Simulation methods in drug screening

O. L. DAVIES

Simulation is used to give a quantitative explanation of the phenomenon frequently observed in routine drug screening, that activities found in the initial tests are not confirmed on retest. Simulation is also used to assess the efficiency of a screening test procedure and to compare the efficiencies of different procedures.

SCREENING tests must be designed to make efficient use of the animals in a test, and to apply the appropriate amount of effort to test each compound so as to maximize the probability of finding a suitable drug. Much literature exists on the design of screening tests, and an extensive bibliography is given in Federer (1963).

A frequently occurring phenomenon is that activity found in an initial test is not confirmed on retest, an observation sometimes referred to as the "first experiment disease". This arises in the following way.

Compounds classified as positive in the primary screening test are usually retested experimentally to obtain more reliable estimates of their activities. But on retest the activities are generally appreciably lower than those observed initially. As a result the test or compound is often uspected of unreliability. Yet it can be readily shown statistically that on average a lower activity can be expected as a natural consequence of the usual testing error, and is not evidence of abnormality in the test. The reasons for this are that testing error alone can give rise to an apparent activity which may be greater than the actual activity, and when selecting compounds with highest apparent activities there is a tendency to select results which are high because of normal testing error. This can be pronounced when the proportion of positives is small.

There is no difficulty in demonstrating this phenomenon mathematically and to measure its magnitude. Another and convincing method of doing this is by simulation. This method does not require any specialized knowledge of mathematics and can be used by any experimenter preferably with some knowledge of computer programming, because simulations are most easily carried out on a computer.

PROCEDURE OF SIMULATION

First the compounds to be tested must be simulated. The compounds are assumed to be selected randomly from a large pool of possible compounds which could be tested. It is necessary to assign a distribution of activities to the pool of compounds. Much information is available on biological activities of compounds in general and there may be enough information available on the particular activity being investigated to allow at least a good guess of the form of the distribution. There must be some expectation of activity, otherwise the screen would not be

From the Department of Statistics, University College of Wales, Aberystwyth, Wales.

undertaken. Since the aim is to demonstrate an effect, a rough indication of the distribution is all that is needed.

In an actual screen which was simulated and the results of which are given later in this paper, the distribution of activities given in Table 1 was assumed. According to this distribution of activities 85% of the compounds are inactive and the remainder have activities ranging from 7·5 to 30% in the proportions shown.

TABLE 1. ASSUMED DISTRIBUTION OF ACTIVITIES IN THE POOL OF COMPOUNDS AVAILABLE FOR TEST

% Activity	0	7·5	15	22·5	30	
Proportion of compounds ..	0·85	0·08	0·04	0·02	0·01	
Cumulative proportions ..	0·85	0·93	0·97	0·99	1·00	

The procedure of selecting a compound randomly from the distribution is as follows. A random number between 0 and 1 is chosen from a table of random numbers (Tippett, 1959). The cumulative distribution is used to identify this number with a compound. If the number is less than or equal to 0·85, then it refers to a compound of zero activity; if the number is more than 0·85 and less than or equal to 0·93, it refers to a compound of activity 7·5% and so on. The complete Table is as shown in Table 2.

TABLE 2. CORRESPONDENCE BETWEEN RANDOM NUMBERS AND ACTIVITIES

Random number x	Activity of compound %
$0 < x \leqslant 0·85$	0
$0·85 < x \leqslant 0·93$	7·5
$0·93 < x \leqslant 0·97$	15
$0·97 < x \leqslant 0·99$	22·5
$x > 0·99$	30

The next step is to simulate the result of a test on the compound. The only information needed about the test is the distribution of the testing errors. If required, a numerical distribution of the errors can be drawn up in much the same way as that for the activities of the compounds, and an error selected randomly from the distribution again using a similar procedure. The distribution of errors will contain negative as well as positive errors. Usually it can be assumed that the testing error distribution is normal and then the standard error is all that needs to be known. Tables exist for random sampling from a normal distribution (Wold, 1948); these tables are based on a standard error of unity, and the selected random number must therefore be multiplied by the standard error of the test. The result represents the error to be added to the activity of the compound selected to give the simulated test results.

The above procedure has to be repeated many times and is best done by computer. Facilities exist in all modern computers to select a random number and also a random Normal deviate.

The steps in the full simulation procedure can now be set out.

