J. Hyg., Camb. (1971), **69**, 95
Printed in Great Britain

A note on the safety testing of vaccines

By P. ARMITAGE
London School of Hygiene and Tropical Medicine

(*Received* 18 *September* 1970)

SUMMARY

The use of the hypergeometric distribution in calculations of sample volumes for safety testing is criticized. Alternative interpretations of the concept of 'infective ml. doses' lead to slightly different solutions, the simplest of which, involving the exponential function, is recommended for general use.

In a recent paper Anderson, Capstick, Mowat & Leech (1970) consider the testing of batches of vaccine to detect small amounts of residual live virus. Given a batch of volume N ml., with 'an assumed number of infective ml. doses', m, the probability P of failing to detect this amount of infection in a test on n ml. is given by these authors (p. 167) as

$$P = \frac{(N-m)!(N-n)!}{N!(N-m-n)!}. \tag{1}$$

This equation enables n to be calculated in terms of N, m and P, and thus provides guidance about the amount of testing necessary if a certain level of contamination is to be detected with reasonably high probability.

The authors do not clearly define an 'infective ml. dose', and the purpose of this note is to point out that alternative interpretations lead to formulae different from (1), although the practical consequences of these differences are usually small.

The basis of (1) appears to be the following model. Suppose the whole volume of N ml. is subdivided into N boxes, each of 1 ml., precisely m of which are 'infective' and $N-m$ of which are 'non-infective'. A random selection is made of n of these N boxes. The test fails to detect the infection if the boxes can be classified as follows.

	In sample	Not in sample	Total
Infective	0	m	m
Non-infective	n	$N-m-n$	$N-m$
Total	n	$N-n$	N

The probability of such an outcome is the term in the hypergeometric distribution, (1).

This model seems inappropriate, because the volume of N ml. is not a finite population of 1 ml. boxes. Thus, the sample of n ml. can be drawn in an infinite number of ways; further, the infective particles are presumably distributed irregularly in space—perhaps in an approximately uniform manner—and it seems

https://doi.org/10.1017/S002217240002129X Published online by Cambridge University Press

artificial to choose one particular partition into N boxes labelled as 'infective' or 'non-infective'.

In any realistic model involving a random distribution of infective particles, the number of 'infective ml. doses' is most naturally thought of as a random variable. Anderson *et al.*, however, clearly regard m as a parameter of the model, not subject to random variation. We therefore consider some alternative models in which m is given various interpretations as a parameter. We assume throughout that infective particles follow a uniform random distribution in space.

Model 1. There are exactly m particles in the volume of N ml. A random selection of n ml. is made. The probability of a negative test result is

$$P = (1-n/N)^m. \tag{2}$$

(This model is used by Peto & Maidment (1969); see the footnote on p. 170 of Anderson *et al.* (1970).)

Model 2. The volume of N ml. under consideration has been effectively selected from a much larger volume in which the density of infective particles is m per N ml. The actual number of infective particles in the batch under consideration is then Poissonly distributed about m. The required probability is

$$P = e^{-mn/N}. \tag{3}$$

Model 1A. There are exactly m' particles in the volume of N ml. An 'infective dose' is defined as a dose containing at least one infective particle, and m/N is defined as the probability that a 1 ml. dose is infective. Then

$$\frac{m}{N} = 1-\left(1-\frac{1}{N}\right)^{m'},$$

whence

$$m' = \frac{\log\{1-(m/N)\}}{\log\{1-(1/N)\}} \tag{4}$$

and, by analogy with (2),

$$P = (1-n/N)^{m'}. \tag{5}$$

Model 2A. An 'infective dose' is defined as in 1A, and m/N is again the probability that a 1 ml. dose is infective, but, as in Model 2, the number of infective particles in the batch is Poissonly distributed with mean m'. Then,

$$\frac{m}{N} = 1-e^{-m'/N},$$

and, by analogy with (3),

$$P = e^{-m'n/N} = \left(1-\frac{m}{N}\right)^n. \tag{6}$$

Equations (1), (2), (3), (5) and (6) give different relationships between P, m, n and N. When m and n are very small in comparison with N (as seems to happen in practical applications), all the formulae tend to the exponential version (3), which has the additional merit of simplicity. Unless there are clear arguments in favour of one of the other models, it would seem sensible to base calculations on (3).

https://doi.org/10.1017/S002217240002129X Published online by Cambridge University Press

REFERENCES

ANDERSON, E. C., CAPSTICK, P. B., MOWAT, G. N. & LEECH, F. B. (1970). *In vitro* method for safety testing of foot-and-mouth disease vaccines. *Journal of Hygiene* **68**, 159.

PETO, S. & MAIDMENT, B. J. (1969). Tables of the upper limit to the estimate of the density of contaminating particles in a medium. *Journal of Hygiene* **67**, 533.

https://doi.org/10.1017/S002217240002129X Published online by Cambridge University Press