TESTS FOR THE STERILITY OF PHARMACEUTICAL PREPARATIONS*

THE DESIGN AND INTERPRETATION OF STERILITY TESTS

By D. Maxwell Bryce, B.Sc., B.Pharm., F.P.S., A.R.I.C.

From Fletcher Fletcher & Co. Ltd., London, N.7

Sterilisation was defined in the United States Pharmacopœia (XIIIth Revision) as "the destruction of all living organisms and their spores in, or their removal from, materials". The word is similarly defined elsewhere, for example, in the British Pharmaceutical Codex 1954 and the Japanese Pharmacopœia 1951. By inference sterility is the state of being free from living organisms and a sterile product is one entirely free from living organisms of all types. This concept is simple enough but unfortunately it is unreal, being incapable of experimental verification. A product is generally regarded as sterile because it has been subjected to a process believed to destroy or remove all micro-organisms and may therefore be expected to pass any sterility tests specified by national pharmacopœias or other authorities. Practical experience however has shown that neither exposure to a process of sterilisation nor passing sterility tests can give absolute certainty of sterility, in the sense of complete absence of living organisms. The most that can be claimed is a probability that the product is sterile although that probability may be very high as, for example, when a needle has been heated until it is red hot or saturated steam under pressure has been properly applied. The designation sterile is therefore to a certain extent arbitrary and official restrictions are generally placed on its use.

A number of pharmacopœias describe processes of sterilisation. Some of these are admitted to be uncertain because it is known that bacterial spores may survive boiling or Tyndallisation, for example. Other processes are considered completely effective: exposure to saturated steam at temperatures of 115° C. and above, heating in aqueous liquids containing 0·2 per cent. chlorocresol or 0·002 per cent. phenylmercuric nitrate at 98–100° C. and heating in a hot air oven at temperatures not less than 150° C. are usually regarded in this light if applied for a sufficient period of time. In fact, none of these can be so accepted. I have had in my possession an organism whose spores regularly survived autoclaving at 115° C. and more for 30 minutes, and Davies and Davison[1] and Davison[2] using heavy inocula of *Bacillus cereus* found that heating with either of the bactericides mentioned above failed to give sterility. The difficulty of controlling physical conditions in the usual type of hot air oven[3] makes dry heat sterilisation a notoriously uncertain procedure and Bowie[4,5] has criticised many of the pressure steam sterilizers at present in use, though not all his criticisms would receive general support[6].

* Papers read at the University (Scientific) Section of the London meeting of the Fédération Internationale Pharmaceutique on Thursday, September 22, 1955.

Admittedly there is no convincing evidence that pathogenic organisms have survived these official processes when they have been efficiently carried out and the survival of non-pathogenic sporing bacteria of abnormally high thermal resistance added in numbers far exceeding those met with in practice is not of great pharmaceutical significance. Nevertheless, the point must be accepted that exposure to one of these processes gives no certainty of sterility although it may give a high degree of probability of sterility.

As certainty cannot be obtained by the knowledge that an officially recognised process has been applied then can it be found by applying sterility tests to the products? The answer is of course an affirmative, but a limited affirmative. It is with these limitations that I hope to deal briefly. Although throughout the paper stress is laid on the limitations of tests for sterility, nevertheless it must be borne in mind that it is on such tests that the whole structure of knowledge of sterilising processes has been built up and when we consider that a process is sufficiently certain in its result to require no subsequent test we are in fact basing our stand on the results of large numbers of tests carried out theretofore.

A sterility test is an experiment carried out with the object of ascertaining certain facts about the flora and fauna of the system under examination. The potential yield of information is limited by the patent impossibility of testing for the whole wide range of possible organisms, and no test at all can be carried out without alteration or destruction of the system under examination. It is therefore impossible to say with certainty that the contents of each container are sterile, or even that they were sterile.

Clearly, what is known as a test for sterility is nothing of the sort, though we might relegate the operative word to inverted commas and designate the procedure as a test for "sterility" i.e., sterility within the meaning of the Act, regulation or pharmacopœia. What, in fact, we carry out is a test for certain contaminant organisms.

