01/2005:50300

1. INTRODUCTION

This chapter provides guidance for the design of bioassays prescribed in the European Pharmacopoeia (Ph. Eur.) and for analysis of their results. It is intended for use by those whose primary training and responsibilities are not in statistics, but who have responsibility for analysis or interpretation of the results of these assays, often without the help and advice of a statistician. The methods of calculation described in this annex are not mandatory for the bioassays which themselves constitute a mandatory part of the Ph. Eur. Alternative methods may be used, provided that they are not less reliable than those described here. A wide range of computer software is available and may be useful depending on the facilities available to, and the expertise of, the analyst.

Professional advice should be obtained in situations where: a comprehensive treatment of design and analysis suitable for research or development of new products is required; the restrictions imposed on the assay design by this chapter are not satisfied, for example particular laboratory constraints may require customized assay designs, or equal numbers of equally spaced doses may not be suitable; analysis is required for extended non-linear dose-response curves, for example as may be encountered in immunoassays. An outline of extended dose-response curve analysis for one widely used model is nevertheless included in Section 3.4 and a simple example is given in Section 5.4.

1.1. GENERAL DESIGN AND PRECISION

Biological methods are described for the assay of certain substances and preparations whose potency cannot be adequately assured by chemical or physical analysis. The principle applied wherever possible throughout these assays is that of comparison with a standard preparation so as to determine how much of the substance to be examined produces the same biological effect as a given quantity, the *Unit*, of the standard preparation. It is an essential condition of such methods of biological assay that the tests on the standard preparation and on the substance to be examined be carried out at the same time and under identical conditions.

For certain assays (determination of virus titre for example) the potency of the test sample is not expressed relative to a standard. This type of assay is dealt with in Section 4.5.

Any estimate of potency derived from a biological assay is subject to random error due to the inherent variability of biological responses and calculations of error should be made, if possible, from the results of each assay, even when the official method of assay is used. Methods for the design of assays and the calculation of their errors are, therefore, described below. In every case, before a statistical method is adopted, a preliminary test is to be carried out with an appropriate number of assays, in order to ascertain the applicability of this method.

The confidence interval for the potency gives an indication of the precision with which the potency has been estimated in the assay. It is calculated with due regard to the experimental design and the sample size. The 95 per cent confidence interval is usually chosen in biological assays. Mathematical statistical methods are used to calculate these limits so as to warrant the statement that there is a 95 per cent probability that these limits include the true potency. Whether this

precision is acceptable to the European Pharmacopoeia depends on the requirements set in the monograph for the preparation concerned.

The terms "mean" and "standard deviation" are used here as defined in most current textbooks of biometry.

The terms "stated potency" or "labelled potency", "assigned potency", "assumed potency", "potency ratio" and "estimated potency" are used in this section to indicate the following concepts:

— "stated potency" or "labelled potency": in the case of a formulated product a nominal value assigned from knowledge of the potency of the bulk material; in the case of bulk material the potency estimated by the manufacturer;

— "assigned potency": the potency of the standard preparation;

— "assumed potency": the provisionally assigned potency of a preparation to be examined which forms the basis of calculating the doses that would be equipotent with the doses to be used of the standard preparation;

— "potency ratio" of an unknown preparation; the ratio of equipotent doses of the standard preparation and the unknown preparation under the conditions of the assay;

— "estimated potency": the potency calculated from assay data.

Section 9 (Glossary of symbols) is a tabulation of the more important uses of symbols throughout this annex. Where the text refers to a symbol not shown in this section or uses a symbol to denote a different concept, this is defined in that part of the text.

2. RANDOMISATION AND INDEPENDENCE OF INDIVIDUAL TREATMENTS

The allocation of the different treatments to different experimental units (animals, tubes, etc.) should be made by some strictly random process. Any other choice of experimental conditions that is not deliberately allowed for in the experimental design should also be made randomly. Examples are the choice of positions for cages in a laboratory and the order in which treatments are administered. In particular, a group of animals receiving the same dose of any preparation should not be treated together (at the same time or in the same position) unless there is strong evidence that the relevant source of variation (for example, between times, or between positions) is negligible. Random allocations may be obtained from computers by using the built-in randomisation function. The analyst must check whether a different series of numbers is produced every time the function is started.

The preparations allocated to each experimental unit should be as independent as possible. Within each experimental group, the dilutions allocated to each treatment are not normally divisions of the same dose, but should be prepared individually. Without this precaution, the variability inherent in the preparation will not be fully represented in the experimental error variance. The result will be an under-estimation of the residual error leading to:

1) an unjustified increase in the stringency of the test for the analysis of variance (see Sections 3.2.3 and 3.2.4),

2) an under-estimation of the true confidence limits for the test which, as shown in Section 3.2.5, are calculated from the estimate of s^2, the residual error mean square.

3. ASSAYS DEPENDING UPON QUANTITATIVE RESPONSES

3.1. STATISTICAL MODELS

3.1.1. GENERAL PRINCIPLES

The bioassays included in the Ph. Eur. have been conceived as "dilution assays", which means that the unknown preparation to be assayed is supposed to contain the same active principle as the standard preparation, but in a different ratio of active and inert components. In such a case the unknown preparation may in theory be derived from the standard preparation by dilution with inert components. To check whether any particular assay may be regarded as a dilution assay, it is necessary to compare the dose-response relationships of the standard and unknown preparations. If these dose-response relationships differ significantly, then the theoretical dilution assay model is not valid. Significant differences in the dose-response relationships for the standard and unknown preparations may suggest that one of the preparations contains, in addition to the active principle, other components which are not inert but which influence the measured responses.

To make the effect of dilution in the theoretical model apparent, it is useful to transform the dose-response relationship to a linear function on the widest possible range of doses. 2 statistical models are of interest as models for the bioassays prescribed: the parallel-line model and the slope-ratio model.

The application of either is dependent on the fulfilment of the following conditions:

1) the different treatments have been randomly assigned to the experimental units,

2) the responses to each treatment are normally distributed,

3) the standard deviations of the responses within each treatment group of both standard and unknown preparations do not differ significantly from one another.

When an assay is being developed for use, the analyst has to determine that the data collected from many assays meet these theoretical conditions.

— Condition 1 can be fulfilled by an efficient use of Section 2.

— Condition 2 is an assumption which in practice is almost always fulfilled. Minor deviations from this assumption will in general not introduce serious flaws in the analysis as long as several replicates per treatment are included. In case of doubt, a test for deviations from normality (e.g. the Shapiro-Wilk[1] test) may be performed.

— Condition 3 can be checked with a test for homogeneity of variances (e.g. Bartlett's[2] test, Cochran's[3] test). Inspection of graphical representations of the data can also be very instructive for this purpose (see examples in Section 5).

When conditions 2 and/or 3 are not met, a transformation of the responses may bring a better fulfilment of these conditions. Examples are $\ln y$, $\sqrt{y}$, y^2.

— Logarithmic transformation of the responses y to $\ln y$ can be useful when the homogeneity of variances is not satisfactory. It can also improve the normality if the distribution is skewed to the right.

— The transformation of y to $\sqrt{y}$ is useful when the observations follow a Poisson distribution i.e. when they are obtained by counting.

— The square transformation of y to y^2 can be useful if, for example, the dose is more likely to be proportional to the area of an inhibition zone rather than the measured diameter of that zone.

For some assays depending on quantitative responses, such as immunoassays or cell-based *in vitro* assays, a large number of doses is used. These doses give responses that completely span the possible response range and produce an extended non-linear dose-response curve. Such curves are typical for all bioassays, but for many assays the use of a large number of doses is not ethical (for example, *in vivo* assays) or practical, and the aims of the assay may be achieved with a limited number of doses. It is therefore customary to restrict doses to that part of the dose-response range which is linear under suitable transformation, so that the methods of Sections 3.2 or 3.3 apply. However, in some cases analysis of extended dose-response curves may be desirable. An outline of one model which may be used for such analysis is given in Section 3.4 and a simple example is shown in Section 5.4.

There is another category of assays in which the response cannot be measured in each experimental unit, but in which only the fraction of units responding to each treatment can be counted. This category is dealt with in Section 4.