(1)   Draw up a distribution of expected activities for the compounds (see Table 1).

(2)   Select compounds successively at random from this distribution and note the results.   This simulates the compounds which are presented for test.   Repeat a sufficient number of times, e.g. 300 or more times.

(3)   Simulate a test result by adding a random error, based on knowledge of the test, to the known activity of the compound selected. Repeat for successive compounds and note the results.

(4)   Identify those compounds which on first test are considered positive, e.g., all compounds giving a test result greater than a given amount decided beforehand.

(5)   Simulate a retest on these compounds by adding a random error to the known activities of the compounds.

(6)   Compare the two series of tests on the "positives".

Example.   Four hundred compounds were screened in a routine biological test and 40 of them were considered positive.   These 40 were retested.   The average degree of activity in the initial tests showed a reduction of 11% after retest.   The standard error of the test was known to be 8%.

RESULTS OF THE SIMULATION

The assumed distribution of activities was that given in Table 1.   The value of 8% was taken for the standard error of the test.

The results of the first simulation are given in Table 3.   Of the 400 compounds tested only those that passed the first screen, i.e., which gave simulated test results of 12% or more, are included, and those test results are given in the second column.   The first column represents the known activities of the compounds as selected randomly from the distribution of expected activities.   The third column gives the results of the retest, and the fourth, the averages of the first and the second tests.

The simulation was repeated 20 times giving the means of Table 4.

These tables show that the grand mean of the retested activities is only about one-half of the corresponding mean of the initial test on the "active" compounds, and indicate that the apparent drop in activity on retest in the routine screening test noted previously, is normal and to be expected.   It is in no way an indication either of loss in activities of compounds or of aberrations in the biological test.

EFFICIENCY OF SCREENING PROCEDURE

Further useful results can be obtained from the above simulations. For example they supply a direct and simple way of assessing the *efficiency* of the screening procedure.   A brief indication of how this can be done will now be given.   Screening divides the compounds tested into two lots, those that are considered to be of interest and those

TABLE 3.   RESULTS OF FIRST SIMULATION

% Activity (1)	Activity on first "test" ≥ 12% (2)	Repeat test (3)	Average test (4)
0·00	19·57	1·96	10·77
0·00	15·27	− 15·05	0·11
0·00	20·38	10·25	15·32
30·00	28·88	29·93	29·41
0·00	14·18	− 4·84	4·67
7·50	18·03	7·34	12·69
0·00	15·76	4·68	5·54
30·00	33·17	49·45	41·31
0·00	17·12	5·91	11·52
15·00	12·60	21·88	17·24
22·50	21·99	23·04	22·52
0·00	22·66	1·88	12·27
0·00	12·82	6·62	3·10
0·00	12·56	4·16	8·36
15·00	21·06	21·96	21·51
22·50	34·42	18·59	26·51
0·00	24·61	4·98	9·82
22·50	23·88	24·86	24·37
0·00	12·69	3·57	4·56
30·00	36·72	29·48	33·10
0·00	18·39	17·13	0·63
22·50	28·07	18·53	23·30
30·00	16·09	44·25	30·17
0·00	14·91	3·16	5·88
7·50	24·65	5·49	9·58
0·00	19·79	8·80	5·49
15·00	13·31	11·77	12·54
0·00	14·35	2·53	5·91
0·00	16·17	18·60	17·39
15·00	18·93	9·38	14·16
0·00	13·51	7·25	3·13
30·00	23·39	24·34	23·87
0·00	13·66	8·94	2·36
15·00	12·31	18·92	15·61
15·00	21·10	11·12	16·11
0·00	15·32	12·73	14·02
22·50	30·90	20·35	25·62
30·00	27·46	43·01	35·23
15·00	32·87	17·41	25·14
15·00	22·07	23·50	22·78
7·50	13·17	6·74	9·96
0·00	17·45	3·70	6·88
22·50	29·27	30·00	29·64
0·00	23·74	10·20	16·97
Mean 10·40	20·44	10·79	15·61

TABLE 4.   MEAN RESULTS FROM EACH OF TWENTY SIMULATION RUNS (each conducted on a group of 400 compounds)