Bearing this in mind, let us now look more closely at such a test. As it is an attempt to infer the state of the whole from the result of an examination of the part, it is essentially a statistical operation. The organisms which most concern us are the pathogenic bacteria, though in passing we note that the viruses are a not unimportant group of pathogens. So, considering the general case, we take a sample of a sample, place it in a tube and provide those conditions we believe to be most suitable to bring about the vigorous reproduction of micro-organisms. Then within an arbitrary number of days we hope to get a clear cut result in terms of visible growth or no visible growth. On this evidence we must then decide whether the original material from which the first sample was drawn is to be accepted as sterile or condemmed. The evidence is indirect and may indeed be flimsy. We have to make a number of assumptions in assessing its significance and these should always be borne in mind. We assume that the growth arose from the sample tested and not from the culture medium. We check this point by examination of samples of the culture medium. We assume that the growth did not originate from a contaminant introduced during the

manipulation of the test—and Fleming found that mould spores make sport occasionally in the best conducted laboratories; and most pharmacopœias by permitted retesting concede the point. Thus the International (1st Ed.), British (1953), United States XV (1955), Swiss †Supp. 1954 and Belgian (1940) Pharmacopœias budget for the contingency. The French (1949) (General Directions) and the Japanese do not. With the contents of the remaining pharmacopœias I am not familiar, but no doubt the above sample is fairly representative.

Returning to the pharmaceutical aspects of the test for "sterility", we next assume that no visible growth within the arbitrary time period means no microbe in the sample drawn—always and inevitably a very moot point, for if insufficient attention is given to likes and dislikes no growth will certainly occur. We make a limited check of this aspect on a sample of the medium but the highly specialised requirements of many organisms are well known. Many sporing aerobes will not grow in blood media and a good many aerobes, I believe, are not at all happy in Brewer's medium—and how certain can we be that growth will occur in any given period of time. Indeed there is no agreement about the incubation period, times specified varying from one day in the French Pharmacopœia to ten days in the Belgian Pharmacopœia in the case of dressings. But then heat damaged spores have been known to take five months or longer to produce visible growth[7]. I believe the present record is about 18 months[8].

It will not have passed unnoticed that a great deal of sampling is involved in the testing procedure and sampling always spells uncertainty, greater uncertainty than is usually appreciated. Mathematics as well as technology therefore has something to say about the inferences we may draw from our evidence.

To arrive at the meaning of our result let us first consider the purely mathematical point of view and in the manner of mathematicians let us simplify the problem by ruling out all technical doubts. We shall then accept the simple equation that no growth = sterile because for practical purposes we must always accept that. And we shall similarly, for the time being accept the converse that growth = not sterile, in order to find out what inferences we may draw from a given result.

"One swallow does not make a summer" says our proverb, and yet everyone would agree that swallows are as good an indication of summer in this part of the northern hemisphere as any other natural phenomenon. How many swallows then do make a summer? To answer this type of question we have to invoke a probability function . . . as the number of swallows increases, the probability of the presence of the æstival season also increases.

In the same way our confidence in the satisfactory nature of a batch of parenteral solutions grows with each sample tested and found to be "sterile". But how does it grow and when can we feel reasonably sure of

† Note added in press—

Bemerkungen zur Prüfung auf Sterilität das Suppl. II der Ph. Helv. V by Metaxas, Linder and Munzel[22] which has come to hand since writing the above is a most useful commentary on the Swiss test in particular and Sterility Testing in general.

what we want to know? If we find a positive in our test series, the batch is labelled not sterile—but if we do not find such a positive, what may we conclude, for even if we test every container but one in a batch and find them satisfactory, we shall still not be sure of the condition of that final container.

A satisfactory conclusion must be that there is a reasonable likelihood that the batch is safe. The actual value of this likelihood of safety depends on the details of the test carried out which in turn depends on the rules of procedure followed. These vary greatly from country to country, different requirements frequently being laid down according to the nature and size of the container. Thus the Belgian Pharmacopœia requires a test in the case of distributed batches on 3 per cent. of vials up to a maximum of 10, the Japanese Pharmacopœia requires 3 containers to be examined if the lot size is less than 100 and then 1 additional container for every additional 50 or less containers in the batch up to a maximum of 10, the Swiss requirement is more detailed and specifies a maximum of 30 containers for batches in excess of 10,000 and the sample examined to be drawn at random. The U.S.P. on the other hand requires a representative sample of 10 units to be examined in the case of products sterilised by steam under pressure and for all other products a representative sample of 20 units is to be examined. It is clear that all these pharmacopœial directives are concerned solely with control of manufacture, a point to which I shall return later, in particular, the insistence of the U.S.P. on representative samples is noteworthy.