3.1.2. ROUTINE ASSAYS

When an assay is in routine use, it is seldom possible to check systematically for conditions 1 to 3, because the limited number of observations per assay is likely to influence the sensitivity of the statistical tests. Fortunately, statisticians have shown that, in symmetrical balanced assays, small deviations from homogeneity of variance and normality do not seriously affect the assay results. The applicability of the statistical model needs to be questioned only if a series of assays shows doubtful validity. It may then be necessary to perform a new series of preliminary investigations as discussed in Section 3.1.1.

2 other necessary conditions depend on the statistical model to be used:

— for the parallel-line model:

   4A) the relationship between the logarithm of the dose and the response can be represented by a straight line over the range of doses used,

   5A) for any unknown preparation in the assay the straight line is parallel to that for the standard.

— for the slope-ratio model:

   4B) the relationship between the dose and the response can be represented by a straight line for each preparation in the assay over the range of doses used,

   5B) for any unknown preparation in the assay the straight line intersects the y-axis (at zero dose) at the same point as the straight line of the standard preparation (i.e. the response functions of all preparations in the assay must have the same intercept as the response function of the standard).

Conditions 4A and 4B can be verified only in assays in which at least 3 dilutions of each preparation have been tested. The use of an assay with only 1 or 2 dilutions may be justified when experience has shown that linearity and parallelism or equal intercept are regularly fulfilled.

---

(1) Wilk, M.B. and Shapiro, S.S. (1968). The joint assessment of normality of several independent samples, *Technometrics* 10, 825-839.
(2) Bartlett, M.S. (1937). Properties of sufficiency and statistical tests, *Proc. Roy. Soc. London*, Series A 160, 280-282.
(3) Cochran, W.G. (1951). Testing a linear relation among variances, *Biometrics* 7, 17-32.

After having collected the results of an assay, and before calculating the relative potency of each test sample, an analysis of variance is performed, in order to check whether conditions 4A and 5A (or 4B and 5B) are fulfilled. For this, the total sum of squares is subdivided into a certain number of sum of squares corresponding to each condition which has to be fulfilled. The remaining sum of squares represents the residual experimental error to which the absence or existence of the relevant sources of variation can be compared by a series of *F*-ratios.

When validity is established, the potency of each unknown relative to the standard may be calculated and expressed as a potency ratio or converted to some unit relevant to the preparation under test e.g. an International Unit. Confidence limits may also be estimated from each set of assay data.

Assays based on the parallel-line model are discussed in Section 3.2 and those based on the slope-ratio model in Section 3.3.

If any of the 5 conditions (1, 2, 3, 4A, 5A or 1, 2, 3, 4B, 5B) are not fulfilled, the methods of calculation described here are invalid and an investigation of the assay technique should be made.

The analyst should not adopt another transformation unless it is shown that non-fulfilment of the requirements is not incidental but is due to a systematic change in the experimental conditions. In this case, testing as described in Section 3.1.1 should be repeated before a new transformation is adopted for the routine assays.

Excess numbers of invalid assays due to non-parallelism or non-linearity, in a routine assay carried out to compare similar materials, are likely to reflect assay designs with inadequate replication. This inadequacy commonly results from incomplete recognition of all sources of variability affecting the assay, which can result in underestimation of the residual error leading to large *F*-ratios.

It is not always feasible to take account of all possible sources of variation within one single assay (e.g. day-to-day variation). In such a case, the confidence intervals from repeated assays on the same sample may not satisfactorily overlap, and care should be exercised in the interpretation of the individual confidence intervals. In order to obtain a more reliable estimate of the confidence interval it may be necessary to perform several independent assays and to combine these into one single potency estimate and confidence interval (see Section 6).

For the purpose of quality control of routine assays it is recommended to keep record of the estimates of the slope of regression and of the estimate of the residual error in control charts.

— An exceptionally high residual error may indicate some technical problem. This should be investigated and, if it can be made evident that something went wrong during the assay procedure, the assay should be repeated. An unusually high residual error may also indicate the presence of an occasional outlying or aberrant observation. A response that is questionable because of failure to comply with the procedure during the course of an assay is rejected. If an aberrant value is discovered after the responses have been recorded, but can then be traced to assay irregularities, omission may be justified. The arbitrary rejection or retention of an apparently aberrant response can be a serious source of bias. In general, the rejection of observations solely because a test for outliers is significant, is discouraged.

— An exceptionally low residual error may once in a while occur and cause the *F*-ratios to exceed the critical values. In such a case it may be justified to replace the residual error estimated from the individual assay, by an average residual error based on historical data recorded in the control charts.

3.1.3. CALCULATIONS AND RESTRICTIONS

According to general principles of good design the following 3 restrictions are normally imposed on the assay design. They have advantages both for ease of computation and for precision.

a) Each preparation in the assay must be tested with the same number of dilutions.

b) In the parallel-line model, the ratio of adjacent doses must be constant for all treatments in the assay; in the slope-ratio model, the interval between adjacent doses must be constant for all treatments in the assay.

c) There must be an equal number of experimental units to each treatment.

If a design is used which meets these restrictions, the calculations are simple. The formulae are given in Sections 3.2 and 3.3. It is recommended to use software which has been developed for this special purpose. There are several programs in existence which can easily deal with all assay-designs described in the monographs. Not all programs may use the same formulae and algorithms, but they should all lead to the same results.

Assay designs not meeting the above mentioned restrictions may be both possible and correct, but the necessary formulae are too complicated to describe in this text. A brief description of methods for calculation is given in Section 7.1. These methods can also be used for the restricted designs, in which case they are equivalent with the simple formulae.

The formulae for the restricted designs given in this text may be used, for example, to create *ad hoc* programs in a spreadsheet. The examples in Section 5 can be used to clarify the statistics and to check whether such a program gives correct results.

3.2. THE PARALLEL-LINE MODEL

3.2.1. INTRODUCTION

The parallel-line model is illustrated in Figure 3.2.1.-I. The logarithm of the doses are represented on the horizontal axis with the lowest concentration on the left and the highest concentration on the right. The responses are indicated on the vertical axis. The individual responses to each treatment are indicated with black dots. The 2 lines are the calculated ln(dose)-response relationship for the standard and the unknown.

Note: the natural logarithm (ln or $\log_e$) is used throughout this text. Wherever the term "antilogarithm" is used, the quantity e^x is meant. However, the Briggs or "common" logarithm (log or $\log_{10}$) can equally well be used. In this case the corresponding antilogarithm is 10^x.

For a satisfactory assay the assumed potency of the test sample must be close to the true potency. On the basis of this assumed potency and the assigned potency of the standard, equipotent dilutions (if feasible) are prepared, i.e. corresponding doses of standard and unknown are expected to give the same response. If no information on the assumed potency is available, preliminary assays are carried out over a wide range of doses to determine the range where the curve is linear.
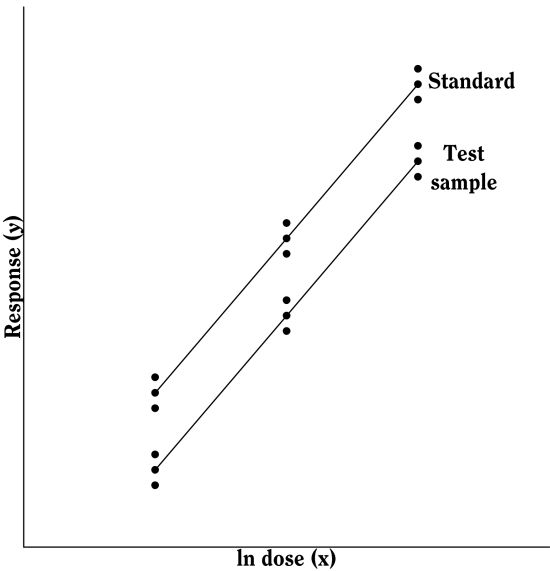
Figure 3.2.1.-I. – *The parallel-line model for a 3 + 3 assay*

The more nearly correct the assumed potency of the unknown, the closer the 2 lines will be together, for they should give equal responses at equal doses. The horizontal distance between the lines represents the "true" potency of the unknown, relative to its assumed potency. The greater the distance between the 2 lines, the poorer the assumed potency of the unknown. If the line of the unknown is situated to the right of the standard, the assumed potency was overestimated, and the calculations will indicate an estimated potency lower than the assumed potency. Similarly, if the line of the unknown is situated to the left of the standard, the assumed potency was underestimated, and the calculations will indicate an estimated potency higher than the assumed potency.