No. of runs (1)	Mean % activity per run (2)	Mean activity on first test per run (3)	Mean repeat tests per run (4)	Mean of (3) and (4) (5)
1	10·40	20·44	10·79	15·61
2	11·03	20·96	11·82	16·39
3	10·37	19·30	10·50	14·90
4	9·91	18·57	9·31	13·94
5	9·96	19·39	10·85	15·12
6	8·61	18·44	7·23	12·84
7	9·48	18·34	10·85	14·59
8	9·84	18·25	10·34	14·30
9	10·42	17·90	9·06	13·48
10	7·90	17·23	9·77	13·50
11	8·14	16·20	7·15	11·67
12	9·68	18·40	8·62	13·51
13	8·11	19·43	7·87	13·65
14	7·62	18·74	6·81	12·78
15	9·48	17·66	9·43	13·55
16	9·88	18·80	9·30	14·05
17	11·48	19·25	10·42	14·83
18	8·90	17·36	6·48	11·92
19	12·09	21·00	12·64	16·82
20	13·06	20·08	13·44	16·76
Grand Mean	9·818	18·787	9·634	14·211

considered not to be of interest. The screening is based on a test which is subject to error, and, therefore, some misclassifications will arise. For the first screen the procedure was to regard compounds giving results $\geqslant 12\%$ activity as of interest and all others of no interest. A total of 8,000 compounds were "tested" in the above simulations, and of these 1,106 were considered active and 6,894 inactive.

By noting the actual activities of the randomly selected compounds in the two groups—those accepted and those rejected—the distributions of the activities of the compounds in the two groups can be derived. These are given in Table 5.

TABLE 5. DISTRIBUTIONS OF ACTIVITY

	% Activity of compounds					Total
	0	7·5	15·0	22·5	30·0	
Compounds which passed the test ..	474	165	199	187	81	1106
Compounds which failed the test ..	6332	429	115	17	1	6894
Total	6806	594	314	204	82	8000
Probability of accepting compound ..	0·070	0·278	0·634	0·917	0·988	
Probability of rejecting compound ..	0·930	0·722	0·366	0·083	0·012	

This Table shows that only 1 out of 82 (i.e. 1·2%) of the compounds with activity of 30% and 17 out of 204 (i.e. 8·3%) of the compounds with activity of 22·5% fail to be detected. The chance of missing active compounds of interest is, therefore, quite small. Similarly, the chance of passing compounds is negligible for those of zero activity and fairly small for compounds with activity of 7·5%. This scheme is suitable in the situation where compounds of 7·5% activity are of no interest, compounds of 15% activity of marginal interest and compounds of activities 22·5% or larger of definite interest.

A greater degree of discrimination arises after the second test. If on retest only those compounds with apparent activities $\geqslant 15\%$ are selected, the probabilities in Table 6 are obtained.

TABLE 6. DISTRIBUTIONS OF ACTIVITIES FOR OTHER ACCEPTANCE CRITERIA

	% Activity of compounds				
	0	7·5	15·0	22·5	30
Number accepted on second test	12	18	93	147	81
Probability of accepting	0·002	0·030	0·296	0·721	0·988
Probability of rejecting	0·988	0·970	0·704	0·279	0·012
(Numbers accepted on alternative procedure) (see text)	(2)	(11)	(86)	(147)	(76)

This second screen has removed most of the inactive compounds at a slight sacrifice of active ones.

Another possible criterion of selection is to calculate the means of the first and second test and to accept compounds with means $\geqslant 18\%$. This is a more stringent criterion which rejects far more of the inactive compounds at some further slight sacrifice of the active ones, as seen from the results in the last row of Table 6.

An initial inspection of the test performance in this way is often sufficient to select a suitable procedure. It is, sometimes, desirable to attempt a more quantitative comparison of the schemes. One way of doing this is to place a cost on passing compounds of low activities and a value on some active compounds which have been detected by the test. A cost might also be placed on those active compounds that have not been detected. These costs may be largely subjective and therefore the sensitivity of this procedure of selecting a suitable scheme to the assumption made will need to be known. This can be done by investigating a range of different costing schemes. A range of different initial distributions of activities might also need to be tried.

References

Federer, W. T. (1963). *Procedures and designs useful for screening material in selection and allocation with a bibliography.* Biometrics vol. 19.
Tippett, L. H. C. (1959). *Random Sampling Numbers.* (Tracts for Computers Series No. XV). Cambridge: University Press.
Wold, H. (1948). *Random Normal Deviates.* (Tracts for Computers Series No. XXV). Cambridge: University Press.