TABLE I

Per cent. infected items in batch:										
0·1	1	2	5	6·5	10	15	20	25	30	40	50
Probability of drawing 20 consecutive sterile items:										
0·98	0·82	0·67	0·36	0·26	0·12	0·039	0·012	0·003	0·0008	0·00004	0·000001

In Britain, the rules of procedure are laid down by the Therapeutic Substances Regulations[9]. These state: "The number of containers for test from every batch shall be 2 per cent. of the containers in the batch or 20 containers whichever is the less, taken at random from the batch, and if so required by the licensing authority, an additional 2 containers for each thousand or part of a thousand after the first". Since information about quality in a homogeneous batch is not related to the size of the batch but to the actual number of samples examined, a point elaborated by Knudsen[18], the efficiency of the test will rise with increasing batch size until the maximum 20 containers are drawn. It is evident that unless contamination of the batch is fairly widespread, it will not be unlikely that 20 successive sterile containers may nevertheless be drawn. In fact if p = the proportion of infected containers in the batch, assuming that an unbiassed sample has been drawn, the probability of obtaining 20 consecutive sterile containers is $(1 - p)^{20}$. A number of values are given in Table I.

These figures, which speak for themselves, show clearly the inability

of a test such as this to detect low levels of contamination. There is, in fact, a 1 in 4 chance that a batch containing 6·7 per cent. of infected containers will give 20 consecutive "steriles" and therefore be passed as "sterile".

When the number of samples drawn is smaller, the test is less stringent and we may consider 500 items as a typical smaller batch. In this case the probability of drawing 10 consecutive sterile items is approximately $(1 - p)^{10}$—this assumes that the probability p remains constant which is not, of course, strictly true. The values given in Table II are then applicable.

TABLE II

Per cent. infected items in batch:										
0·1	1	2	5	6·5	10	15	20	25	30	40	50
Probability of drawing 10 consecutive sterile items:											
0·99	0·90	0·82	0·60	0·51	0·35	0·20	0·11	0·056	0·028	0·006	0·001

So that there is now about a 50:50 chance of failing to detect contamination in a batch containing 6·7 per cent. of infected containers.

But the regulations recognise the possibility of adventitious infection entering during testing with the result that growth is not indicative of an infected container and make provision for the contingency as follows: "If at the examination a growth of micro-organisms is found in any tube, a further sample may be taken from the batch in the quantity specified . . . and the tests may be repeated on the further sample so taken. If on examination of the further sample no micro-organism is found, the sample shall be regarded as having passed the test; but if the same organism is found as was found in the first sample tested, the batch shall be treated as not sterile and the material contained in the batch shall not be issued or used as part of a further batch. If on such examination, a micro-organism is found, but the same micro-organism as was found in the sample first tested is not found, the test may be repeated on a third sample taken from the batch in the quantity aforesaid. If on examination of this sample no micro-organism is found the batch shall be regarded as having passed the test; but if any micro-organism is found the batch shall be treated as not sterile and the material contained in the batch shall not be issued or used as part of a further batch"*.

It is clear that the effect of these additional rules is to make the sterility

TABLE III

Per cent. infected items in batch:									
1	2	5	10	15	20	25	30	40	50
Proportion of such batches which would be passed as sterile:									
99·1	96·7	84	58	36	20	11	5·6	1·2	0·2

* Note added in press—

I am obliged to Mr. C. L. Sargent of the M.O.H. for pointing out that the 1953 amendment[23] adds the following words to the paragraph "unless or until the material has been resterilised and has passed the foregoing tests".

test even less stringent. The values in Table III are quoted from Davies and Fishburn[11] and are based on the assumption that only one infecting organism is present in a batch of 500 items. If more than one infecting organism is present the chance of passing a seriously infected batch is still further increased.

This excursion into the mathematics of induction has been based on simple sampling statistics applied to a rather simplified model of the actual test carried out. The figures apply, provided random samples have been drawn, fairly closely to small containers such as ampoules where the whole of the contents are examined. May we wonder in passing if a test which depends for its interpretation simply on visible growth, gains or loses by subdivision in order to test specifically for less common contaminants?