3.2.2. ASSAY DESIGN

The following considerations will be useful in optimising the precision of the assay design:

1) the ratio between the slope and the residual error should be as large as possible,

2) the range of doses should be as large as possible,

3) the lines should be as close together as possible, i.e. the assumed potency should be a good estimate of the true potency.

The allocation of experimental units (animals, tubes, etc.) to different treatments may be made in various ways.

3.2.2.1. Completely randomised design

If the totality of experimental units appears to be reasonably homogeneous with no indication that variability in response will be smaller within certain recognisable sub-groups, the allocation of the units to the different treatments should be made randomly.

If units in sub-groups such as physical positions or experimental days are likely to be more homogeneous than the totality of the units, the precision of the assay may be increased by introducing one or more restrictions into the design. A careful distribution of the units over these restrictions permits irrelevant sources of variation to be eliminated.

3.2.2.2. Randomised block design

In this design it is possible to segregate an identifiable source of variation, such as the sensitivity variation between litters of experimental animals or the variation between Petri dishes in a diffusion microbiological assay. The design requires that every treatment be applied an equal number of times in every block (litter or Petri dish) and is suitable only when the block is large enough to accommodate all treatments once. This is illustrated in Section 5.1.3. It is also possible to use a randomised design with repetitions. The treatments should be allocated randomly within each block. An algorithm to obtain random permutations is given in Section 8.5.

3.2.2.3. Latin square design

This design is appropriate when the response may be affected by two different sources of variation each of which can assume k different levels or positions. For example, in a plate assay of an antibiotic the treatments may be arranged in a $k \times k$ array on a large plate, each treatment occurring once in each row and each column. The design is suitable when the number of rows, the number of columns and the number of treatments are equal. Responses are recorded in a square format known as a Latin square. Variations due to differences in response among the k rows and among the k columns may be segregated, thus reducing the error. An example of a Latin square design is given in Section 5.1.2. An algorithm to obtain Latin squares is given in Section 8.6.

3.2.2.4. Cross-over design

This design is useful when the experiment can be sub-divided into blocks but it is possible to apply only 2 treatments to each block. For example, a block may be a single unit that can be tested on 2 occasions. The design is intended to increase precision by eliminating the effects of differences between units while balancing the effect of any difference between general levels of response at the 2 occasions. If 2 doses of a standard and of an unknown preparation are tested, this is known as a twin cross-over test.

The experiment is divided into 2 parts separated by a suitable time interval. Units are divided into 4 groups and each group receives 1 of the 4 treatments in the first part of the test. Units that received one preparation in the first part of the test receive the other preparation on the second occasion, and units receiving small doses in one part of the test receive large doses in the other. The arrangement of doses is shown in Table 3.2.2.-I. An example can be found in Section 5.1.5.

Table 3.2.2.-I. — *Arrangement of doses in cross-over design*

Group of units	Time I	Time II
1	S_1	T_2
2	S_2	T_1
3	T_1	S_2
4	T_2	S_1

3.2.3. ANALYSIS OF VARIANCE

This section gives formulae that are required to carry out the analysis of variance and will be more easily understood by reference to the worked examples in Section 5.1. Reference should also be made to the glossary of symbols (Section 9).

The formulae are appropriate for symmetrical assays where one or more preparations to be examined (T, U, etc.) are compared with a standard preparation (S). It is stressed that the formulae can only be used if the doses are equally spaced, if equal numbers of treatments per preparation are applied, and each treatment is applied an equal number of times. It should not be attempted to use the formulae in any other situation.

Apart from some adjustments to the error term, the basic analysis of data derived from an assay is the same for completely randomised, randomised block and Latin square designs. The formulae for cross-over tests do not entirely fit this scheme and these are incorporated into Example 5.1.5.

Having considered the points discussed in Section 3.1 and transformed the responses, if necessary, the values should be averaged over each treatment and each preparation, as shown in Table 3.2.3.-I. The linear contrasts, which relate to the slopes of the ln(dose)-response lines, should also be formed. 3 additional formulae, which are necessary for the construction of the analysis of variance, are shown in Table 3.2.3.-II.

The total variation in response caused by the different treatments is now partitioned as shown in Table 3.2.3.-III the sums of squares being derived from the values obtained in Tables 3.2.3.-I and 3.2.3.-II. The sum of squares due to non-linearity can only be calculated if at least 3 doses per preparation are included in the assay.

The residual error of the assay is obtained by subtracting the variations allowed for in the design from the total variation in response (Table 3.2.3.-IV). In this table $\overline{y}$ represents the mean of all responses recorded in the assay. It should be noted that for a Latin square the number of replicate responses (n) is equal to the number of rows, columns or treatments (dh).

The analysis of variance is now completed as follows. Each sum of squares is divided by the corresponding number of degrees of freedom to give mean squares. The mean square for each variable to be tested is now expressed as a ratio to the residual error (s^2) and the significance of these values (known as F-ratios) are assessed by use of Table 8.1 or a suitable sub-routine of a computer program.

3.2.4. TESTS OF VALIDITY

Assay results are said to be "statistically valid" if the outcome of the analysis of variance is as follows.

1) The linear regression term is significant, i.e. the calculated probability is less than 0.05. If this criterion is not met, it is not possible to calculate 95 per cent confidence limits.

2) The term for non-parallelism is not significant, i.e. the calculated probability is not less than 0.05. This indicates that condition 5A, Section 3.1, is satisfied;

3) The term for non-linearity is not significant, i.e. the calculated probability is not less than 0.05. This indicates that condition 4A, Section 3.1, is satisfied.

A significant deviation from parallelism in a multiple assay may be due to the inclusion in the assay-design of a preparation to be examined that gives an ln(dose)-response line with a slope different from those for the other preparations. Instead of declaring the whole assay invalid, it may then be decided to eliminate all data relating to that preparation and to restart the analysis from the beginning.

When statistical validity is established, potencies and confidence limits may be estimated by the methods described in the next section.

Table 3.2.3.-I. — *Formulae for parallel-line assays with d doses of each preparation*

	Standard (S)	1st Test sample (T)	2nd Test sample (U, etc.)
Mean response lowest dose	S_1	T_1	U_1
Mean response 2nd dose	S_2	T_2	U_2
...	...	...	...
Mean response highest dose	S_d	T_d	U_d
Total preparation	$P_S = S_1 + S_2 + ... + S_d$	$P_T = T_1 + T_2 + ... + T_d$	$P_U = ...\text{etc.}$
Linear contrast	$L_S = 1S_1 + 2S_2 + ... + dS_d + -\frac{1}{2}(d+1)P_S$	$L_T = 1T_1 + 2T_2 + ... + dT_d + -\frac{1}{2}(d+1)P_T$	$L_U = ...\text{etc.}$

Table 3.2.3.-II. — *Additional formulae for the construction of the analysis of variance*

$H_P = \dfrac{n}{d}$	$H_L = \dfrac{12n}{d^3 - d}$	$K = \dfrac{n(P_S + P_T + ...)^2}{hd}$

Table 3.2.3.-III. — *Formulae to calculate the sum of squares and degrees of freedom*

Source of variation	Degrees of freedom (f)	Sum of squares
Preparations	$h - 1$	$SS_{\text{prep}} = H_P(P_S^2 + P_T^2 + ...) - K$
Linear regression	1	$SS_{\text{reg}} = \frac{1}{h}H_L(L_S + L_T + ...)^2$
Non-parallelism	$h - 1$	$SS_{\text{par}} = H_L(L_S^2 + L_T^2 + ...) - SS_{\text{reg}}$
Non-linearity[*]	$h(d - 2)$	$SS_{\text{lin}} = SS_{\text{treat}} - SS_{\text{prep}} - SS_{\text{reg}} - SS_{\text{par}}$
Treatments	$hd - 1$	$SS_{\text{treat}} = n(S_1^2 + ... + S_d^2 + T_1^2 + ... + T_d^2 + ...) - K$
[*] Not calculated for two-dose assays		

