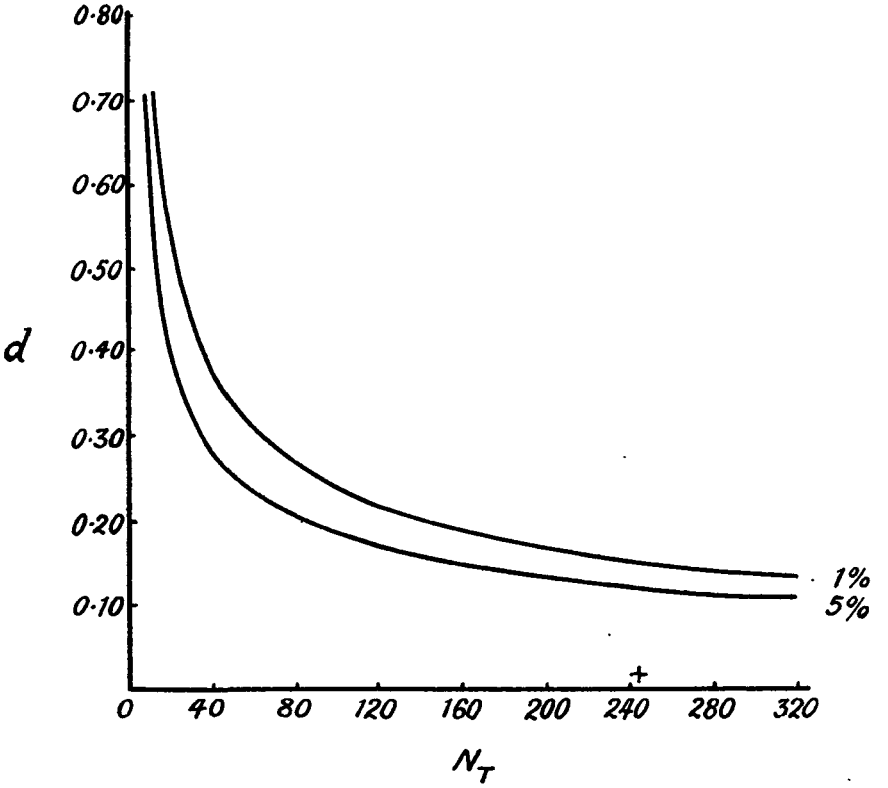

Fig. 6.

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

The curves are given in Fig. 6. The actual data gave $N_T = 239$, and

$$p_C = \tfrac{76}{90} = 0{\cdot}84444.$$

Then $d = 0{\cdot}0152$ and on entering on the diagram it is clearly not significant. Furthermore, a deviation as small as this would not become significant unless N_T were so large as to be unobtainable by ordinary experimental methods.

4. SUMMARY

The question of testing for synergic action of treatments of *Vibrion septique* infection in mice is considered. Henderson and Gorer's data are used as the basis of the analysis.

Their six experiments are tested for heterogeneity and one experiment is found to be discrepant. This is then removed and only the remaining five used in the later working.

A χ^2 formula is derived for testing the significance of the evidence for synergic action. Sulphapyridine shows such interaction in effect with both antitoxin and antibacterial serum. These last two show no interaction.

The most efficient design of such experiments is then considered, and it is shown that the death rate in the single treatment tests should be kept at least as high as 90%, and that whereas the numbers of mice assigned to the single treatment tests should be equal, that assigned to the joint treatment should be 1·5 times as great.

A graphical method for testing the significance of the results is described. These graphs also allow of an opinion being formed as to the number of mice which would be necessary to detect an interaction of given magnitude.

I am indebted to Prof. J. B. S. Haldane for suggesting the problem and to Drs Henderson and Gorer for giving me access to their data before publication.

APPENDIX

The following alternative approach to the problem of detecting synergic interaction has been developed by Prof. R. A. Fisher, at whose request I append its outline to the account of my own analysis. It differs from my method in the stage at which large sample theory is introduced.