When we test larger containers an additional uncertainty enters the picture. At low contamination density it will be quite possible to withdraw a sterile sample from an infected container. This is even more true in the case of solids than in the case of liquids where distribution occurs more readily. How in fact should one test a large amount of solid for sterility? Presumably, and I have never been faced with the problem except in the special case of surgical dressings, the procedure would be dictated by the history of the solid e.g., the surface might first be examined, followed by examination of a representative sample obtained after thorough mixing or by core sampler. But whereas the sampling error in the case of the liquid is readily estimated, it is quite unpredictable in the case of the solid.

Sterility control, like any other form of quality control, is achieved not by the inspection operation but by getting at causes. Small samples considered in isolation tell very little about the bulk from which they were drawn, and the Therapeutic Substances Regulations or any other regulations based on the examination of small samples can detect only widespread contamination within a batch, and become increasingly less stringent with decreasing batch size, until with a batch size of 50, for example, we reach the position where it will be the exception and not the rule to throw out a batch containing 50 per cent. of infected containers. Thus many a batch of 500 containers which would fail to pass the test may prove quite acceptable as 10 sub-batches of 50. But if such a test is the best that can be done for the protection of the patient, is it fair to the manufacturer; in fact, is it technologically sound?

We have already noted that most official sampling and examination procedures seem to be designed for the guidance of the manufacturer. They are also frequently used as a basis for subsequent examination by buyers or other interested bodies.

Now when I ask is the test technologically sound, I am bound to recall the warning in the U.S.P. which will bear repetition "Sterility Tests are highly exacting and should be conducted by personnel having had expert training and experience in rigid aseptic technique". Further, the tests must be made under near ideal conditions, as we may be sure that indeed they are in the manufacturers' control laboratories, for while it is a

prime necessity to protect the patient, there is no virtue in throwing away or unnecessarily reprocessing good material due to misleading test results. Real evidence for the rejection of a batch requires the existence of data supplementary to the specified test, e.g., the existence of some form of control test designed to assess the testing conditions and expertise of the operator. Even so, the number of samples tested and the number of control tests performed is likely to be so small that any inference will be subject to a very high degree of uncertainty.

These problems assume an acute form in the testing of sterile surgical dressings and it is in this connection that I have been interested in the problem.

From the point of view of sterilisation, dressings differ in several major respects from other pharmaceutical items. Thus they are highly contaminated with microbes when they enter the steriliser- this is particularly true of cotton wool which after bleaching and washing is dried with hot air, rather than of gauze which is dried on a hot drum, though even gauze will contain on the average about 100 organisms per square inch—they are difficult to manipulate in testing and they are frequently presented for testing in a bacteriologically filthy wrapping. As a result even when the test is carried out under the best conditions of asepsis, appreciable contamination by airborne organisms occurs. For example, Pulvertaft[12], who identified *Cl. tetanii* and *Cl. welchii* in sanitary pads and accouchement sets, found that about 1 in 4 tests was contaminated during testing when working with dressings of undoubted sterility and Savage[13] found 16 positives in a series of 69 routine control tests. The probability of accidental infection during testing clearly depends on the testing conditions and the investigation by Savage stressed the importance of the very local conditions about the dressing in determining this. He concluded that the probability of adventitious contamination could be as high as 0·2. This state of affairs which is not widely appreciated leads to the serious situation when deliveries of dressings which are undoubtedly sound are rejected due to the lack of experience or inadequate technique of the testing bacteriologist. It is not unfair to say that some medical bacteriologists are prone to underestimate the difficulties of sterility testing.

If it be accepted as not abnormal for a substantial proportion of tests to be contaminated by aerial organisms, then it follows that the presence of low levels of contamination cannot be established. Savage was the first worker to clearly recognise this and he extended his argument to saying that since surgical dressings are massively contaminated before sterilisation, failure of the sterilising process may be expected to result generally in overall lack of sterility. (The same argument seems applicable to control of a filtration process in certain cases) i.e., the likelihood of a sterilisation process resulting in a load of wrapped dressings from which it will be possible to draw both sterile and non-sterile samples is small—this supposition has been borne out by many years of practical operation; only very rarely is it possible to obtain such a result as 9 infected samples out of 10. Therefore an inadequately sterilised load

of dressings will result in a complete run of contaminated samples and he stated his users' control test explicitly in the *Brit. med. J.*[14], in the following words: The test shall be made upon 10 dressings (or portions of a dressing) simultaneously and in random order with a control test upon 10 dressings similar to those under examination except that they are certainly known to be sterile as a result of laboratory treatment. Two conditions are essential: (1) not more than 4 of these control cultures may be positive; (2) the size of the test portion must be chosen so that (if the dressing is not sterile) the average number of organisms in each portion is at least 9. When these conditions are satisfied, dressings may be taken as infected when all the test cultures but not more than 4 of the controls, are positive and as sterile when at least one test culture is negative whatever the condition of the controls. If more than four of the controls are positive infection of the dressings cannot be inferred with enough certainty and the whole test is rejected without drawing any conclusions.