Table 3.2.3.-IV. — *Estimation of the residual error*

Source of variation		Degrees of freedom	Sum of squares
Blocks (rows)[*]		$n - 1$	$SS_{\text{block}} = hd\,(R_1^2 + ... + R_n^2) - K$
Columns[**]		$n - 1$	$SS_{\text{col}} = hd\,(C_1^2 + ... + C_n^2) - K$
Residual error[***]	**Completely randomised**	$hd\,(n - 1)$	$SS_{\text{res}} = SS_{\text{tot}} - SS_{\text{treat}}$
	Randomised block	$(hd - 1)\,(n - 1)$	$SS_{\text{res}} = SS_{\text{tot}} - SS_{\text{treat}} - SS_{\text{block}}$
	Latin square	$(hd - 2)\,(n - 1)$	$SS_{\text{res}} = SS_{\text{tot}} - SS_{\text{treat}} - SS_{\text{block}} - SS_{\text{col}}$
Total		$nhd - 1$	$SS_{\text{tot}} = \sum (y - \overline{y})^2$

[*] Not calculated for completely randomised designs
[**] Only calculated for Latin square designs
[***] Depends on the type of design

3.2.5. ESTIMATION OF POTENCY AND CONFIDENCE LIMITS

If I is the ln of the ratio between adjacent doses of any preparation, the common slope (b) for assays with d doses of each preparation is obtained from:

$$b = \frac{H_L\,(L_S + L_T + ...)}{Inh} \qquad (3.2.5.\text{-}1)$$

and the logarithm of the potency ratio of a test preparation, for example T, is:

$$M_T' = \frac{P_T - P_S}{db} \qquad (3.2.5.\text{-}2)$$

The calculated potency is an estimate of the "true potency" of each unknown. Confidence limits may be calculated as the antilogarithms of:

$$CM_T' \pm \sqrt{(C - 1)\,(CM_T'^2 + 2V)} \qquad (3.2.5.\text{-}3)$$

where $\; C = \dfrac{SS_{\text{reg}}}{SS_{\text{reg}} - s^2 t^2} \;$ and $\; V = \dfrac{SS_{\text{reg}}}{b^2 dn}$

The value of t may be obtained from Table 8.2 for $p = 0.05$ and degrees of freedom equal to the number of the degrees of freedom of the residual error. The estimated potency (R_T) and associated confidence limits are obtained by multiplying the values obtained by A_T after antilogarithms have been taken. If the stock solutions are not exactly equipotent on the basis of assigned and assumed potencies, a correction factor is necessary (See Examples 5.1.2 and 5.1.3).

3.2.6. MISSING VALUES

In a balanced assay, an accident totally unconnected with the applied treatments may lead to the loss of one or more responses, for example because an animal dies. If it is considered that the accident is in no way connected with the composition of the preparation administered, the exact calculations can still be performed but the formulae are necessarily more complicated and can only be given within the framework of general linear models (see Section 7.1). However, there exists an approximate method which keeps the simplicity of the balanced design by replacing the missing response by a calculated value. The loss of information is taken into account by diminishing the degrees of freedom for the total sum of squares and for the residual error by the number of missing values and using one of the formulae below for the missing values. It should be borne in mind that this is only an approximate method, and that the exact method is to be preferred.

If more than one observation is missing, the same formulae can be used. The procedure is to make a rough guess at all the missing values except one, and to use the proper formula for this one, using all the remaining values including the rough guesses. Fill in the calculated value. Continue by

similarly calculating a value for the first rough guess. After calculating all the missing values in this way the whole cycle is repeated from the beginning, each calculation using the most recent guessed or calculated value for every response to which the formula is being applied. This continues until 2 consecutive cycles give the same values; convergence is usually rapid.

Provided that the number of values replaced is small relative to the total number of observations in the full experiment (say less than 5 per cent), the approximation implied in this replacement and reduction of degrees of freedom by the number of missing values so replaced is usually fairly satisfactory. The analysis should be interpreted with great care however, especially if there is a preponderance of missing values in one treatment or block, and a biometrician should be consulted if any unusual features are encountered. Replacing missing values in a test without replication is a particularly delicate operation.

Completely randomised design

In a completely randomised assay the missing value can be replaced by the arithmetic mean of the other responses to the same treatment.

Randomised block design

The missing value is obtained using the equation:

$$y' = \frac{nB' + kT' - G'}{(n - 1)\,(k - 1)} \qquad (3.2.6.\text{-}1)$$

where B' is the sum of the responses in the block containing the missing value, T' the corresponding treatment total and G' is the sum of all responses recorded in the assay.

Latin square design

The missing value y' is obtained from:

$$y' = \frac{k\,(B' + C' + T') - 2G'}{(k - 1)\,(k - 2)} \qquad (3.2.6.\text{-}2)$$

where B' and C' are the sums of the responses in the row and column containing the missing value. In this case $k = n$.

Cross-over design

If an accident leading to loss of values occurs in a cross-over design, a book on statistics should be consulted (e.g. D.J. Finney, see Section 10), because the appropriate formulae depend upon the particular treatment combinations.

3.3. THE SLOPE-RATIO MODEL

3.3.1. INTRODUCTION

This model is suitable, for example, for some microbiological assays when the independent variable is the concentration of an essential growth factor below the optimal concentration of the medium. The slope-ratio model is illustrated in Figure 3.3.1.-I.
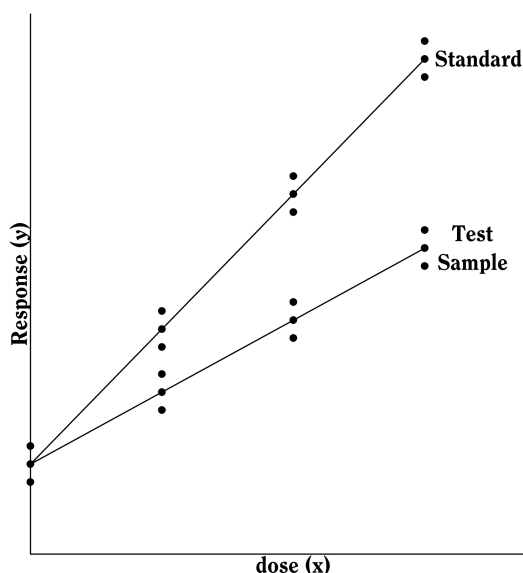
Figure 3.3.1.-I. − *The slope-ratio model for a 2 × 3 + 1 assay*

The doses are represented on the horizontal axis with zero concentration on the left and the highest concentration on the right. The responses are indicated on the vertical axis. The individual responses to each treatment are indicated with black dots. The 2 lines are the calculated dose-response relationship for the standard and the unknown under the assumption that they intersect each other at zero-dose. Unlike the parallel-line model, the doses are not transformed to logarithms.

Just as in the case of an assay based on the parallel-line model, it is important that the assumed potency is close to the true potency, and to prepare equipotent dilutions of the test preparations and the standard (if feasible). The more nearly correct the assumed potency, the closer the 2 lines will be together. The ratio of the slopes represents the "true" potency of the unknown, relative to its assumed potency. If the slope of the unknown preparation is steeper than that of the standard, the potency was underestimated and the calculations will indicate an estimated potency higher than the assumed potency. Similarly, if the slope of the unknown is less steep than that of the standard, the potency was overestimated and the calculations will result in an estimated potency lower than the assumed potency.

In setting up an experiment, all responses should be examined for the fulfilment of the conditions 1, 2 and 3 in Section 3.1. The analysis of variance to be performed in routine is described in Section 3.3.3 so that compliance with conditions 4B and 5B of Section 3.1 may be examined.

3.3.2. ASSAY DESIGN
The use of the statistical analysis presented below imposes the following restrictions on the assay:

a) the standard and the test preparations must be tested with the same number of equally spaced dilutions,

b) an extra group of experimental units receiving no treatment may be tested (the blanks),

c) there must be an equal number of experimental units to each treatment.

As already remarked in Section 3.1.3, assay designs not meeting these restrictions may be both possible and correct, but the simple statistical analyses presented here are no longer applicable and either expert advice should be sought or suitable software should be used.