In the absence of synergic interaction, the expected numbers of mice dying after treatment A is $p_A N_A$, after treatment B is $p_B N_B$, and after the joint treatment A and B is $p_A p_B N_C$, where the notation is as before. The observed frequencies are X_A, X_B, X_C. The joint log likelihood expression is then

$$L = X_A \log p_A + Y_A \log (1-p_A) + X_B \log p_S + Y_S \log (1-p_B) + X_C \log p_A p_B + Y_C \log (1-p_A p_B),$$

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

and the best estimates of p_A and p_B are obtained by solving the equations,

$$\frac{\partial L}{\partial p_A}=\frac{X_A}{p_A}-\frac{Y_A}{1-p_A}+\frac{X_C p_B}{p_A p_B}-\frac{Y_C p_B}{1-p_A p_B}=0,$$
$$\frac{\partial L}{\partial p_B}=\frac{X_B}{p_B}-\frac{Y_B}{1-p_B}+\frac{X_C p_A}{p_A p_B}-\frac{Y_C p_A}{1-p_A p_B}=0.$$

It can be shown from these equations that

$$p_A=\frac{X_A+\lambda}{N_A+\lambda},\quad p_B=\frac{X_B+\lambda}{N_B+\lambda}\quad\text{and}\quad p_A p_B=\frac{X_C-\lambda}{N_C-\lambda}$$

where λ is calculated from the equation

$$(N_A+\lambda)\,(N_B+\lambda)\,(X_C-\lambda)=(X_A+\lambda)\,(X_B+\lambda)\,(N_C-\lambda).$$

Having estimated p_A and p_B in this way their values may be used to formulate expectations for the numbers of mice surviving and dying in the three tests. Then the χ^2 testing the agreement of the observed numbers of mice in each class with the expected numbers has one degree of freedom and is the required test of synergic action.

The expected numbers are

Treatment	A	B	Joint
Dying	$\frac{N_A(X_A+\lambda)}{N_A+\lambda}$	$\frac{N_B(X_B+\lambda)}{N_B+\lambda}$	$\frac{N_C(X_C-\lambda)}{N_C-\lambda}$
Surviving	$\frac{N_A Y_A}{N_A+\lambda}$	$\frac{N_B Y_B}{N_B+\lambda}$	$\frac{N_C Y_C}{N_C-\lambda}$
Total	N_A	N_B	N_C

Then in experiment A the deviation of observation from expectation in each class is

$$\frac{N_A(X_A+\lambda)-X_A(N_A+\lambda)}{N_A+\lambda}=-\frac{Y_A\lambda}{N_A+\lambda}.$$

The contribution of this experiment to χ^2 will be

$$\frac{Y_A{}^2\lambda^2}{(N_A+\lambda)^2}\left[\frac{N_A+\lambda}{N_A(X_A+\lambda)}+\frac{N_A+\lambda}{N_A Y_A}\right],$$

which reduces to

$$\frac{Y_A\lambda^2}{N_A(X_A+\lambda)},$$

and so, summing over all three experiments

$$\chi^2=\lambda^2\left[\frac{Y_A}{N_A(X_A+\lambda)}+\frac{Y_B}{N_B(X_B+\lambda)}+\frac{Y_C}{N_C(X_C-\lambda)}\right].$$

Calculations of χ^2 in this way from Henderson and Gorer's data give for S and A 74·84, for S and B 30·56 and for A and B 0·069. These values should be compared with 129·65, 87·18 and 0·066 found by my method. The agreement between the two χ^2's is good, as it should be, when there is no evidence of synergic action; but when there is strong interaction of the treatments the

https://doi.org/10.1017/S0022172400028011 Published online by Cambridge University Press

two χ^2's give divergent results. This is to be expected, as in this case the original hypothesis is proven wrong and so all calculations which relate to one another through the hypothesis will become discrepant. The discrepancy has no serious implications. The two methods agree in showing the initial hypothesis, of no interaction, to be wrong. As they were developed purely to give tests of significance, they have thus both fulfilled their purpose.

Prof. Fisher's approach has a certain theoretical advantage over my own in that large sample theory is not so extensively involved. It is, however, a more laborious method to apply, as every determination of χ^2 involves the calculation of λ from a quadratic equation. In the case of experiments such as those of Henderson and Gorer, where the number of mice is reasonably large, the difference in treatment makes practically no difference to the results and so the easier arithmetic gives my method some practical advantage.

REFERENCES

FISHER, R. A. (1938). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

HENDERSON, D. W. & GORER, P. A. (1940). *J. Hyg., Camb.*, **40**, 345.

(*Received for publication* 25. v. 40.—Ed.)