For further information about the test the original papers should be consulted and attention is particularly drawn to the difficulty of preparing a suitable set of control dressings.

Thus by showing that the occurrence of even a single negative in a test series is real evidence of sterility. Savage largely circumvented difficulties in the interpretation of results.

In an earlier paper[15] I attempted a mathematical justification of the test based on the assumption of a continuum of testing conditions characterised by a definite infection probability.

The control series of tests is performed in order to check the suitability of testing conditions, but the very limited amount of information available from a control run of only 10 tubes is not generally appreciated. If we regard these tubes as a random sample of overall testing conditions then the mathematical argument goes as follows[16]: if an event is observed to occur a times out of N, an upper limit p to the probability of this event may be assigned such that if the probability were actually p, then an observed number of occurrences as small or smaller than a would occur with a frequency P and corresponding to this probability is the limit of expectation of the number of occurrences in N trials, namely pN. In the above case $a = 4$ and the probabilities are calculated by solution of the equation:

$$P = \frac{10!\,p^4(1-p)^6}{4!\,6!} + \frac{10!\,p^3(1-p)^7}{3!\,7!} + \frac{10!\,p^2(1-p)^8}{2!\,8!} + \frac{10!\,p(1-p)^9}{9!} + (1-p)^{10}$$

See, for example, Davies[17].

The equation may be solved for different values of P and the values in Table IV are taken directly from Fisher and Yates[16] Table VIIIi: Binomial & Poisson Distributions: Limits of Expectation.

Therefore while in the case when 4 positives are observed in the control series the most likely value for the chance of adventitious contamination is 0·4, we can say only that on the average 9 times out of 10 the true values will be less than 0·65 (using the normal probability scale which extends

TABLE IV

No. of positives in 10 controls *a*	Upper limit of expectation when probability of a or fewer is		
	10 per cent.	2·5 per cent.	0·5 per cent.
5	7·33	8·13	8·72
4	6·46	7·38	8·09
3	5·52	6·52	7·35
2	4·50	5·56	6·48

TABLE V

Probability of chance infection	Probability of rejection of sterile dressings
0·8	0·07 per cent.
0·7	0·13 ,,
0·4	0·007 ,,
0·2	0·00001 ,,

from 0 to 1, in which 1 represents certainty) 39 times out of 40 it will be less than 0·74, but once in 200 tests it will exceed 0·81.

Assuming that the probability of chance infection is the same for a test dressing as for a control dressing and that it remains constant, we may calculate the probability of rejecting sterile dressings. Suppose that we carry out a test in such circumstances that the chance of aerial contamination remains constant at 0·7; there are 3 possible results:

(1) <10 positives in test .. .. Pass
(2) 10 positives in test $<$ 5 in control series.. .. Fail
(3) 10 positives in test $>$ 4 in control series.. .. Repeat test

The probability of each of these categories roughly computed is:

Pass	0·972
Fail	0·0013
Repeat	0·027

Further, 97·2 per cent. of the repeat tests would comply with the requirements for proof of sterility. There is then little error in calculating the probability of rejecting sterile dressings as the product of the probability of the result in the test series and the probability of the result in the control series—see Table V.

It is evident that it varies considerably with the probability of infection with a maximum value of about 0·13 per cent. However, in general, the probability of infection is not known and so the probability of wrong rejection can only be expressed as a function of this unknown value which we may call p.

$$\text{Probability of rejection} = \frac{10!\,10!}{10!\,4!\,6!}\,p^{14}(1-p)^6$$

Treating the extreme case (4 control and 10 test cultures positive) by the method of Fisher based on the multinomial expansion, we may calculate the probability of wrong rejection. This is found to be 0·58 per cent. Therefore approximately 1 out of every 200 batches of dressings rejected by this criterion will be wrongly rejected. The corresponding value for the case of 3 positives in the control run is 0·16 per cent. and the value falls to 0·00054 per cent. when there are no positives, i.e., even when 10 positives are found in the test series and none in the control series there is still a definite chance that a sterile batch of dressings may be unjustly condemned. The chance is small—but does not fall into the

same category as the well-known risk alleged by students of thermodynamics that a kettle of water may boil when you put it on a lump of ice. Clearly the manufacturer can regard the test with equanimity.