A design with 2 doses per preparation and 1 blank, the "common zero (2*h* + 1)-design", is usually preferred, since it gives the highest precision combined with the possibility to check validity within the constraints mentioned above. However, a linear relationship cannot always be assumed to be valid down to zero-dose. With a slight loss of precision a design without blanks may be adopted. In this case 3 doses per preparation, the "common zero (3*h*)-design", are preferred to 2 doses per preparation. The doses are thus given as follows:

1) the standard is given in a high dose, near to but not exceeding the highest dose giving a mean response on the straight portion of the dose-response line,

2) the other doses are uniformly spaced between the highest dose and zero dose,

3) the test preparations are given in corresponding doses based on the assumed potency of the material.

A completely randomised, a randomised block or a Latin square design may be used, such as described in Section 3.2.2. The use of any of these designs necessitates an adjustment to the error sum of squares as described for assays based on the parallel-line model. The analysis of an assay of one or more test preparations against a standard is described below.

3.3.3. ANALYSIS OF VARIANCE

3.3.3.1. The (*hd* + 1)-design

The responses are verified as described in Section 3.1 and, if necessary, transformed. The responses are then averaged over each treatment and each preparation as shown in Table 3.3.3.1.-I. Additionally, the mean response for blanks (*B*) is calculated.

The sums of squares in the analysis of variance are calculated as shown in Tables 3.3.3.1.-I to 3.3.3.1.-III. The sum of squares due to non-linearity can only be calculated if at least 3 doses of each preparation have been included in the assay. The residual error is obtained by subtracting the variations allowed for in the design from the total variation in response (Table 3.3.3.1.-IV).

The analysis of variance is now completed as follows. Each sum of squares is divided by the corresponding number of degrees of freedom to give mean squares. The mean square for each variable to be tested is now expressed as a ratio to the residual error (s^2) and the significance of these values (known as *F*-ratios) are assessed by use of Table 8.1 or a suitable sub-routine of a computer program.

3.3.3.2. The (*hd*)-design

The formulae are basically the same as those for the (*hd* + 1)-design, but there are some slight differences.

— *B* is discarded from all formulae.

— $K = \dfrac{n\,(P_S + P_T + ...)^2}{hd}$

— SS_{blank} is removed from the analysis of variance.

— The number of degrees of freedom for treatments becomes $hd - 1$.

— The number of degrees of freedom of the residual error and the total variance is calculated as described for the parallel-line model (see Table 3.2.3.-IV).

Validity of the assay, potency and confidence interval are found as described in Sections 3.3.4 and 3.3.5.

Table 3.3.3.1.-I. — *Formulae for slope-ratio assays with d doses of each preparation and a blank*

	Standard (S)	1st Test sample (T)	2nd Test sample (U, etc.)
Mean response lowest dose	S_1	T_1	U_1
Mean response 2nd dose	S_2	T_2	U_2
...	...	...	...
Mean response highest dose	S_d	T_d	U_d
Total preparation	$P_S = S_1 + S_2 + ... + S_d$	$P_T = T_1 + T_2 + ... + T_d$	$P_U = ...$
Linear product	$L_S = 1S_1 + 2S_2 + ... + dS_d$	$L_T = 1T_1 + 2T_2 + ... + dT_d$	$L_U = ...$
Intercept value	$a_S = (4d + 2) P_S - 6L_S$	$a_T = (4d + 2) P_T - 6L_T$	$a_U = ...$
Slope value	$b_S = 2L_S - (d + 1) P_S$	$b_T = 2L_T - (d + 1) P_T$	$b_U = ...$
Treatment value	$G_S = S_1^2 + ... + S_d^2$	$G_T = T_1^2 + ... + T_d^2$	$G_U = ...$
Non-linearity(*)	$J_S = G_S - \dfrac{P_S^2}{d} - \dfrac{3b_S^2}{d^3 - d}$	$J_T = G_T - \dfrac{P_T^2}{d} - \dfrac{3b_T^2}{d^3 - d}$	$J_U = ...$

(*) Not calculated for two-dose assays

Table 3.3.3.1.-II. — *Additional formulae for the construction of the analysis of variance*

$H_B = \dfrac{nhd^2 - nhd}{hd^2 - hd + 4d + 2}$	$H_I = \dfrac{n}{4d^3 - 2d^2 - 2d}$	$a = \dfrac{a_S + a_T + ...}{h (d^2 - d)}$	$K = \dfrac{n (B + P_S + P_T + ...)^2}{hd + 1}$

Table 3.3.3.1.-III. — *Formulae to calculate the sum of squares and degrees of freedom*

Source of variation	Degrees of freedom (*f*)	Sum of squares
Regression	h	$SS_{reg} = SS_{treat} - SS_{blank} - SS_{int} - SS_{lin}$
Blanks	1	$SS_{blank} = H_B (B - a)^2$
Intersection	$h - 1$	$SS_{int} = H_I \left((a_S^2 + a_T^2 + ...) - h (d^2 - d)^2 a^2 \right)$
Non-linearity(*)	$h (d - 2)$	$SS_{lin} = n (J_S + J_T + ...)$
Treatments	hd	$SS_{treat} = n (B^2 + G_S + G_T + ...) - K$

(*) Not calculated for two-dose assays

3.3.4. TESTS OF VALIDITY

Assay results are said to be "statistically valid" if the outcome of the analysis of variance is as follows:

1) the variation due to blanks in (hd + 1)-designs is not significant, i.e. the calculated probability is not smaller than 0.05. This indicates that the responses of the blanks do not significantly differ from the common intercept and the linear relationship is valid down to zero dose;

2) the variation due to intersection is not significant, i.e. the calculated probability is not less than 0.05. This indicates that condition 5B, Section 3.1 is satisfied;

3) in assays including at least 3 doses per preparation, the variation due to non-linearity is not significant, i.e. the calculated probability is not less than 0.05. This indicates that condition 4B, Section 3.1 is satisfied.

A significant variation due to blanks indicates that the hypothesis of linearity is not valid near zero dose. If this is likely to be systematic rather than incidental for the type of assay, the (hd-design) is more appropriate. Any response to blanks should then be disregarded.

When these tests indicate that the assay is valid, the potency is calculated with its confidence limits as described in Section 3.3.5.

3.3.5. ESTIMATION OF POTENCY AND CONFIDENCE LIMITS

3.3.5.1. The (hd + 1)-design

The common intersection a' of the preparations can be calculated from:

$$a' = \frac{(2d + 1) B + (2d - 3) ha}{h (2d - 3) + 2d + 1} \qquad (3.3.5.1.\text{-}1)$$

The slope of the standard, and similarly for each of the other preparations, is calculated from:

$$b'_S = \frac{6L_S - 3d (d + 1) a'}{2d^3 + 3d^2 + d} \qquad (3.3.5.1.\text{-}2)$$

The potency ratio of each of the test preparations can now be calculated from:

Table 3.3.3.1.-IV. — *Estimation of the residual error*

Source of variation		Degrees of freedom	Sum of squares
Blocks (rows)[*]		$n - 1$	$SS_{block} = hd\,(R_1^2 + ... + R_n^2) - K$
Columns[**]		$n - 1$	$SS_{col} = hd\,(C_1^2 + ... + C_n^2) - K$
Residual error[***]	Completely randomised	$(hd + 1)(n - 1)$	$SS_{res} = SS_{tot} - SS_{treat}$
	Randomised block	$hd\,(n - 1)$	$SS_{res} = SS_{tot} - SS_{treat} - SS_{block}$
	Latin square	$(hd - 1)(n - 1)$	$SS_{res} = SS_{tot} - SS_{treat} - SS_{block} - SS_{col}$
Total		$nhd + n - 1$	$SS_{tot} = \sum (y - \overline{y})^2$

[*] Not calculated for completely randomised designs
[**] Only calculated for Latin square designs
[***] Depends on the type of design

$$R'_T = \frac{b'_T}{b'_S} \qquad (3.3.5.1.\text{-}3)$$

which has to be multiplied by A_T, the assumed potency of the test preparation, in order to find the estimated potency R_T. If the step between adjacent doses was not identical for the standard and the test preparation, the potency has to be multiplied by I_S/I_T. Note that, unlike the parallel-line analysis, no antilogarithms are calculated.

The confidence interval for R_T' is calculated from:

$$CR'_T - K' \pm \sqrt{(C - 1)(CR_T'^2 + 1) + K'(K' - 2CR'_T)}$$
$$(3.3.5.1.\text{-}4)$$

where $C = \dfrac{b_S'^2}{b_S'^2 - s^2 t^2 V_1}$ and $K' = (C - 1)V_2$

V_1 are V_2 are related to the variance and covariance of the numerator and denominator of R_T. They can be obtained from:

$$V_1 = \frac{6}{n\,(2d + 1)}\left(\frac{1}{d\,(d + 1)} + \frac{3}{2\,(2d + 1) + hd\,(d - 1)}\right)$$
$$(3.3.5.1.\text{-}5)$$

$$V_2 = \frac{3d\,(d + 1)}{(3d + 1)(d + 2) + hd\,(d - 1)} \qquad (3.3.5.1.\text{-}6)$$

The confidence limits are multiplied by A_T, and if necessary by I_S/I_T.

3.3.5.2. The (*hd*)-design

The formulae are the same as for the (*hd* + 1)-design, with the following modifications:

$$a' = a \qquad (3.3.5.2.\text{-}1)$$

$$V_1 = \frac{6}{nd\,(2d + 1)}\left(\frac{1}{d + 1} + \frac{3}{h\,(d - 1)}\right) \qquad (3.3.5.2.\text{-}2)$$

$$V_2 = \frac{3\,(d + 1)}{3\,(d + 1) + h\,(d - 1)} \qquad (3.3.5.2.\text{-}3)$$

3.4. EXTENDED SIGMOID DOSE-RESPONSE CURVES

This model is suitable, for example, for some immunoassays when analysis is required of extended sigmoid dose-response curves. This model is illustrated in Figure 3.4.-I.
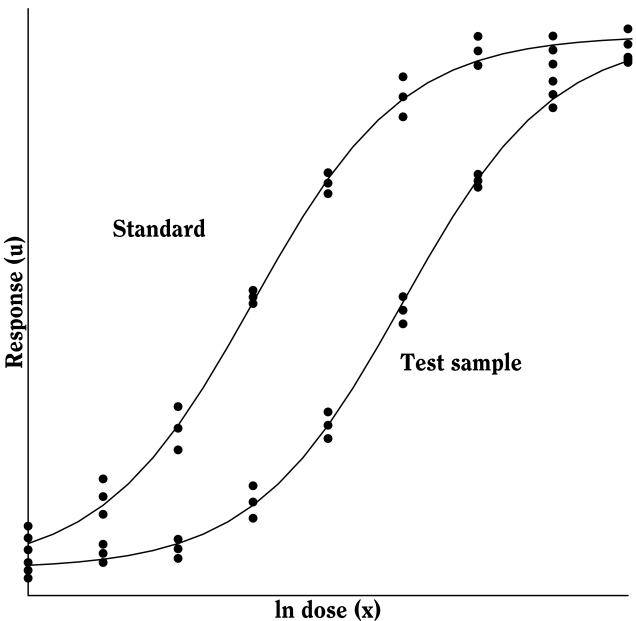
Figure 3.4.-I. − *The four-parameter logistic curve model*

The logarithms of the doses are represented on the horizontal axis with the lowest concentration on the left and the highest concentration on the right. The responses are indicated on the vertical axis. The individual responses to each treatment are indicated with black dots. The 2 curves are the calculated ln(dose)-response relationship for the standard and the test preparation.

The general shape of the curves can usually be described by a logistic function but other shapes are also possible. Each curve can be characterised by 4 parameters: The upper asymptote (α), the lower asymptote (δ), the slope-factor (β), and the horizontal location (γ). This model is therefore often referred to as a four-parameter model. A mathematical representation of the ln(dose)-response curve is:

$$u = \delta + \frac{\alpha - \delta}{1 + e^{-\beta(x - \gamma)}}$$

For a valid assay it is necessary that the curves of the standard and the test preparations have the same slope-factor, and the same maximum and minimum response level at the extreme parts. Only the horizontal location (γ) of the curves may be different. The horizontal distance between the curves is related to the "true" potency of the unknown. If the assay is used routinely, it may be sufficient to test the condition of equal upper and lower response levels when the assay is

developed, and then to retest this condition directly only at suitable intervals or when there are changes in materials or assay conditions.

The maximum-likelihood estimates of the parameters and their confidence intervals can be obtained with suitable computer programs. These computer programs may include some statistical tests reflecting validity. For example, if the maximum likelihood estimation shows significant deviations from the fitted model under the assumed conditions of equal upper and lower asymptotes and slopes, then one or all of these conditions may not be satisfied.

The logistic model raises a number of statistical problems which may require different solutions for different types of assays, and no simple summary is possible. A wide variety of possible approaches is described in the relevant literature. Professional advice is therefore recommended for this type of analysis. A simple example is nevertheless included in Section 5.4 to illustrate a "possible" way to analyse the data presented. A short discussion of alternative approaches and other statistical considerations is given in Section 7.5.

If professional advice or suitable software is not available, alternative approaches are possible: 1) if "reasonable" estimates of the upper limit (α) and lower limit (δ) are available, select for all preparations the doses with mean of the responses (u) falling between approximately 20 per cent and 80 per cent of the limits, transform responses of the selected doses to $y = \ln\left(\dfrac{u - \delta}{\alpha - u}\right)$ and use the parallel line model (Section 3.2) for the analysis; 2) select a range of doses for which the responses (u) or suitably transformed responses, for example $\ln(u)$, are approximately linear when plotted against $\ln(\text{dose})$; the parallel line model (Section 3.2) may then be used for analysis.

4. ASSAYS DEPENDING UPON QUANTAL RESPONSES

4.1. INTRODUCTION

In certain assays it is impossible or excessively laborious to measure the effect on each experimental unit on a quantitative scale. Instead, an effect such as death or hypoglycaemic symptoms may be observed as either occurring or not occurring in each unit, and the result depends on the number of units in which it occurs. Such assays are called quantal or all-or-none.

The situation is very similar to that described for quantitative assays in Section 3.1, but in place of n separate responses to each treatment a single value is recorded, i.e. the fraction of units in each treatment group showing a response. When these fractions are plotted against the logarithms of the doses the resulting curve will tend to be sigmoid (S-shaped) rather than linear. A mathematical function that represents this sigmoid curvature is used to estimate the dose-response curve. The most commonly used function is the cumulative normal distribution function. This function has some theoretical merit, and is perhaps the best choice if the response is a reflection of the tolerance of the units. If the response is more likely to depend upon a process of growth, the logistic distribution model is preferred, although the difference in outcome between the 2 models is usually very small.

The maximum likelihood estimators of the slope and location of the curves can be found only by applying an iterative procedure. There are many procedures which lead to the same outcome, but they differ in efficiency due to the speed of convergence. One of the most rapid methods is direct optimisation of the maximum-likelihood function (see Section 7.1), which can easily be performed with computer programs having a built-in procedure for this purpose. Unfortunately, most of these procedures do not yield an estimate of the confidence interval, and the technique to obtain it is too complicated to describe here. The technique described below is not the most rapid, but has been chosen for its simplicity compared to the alternatives. It can be used for assays in which one or more test preparations are compared to a standard. Furthermore, the following conditions must be fulfilled:

1) the relationship between the logarithm of the dose and the response can be represented by a cumulative normal distribution curve,

2) the curves for the standard and the test preparation are parallel, i.e. they are identically shaped and may only differ in their horizontal location,

3) in theory, there is no natural response to extremely low doses and no natural non-response to extremely high doses.

4.2. THE PROBIT METHOD

The sigmoid curve can be made linear by replacing each response, i.e. the fraction of positive responses per group, by the corresponding value of the cumulative standard normal distribution. This value, often referred to as "normit", ranges theoretically from $-\infty$ to $+\infty$. In the past it was proposed to add 5 to each normit to obtain "probits". This facilitated the hand-performed calculations because negative values were avoided. With the arrival of computers the need to add 5 to the normits has disappeared. The term "normit method" would therefore be better for the method described below. However, since the term "probit analysis" is so widely spread, the term will, for historical reasons, be maintained in this text.

Once the responses have been linearised, it should be possible to apply the parallel-line analysis as described in Section 3.2. Unfortunately, the validity condition of homogeneity of variance for each dose is not fulfilled. The variance is minimal at normit = 0 and increases for positive and negative values of the normit. It is therefore necessary to give more weight to responses in the middle part of the curve, and less weight to the more extreme parts of the curve. This method, the analysis of variance, and the estimation of the potency and confidence interval are described below.