Considerations such as the above led to the introduction of a modified test for sterility in surgical dressings in the British Pharmaceutical Codex, 1954. This conceded the necessity for a control series as a check on bacteriological conditions and aseptic technique, and required that no positives appear in the control series, and not more than 3 in the test series for acceptance as sterile. Apart from the technical difficulties of the test, the soundness of the numerical requirements may readily be justified in a number of ways. Thus we may look up Fisher and Yates, Table VIIIi; and read off that when no positive occurs in the control series of 10 tubes we may be 90 per cent. certain that the true value estimating contamination is less than 2·06, 97·5 per cent. certain that it is less than 3·09 and 99·5 per cent. certain that it is less than 4·11. This suggests that from the mathematical point of view not more than 4 positives rather than not more than 3 would be doing better justice to the manufacturer. However, the B.P.C. covers the point by stating that conclusions drawn from tests in which the numbers of positives are on or near these limits are subject to a chance of error and the test should be repeated using larger inocula. The non-mathematical may feel that the test can be justified by common sense along the following lines: if no positive appears in the control series of tubes, we may infer that contamination of a tube by chance is a fairly unlikely event (indeed the tabulated figures mentioned above are based on this argument—that a Poisson distribution is involved). That being so, two such events will not frequently occur together, three most infrequently and four is so unlikely to occur that we may infer that chance is not a sufficient explanation.

This type of problem is generally treated statistically by calculation of the function χ^2 which is applied to problems where it is necessary to determine if a given event has occurred with a frequency significantly different from expectation. Calculation of χ^2 for homogeneity implies acceptance of the results of both classifications as their own expectations so to speak[18] and testing them for independence of classification. Thus we have the well-known 2×2 contingency table:

$$\begin{array}{c|c} 0 & 10 \\ \hline 3 & 7 \end{array}$$ where the expected cell values would be $$\begin{array}{c|c} 1\frac{1}{2} & 8\frac{1}{2} \\ \hline 1\frac{1}{2} & 8\frac{1}{2} \end{array}$$

Unfortunately in the case under consideration the numbers are very small and unbalanced so that this approach is not possible even making Yate's correction for continuity. However, Fisher and Yates, Table VIII, suggests that the value of χ_c obtained in the case of the distribution: $$\begin{array}{c|c} 0 & 10 \\ \hline 4 & 6 \end{array}$$ which is the one which mainly concerns us, does not reach the 1 in 40 level of significance, in other words again, it is not, mathematically, a wholly satisfactory criterion. That is not to say that the mathematician asserts that a given result is indicative of sterility or otherwise. He can

merely comment on the adequacy of the evidence and the above finding is confirmed by making the exact calculation of the probability of heterogeneity by the method of Fisher[19] which gives the value:

$$p = \frac{4!\,16!\,10!\,10!}{20!}\left(\frac{1}{10!\,4!\,6!}\right) = 0{\cdot}043.$$

The precision of the test could be improved by taking a greater number of samples, but it is generally felt that the examination of 20 samples of surgical dressings is enough for anyone—the mathematical assumption of constant testing conditions might be invalidated in a larger series due to human nature alone, which also may provide some corrective to the strictness of the test, in that unless testing conditions can be made really first rate, there is going to be a great deal of repetitive testing. On the other hand, if the control series of dressings has not been properly prepared, particularly in the case of dressings whose wrappings have become dusty or dirty in store, the test will be heavily biassed and the basis of the above mathematics will be invalidated.