4.2.1. TABULATION OF THE RESULTS

Table 4.2.1.-I is used to enter the data into the columns indicated by numbers:

(1) the dose of the standard or the test preparation,

(2) the number n of units submitted to that treatment,

(3) the number of units r giving a positive response to the treatment,

(4) the logarithm x of the dose,

(5) the fraction $p = r/n$ of positive responses per group.

The first cycle starts here.

(6) column Y is filled with zeros at the first iteration,

(7) the corresponding value $\Phi = \Phi(Y)$ of the cumulative standard normal distribution function (see also Table 8.4).

The columns (8) to (10) are calculated with the following formulae:

(8)
$$Z = \frac{e^{-Y^2/2}}{\sqrt{2\pi}}$$
(4.2.1.-1)

(9)
$$y = Y + \frac{p - \Phi}{Z}$$
(4.2.1.-2)

(10)
$$w = \frac{nZ^2}{\Phi - \Phi^2}$$
(4.2.1.-3)

The columns (11) to (15) can easily be calculated from columns (4), (9) and (10) as wx, wy, wx^2, wy^2 and wxy respectively, and the sum (Σ) of each of the columns (10) to (15) is calculated separately for each of the preparations.

The sums calculated in Table 4.2.1.-I are transferred to columns (1) to (6) of Table 4.2.1.-II and 6 additional columns (7) to (12) are calculated as follows:

(7)
$$S_{xx} = \sum wx^2 - \frac{\left(\sum wx\right)^2}{\sum w} \qquad (4.2.1.\text{-}4)$$

(8)
$$S_{xy} = \sum wxy - \frac{\left(\sum wx\right)\left(\sum wy\right)}{\sum w} \qquad (4.2.1.\text{-}5)$$

(9)
$$S_{yy} = \sum wy^2 - \frac{\left(\sum wy\right)^2}{\sum w} \qquad (4.2.1.\text{-}6)$$

(10)
$$\overline{x} = \frac{\sum wx}{\sum w} \qquad (4.2.1.\text{-}7)$$

(11)
$$\overline{y} = \frac{\sum wy}{\sum w} \qquad (4.2.1.\text{-}8)$$

The common slope b can now be obtained as:

$$b = \frac{\sum S_{xy}}{\sum S_{xx}} \qquad (4.2.1.\text{-}9)$$

and the intercept a of the standard, and similarly for the test preparations is obtained as:

(12)
$$a = \overline{y} - b\overline{x} \qquad (4.2.1.\text{-}10)$$

Column (6) of the first working table can now be replaced by $Y = a + bx$ and the cycle is repeated until the difference between 2 cycles has become small (e.g. the maximum difference of Y between 2 consecutive cycles is smaller than 10^{-8}).

4.2.2. TESTS OF VALIDITY

Before calculating the potencies and confidence intervals, validity of the assay must be assessed. If at least 3 doses for each preparation have been included, the deviations from linearity can be measured as follows: add a 13th column to Table 4.2.1.-II and fill it with:

$$S_{yy} - \frac{S_{xy}^2}{S_{xx}} \qquad (4.2.2.\text{-}1)$$

The column total is a measure of deviations from linearity and is approximately χ^2 distributed with degrees of freedom equal to $N - 2h$. Significance of this value may be assessed with the aid of Table 8.3 or a suitable sub-routine in a computer program. If the value is significant at the 0.05 probability level, the assay must probably be rejected (see Section 4.2.4).

When the above test gives no indication of significant deviations from linear regression, the deviations from parallelism are tested at the 0.05 significance level with:

$$\chi^2 = \sum \frac{S_{xy}^2}{S_{xx}} - \frac{\left(\sum S_{xy}\right)^2}{\sum S_{xx}} \qquad (4.2.2.\text{-}2)$$

with $h - 1$ degrees of freedom.

4.2.3. ESTIMATION OF POTENCY AND CONFIDENCE LIMITS

When there are no indications for a significant departure from parallelism and linearity the ln(potency ratio) M'_T is calculated as:

$$M'_T = \frac{a_T - a_S}{b} \qquad (4.2.3.\text{-}1)$$

and the antilogarithm is taken. Now let $t = 1.96$ and $s = 1$. Confidence limits are calculated as the antilogarithms of:

$$CM'_T - (C-1)(\overline{x}_S - \overline{x}_T) \pm \sqrt{(C-1)\left(V\sum S_{xx} + C\left(M'_T - \overline{x}_S + \overline{x}_T\right)^2\right)} \qquad (4.2.3.\text{-}2)$$

where $C = \dfrac{b^2 \sum S_{xx}}{b^2 \sum S_{xx} - s^2 t^2}$ and $V = \dfrac{1}{\sum\limits_{S} w} + \dfrac{1}{\sum\limits_{T} w}$

Table 4.2.1.-I. — *First working table*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	dose	n	r	x	p	Y	Φ	Z	y	w	wx	wy	wx^2	wy^2	wxy
S	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
										$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$
T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
										$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$	$\Sigma =$
etc.															

Table 4.2.1.-II. — *Second working table*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Σw	Σwx	Σwy	Σwx^2	Σwy^2	Σwxy	S_{xx}	S_{xy}	S_{yy}	$\overline{x}$	$\overline{y}$	a
S	.	.	.	.	.	.	.	.	.	.	.	.
T	.	.	.	.	.	.	.	.	.	.	.	.
etc.	.	.	.	.	.	.	.	.	.	.	.	.
							$\Sigma =$	$\Sigma =$				

4.2.4. INVALID ASSAYS

If the test for deviations from linearity described in Section 4.2.2 is significant, the assay should normally be rejected. If there are reasons to retain the assay, the formulae are slightly modified. t becomes the t-value ($p = 0.05$) with the same number of degrees of freedom as used in the check for linearity and s^2 becomes the χ^2 value divided by the same number of degrees of freedom (and thus typically is greater than 1).

The test for parallelism is also slightly modified. The χ^2 value for non-parallelism is divided by its number of degrees of freedom. The resulting value is divided by s^2 calculated above to obtain an F-ratio with h - 1 and N - $2h$ degrees of freedom, which is evaluated in the usual way at the 0.05 significance level.

4.3. THE LOGIT METHOD

As indicated in Section 4.1 the logit method may sometimes be more appropriate. The name of the method is derived from the logit function which is the inverse of the logistic distribution. The procedure is similar to that described for the probit method with the following modifications in the formulae for Φ and Z.

$$\Phi = \frac{1}{1 + e^{-Y}} \qquad (4.3.\text{-}1)$$

$$Z = \frac{e^{-Y}}{\left(1 + e^{-Y}\right)^2} \qquad (4.3.\text{-}2)$$

4.4. OTHER SHAPES OF THE CURVE

The probit and logit method are almost always adequate for the analysis of quantal responses called for in the European Pharmacopoeia. However, if it can be made evident that the ln(dose)-response curve has another shape than the 2 curves described above, another curve Φ may be adopted. Z is then taken to be the first derivative of Φ.

For example, if it can be shown that the curve is not symmetric, the Gompertz distribution may be appropriate (the gompit method) in which case $\Phi = 1 - e^{-e^Y}$ and $Z = e^{Y-e^Y}$.

4.5. THE MEDIAN EFFECTIVE DOSE

In some types of assay it is desirable to determine a median effective dose which is the dose that produces a response in 50 per cent of the units. The probit method can be used to determine this median effective dose (ED_{50}), but since there is no need to express this dose relative to a standard, the formulae are slightly different.

Note: a standard can optionally be included in order to validate the assay. Usually the assay is considered valid if the calculated ED_{50} of the standard is close enough to the assigned ED_{50}. What "close enough" in this context means depends on the requirements specified in the monograph.

The tabulation of the responses to the test samples, and optionally a standard, is as described in Section 4.2.1. The test for linearity is as described in Section 4.2.2. A test for parallelism is not necessary for this type of assay. The ED_{50} of test sample T, and similarly for the other samples, is obtained as described in Section 4.2.3, with the following modifications in formulae 4.2.3.-1 and 4.2.3.-2).

$$M'_T = \frac{-a_T}{b} \qquad (4.5.\text{-}1)$$

$$C M'_T - (C-1)\overline{x}_T \pm \sqrt{(C-1)\left(V \sum S_{xx} + C\left(M'_T + \overline{x}_T\right)^2\right)} \qquad (4.5.\text{-}2)$$

where $V = \dfrac{1}{\sum\limits_T w}$ and C is left unchanged

5. EXAMPLES

This section consists of worked examples illustrating the application of the formulae. The examples have been selected primarily to illustrate the statistical method of calculation. They are not intended to reflect the most suitable method of assay, if alternatives are permitted in the individual monographs. To increase their value as program checks, more decimal places are given than would usually be necessary. It should also be noted that other, but equivalent methods of calculation exist. These methods should lead to exactly the same final results as those given in the examples.