So we see that it is very difficult to achieve full control of surgical dressing sterilisation by means of orthodox sterility testing, though in the manufacturer's laboratory the yield of information is greater than has been suggested above. Thus the assembly of control tests plotted as a control chart together with plate counts give a useful guide to testing conditions, and the identification of organisms found gives further useful information about sources of contamination. In fact because of the invariable presence of sporing organisms in surgical dressings, I have suggested that under many circumstances pasteurisation of the medium immediately after inoculation would result in much more information per test. This procedure was carried out for some time as a check on other methods of control. And while on the subject of refinements of technique, I should like to draw attention to the B.P.C. instruction to test with larger inocula in the cases of doubt. In many instances this is sounder than the more usual instruction to draw and test a larger number of samples since it increases the probability of finding contaminants without appreciably affecting the adventitious contamination rate.

In practice, in the company with which I was associated, primary control of sterilisation was based on the examination of earth packets containing thermoresistant spores which were strategically sited about the load. This is the method of control required by the Belgian Pharmacopœia which suggests however that *B. subtilis* is a suitable organism. Berry has shown that cultures of that organism vary very widely in their resistance to heat and the same is probably true of most other pure cultures. So that the German test material, based on the work of Konrich[20], dried and sieved garden earth with a sufficient content of native spores to resist steam at 120° C. for 5 minutes, is much more satisfactory.

Not any earth will do. Samples from some parts of the factory grounds were useless, but samples from one particular spot have been in constant use now for 20 years and have found their way into a number of other institutions as test organisms.

The actual organisms which yields the resistant spores has been described by Savage[21] as a slender rod which grows slowly to form chains. Despite its good record, the thermal death-point is kept under constant observation. About ¼–1 g. of earth is used per test packet and this is subsequently incubated in 15 ml. of B.P. aerobic medium. Cases of doubtful growth are resolved by microscope since subculturing is frequently unsuccessful due to the necessity for a growth factor present in the earth.

There seems no reason why this method of control should not be more widely utilised since, if a large inoculum of this organism, which is so much more resistant than pathogens such as *Cl. tetani* is sited where steam penetration is likely to be poorest and where air layering or trapping is most likely to occur, is killed, there can be no doubt about the effectiveness of the process.

The primary division of sterility testing is between systematic control tests carried out by those responsible for the production of sterile products, and any other tests. Testing cannot be divorced from the technology of sterilising procedures without great loss of information, and subsequent examinations are to be regarded only as safeguards against the occurrence of gross contamination or complete process failure.

I gratefully acknowledge the co-operation in the preparation of this paper of Professor Berry of the University of London, Dr. Capper of the Pharmaceutical Society and Dr. Maxwell Savage of Messrs. S. Maw, Son & Sons Ltd., New Barnet.

References

1. Davies and Davison, *Quart. J. Pharm.*, 1947, **20**, 212.
2. Davison, *J. Pharm. Pharmacol.*, 1951, **3**, 734.
3. Darmady and Brock, *J. clin. Path.*, 1954, **7**, 290.
4. Bowie, *Brit. med. J.*, 1955, **1**, 350.
5. Bowie, *Pharm. J.*, 1955, **174**, 492.
6. Savage, *J. appl. Bact.*, 1954, **17**, 283.
7. Savage, Personal communication.
8. Esty and Meyer, *J. infect. Dis.*, 1922, **31**, 650.
9. Statutory Instrument, 1952, No. 1937. H.M.S.O., London.
10. Knudsen, *J. Amer. pharm. Ass., Sci. Ed.*, 1949, **38**, 332.
11. Davies and Fishburn, *Pharm. J.*, 1948, **160**, 184.
12. Pulvertaft, *Brit. med. J.*, 1937, **1**, 441.
13. Savage, *Quart. J. Pharm. Pharmacol.*, 1940, **13**, 237.
14. Savage, *Brit. med. J.*, 1942, **1**, 472.
15. Bryce, *Pharm. J.*, 1952, **169**, 318.
16. Fisher and Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd Ed., London, 1948.
17. Davies, *Statistical Methods in Research and Production*, 2nd Ed., London, 1949.
18. Mather, *Statistical Analysis in Biology*, 3rd Ed., London, 1949.
19. Fisher, *Statistical Methods for Research Workers*, 11th Ed., London, 1950.
20. Konrich, *Dtsch. Z. Chir.*, 1929, **221**, 28.
21. Savage, *Quart. J. Pharm. Pharmacol.*, 1941, **14**, 158.
22. Metaxas, Linder and Munzel, *Pharm. Act. Helvet.*, 1955, **30**, 137.
23. Statutory Instrument, 1953, No. 1172, H.M.S.O., London.