5.1. PARALLEL-LINE MODEL

5.1.1. TWO-DOSE MULTIPLE ASSAY WITH COMPLETELY RANDOMISED DESIGN
An assay of corticotrophin by subcutaneous injection in rats

The standard preparation is administered at 0.25 and 1.0 unit per 100 g of body mass. 2 preparations to be examined are both assumed to have a potency of 1 unit per milligram and they are administered in the same quantities as the standard. The individual responses and means per treatment are given in Table 5.1.1.-I. A graphical presentation (Figure 5.1.1.I) gives no rise to doubt the homogeneity of variance and normality of the data, but suggests problems with parallelism for preparation U.

Table 5.1.1.-I. — *Response metameter y - mass of ascorbic acid (mg) per 100 g of adrenal gland*

	Standard S		Preparation T		Preparation U	
	S_1	S_2	T_1	T_2	U_1	U_2
	300	289	310	230	250	236
	310	221	290	210	268	213
	330	267	360	280	273	283
	290	236	341	261	240	269
	364	250	321	241	307	251
	328	231	370	290	270	294
	390	229	303	223	317	223
	360	269	334	254	312	250
	342	233	295	216	320	216
	306	259	315	235	265	265
Mean	332.0	248.4	323.9	244.0	282.2	250.0
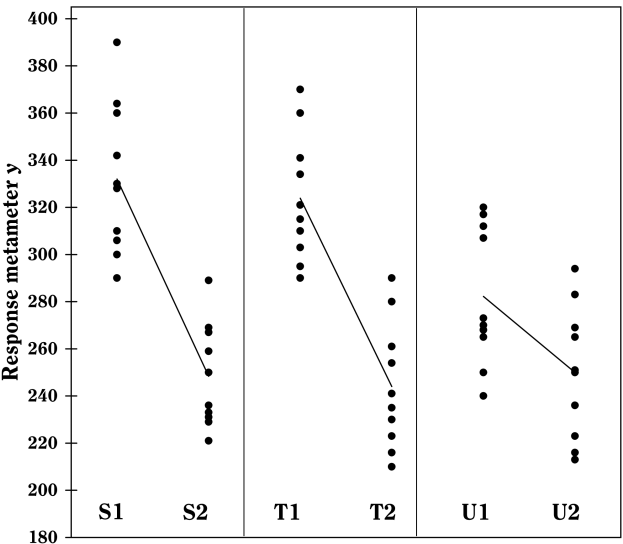
Figure 5.1.1.-I.

The formulae in Tables 3.2.3.-I and 3.2.3.-II lead to:

P_S = 580.4 $\qquad$ L_S = −41.8

P_T = 567.9 $\qquad$ L_T = −39.95

P_U = 532.2 $\qquad$ L_U = −16.1

$H_P = \dfrac{10}{2} = 5$ $\qquad$ $H_L = \dfrac{120}{6} = 20$

The analysis of variance can now be completed with the formulae in Tables 3.2.3-III and 3.2.3.-IV. This is shown in Table 5.1.1.-II.

Table 5.1.1.-II. — *Analysis of variance*

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F*-ratio	Proba-bility
Preparations	2	6256.6	3128.3		
Regression	1	63 830.8	63 830.8	83.38	0.000
Non-parallelism	2	8218.2	4109.1	5.37	0.007
Treatments	5	78 305.7			
Residual error	54	41 340.9	765.57		
Total	59	119 646.6			

The analysis confirms a highly significant linear regression. Departure from parallelism, however, is also significant ($p = 0.0075$) which was to be expected from the graphical observation that preparation U is not parallel to the standard. This preparation is therefore rejected and the analysis repeated using only preparation T and the standard.

Table 5.1.1.-III. — *Analysis of variance without sample U*

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F*-ratio	Proba-bility
Preparations	1	390.6	390.6		
Regression	1	66 830.6	66 830.6	90.5	0.000
Non-parallelism	1	34.2	34.2	0.05	0.831
Treatments	3	67 255.5			
Residual error	36	26 587.3	738.54		
Total	39	93 842.8			

The analysis without preparation U results in compliance with the requirements with respect to both regression and parallelism and so the potency can be calculated. The formulae in Section 3.2.5 give:

— for the common slope:

$$b = \frac{20\,(-41.8 - 39.95)}{\ln4 \times 10 \times 2} = -58.970$$

— the ln(potency ratio) is:

$$M'_T = \frac{567.9 - 580.4}{2 \times (-58.970)} = 0.1060$$

$$C = \frac{66\,830.6}{66\,830.6 - 738.54 \times 2.028^2} = 1.0476$$

$$V = \frac{66\,830.6}{(-58.970)^2 \times 2 \times 10} = 0.9609$$

— and ln(confidence limits) are:

$$1.0476 \times 0.1060 \pm \sqrt{0.0476 \times (1.0476 \times 0.1060^2 + 2 \times 0.9609)}$$
$$= 0.1110 \pm 0.3034$$

By taking the antilogarithms we find a potency ratio of 1.11 with 95 per cent confidence limits from 0.82-1.51.

Multiplying by the assumed potency of preparation T yields a potency of 1.11 units/mg with 95 per cent confidence limits from 0.82 to 1.51 units/mg.

5.1.2. THREE-DOSE LATIN SQUARE DESIGN
Antibiotic agar diffusion assay using a rectangular tray

The standard has an assigned potency of 4855 IU/mg. The test preparation has an assumed potency of 5600 IU/mg. For the stock solutions 25.2 mg of the standard is dissolved in 24.5 ml of solvent and 21.4 mg of the test preparation is dissolved in 23.95 ml of solvent. The final solutions are prepared by first diluting both stock solutions to 1/20 and further using a dilution ratio of 1.5.

A Latin square is generated with the method described in Section 8.6 (see Table 5.1.2.-I). The responses of this routine assay are shown in Table 5.1.2.-II (inhibition zones in mm × 10). The treatment mean values are shown in Table 5.1.2.-III. A graphical representation of the data (see Figure 5.1.2.-I) gives no rise to doubt the normality or homogeneity of variance of the data.

The formulae in Tables 3.2.3.-I and 3.2.3.-II lead to:

P_S = 529.667 $\qquad$ L_S = 35.833

P_T = 526.333 $\qquad$ L_T = 39.333

$H_P = \dfrac{6}{3} = 2$ $\qquad$ $H_L = \dfrac{72}{24} = 3$

The analysis of variance can now be completed with the formulae in Tables 3.2.3.-III and 3.2.3.-IV. The result is shown in Table 5.1.2.-IV.

The analysis shows significant differences between the rows. This indicates the increased precision achieved by using a Latin square design rather than a completely randomised design. A highly significant regression and no significant departure of the individual regression lines from parallelism and linearity confirms that the assay is satisfactory for potency calculations.

Table 5.1.2.-I. — *Distribution of treatments over the plate*

	1	2	3	4	5	6
1	S_1	T_1	T_2	S_3	S_2	T_3
2	T_1	T_3	S_1	S_2	T_2	S_3
3	T_2	S_3	S_2	S_1	T_3	T_1
4	S_3	S_2	T_3	T_1	S_1	T_2
5	S_2	T_2	S_3	T_3	T_1	S_1
6	T_3	S_1	T_1	T_2	S_3	S_2

Table 5.1.2.-II. — *Measured inhibition zones in mm × 10*

	1	2	3	4	5	6	Row mean
1	161	160	178	187	171	194	$175.2 = R_1$
2	151	192	150	172	170	192	$171.2 = R_2$
3	162	195	174	161	193	151	$172.7 = R_3$
4	194	184	199	160	163	171	$178.5 = R_4$
5	176	181	201	202	154	151	$177.5 = R_5$
6	193	166	161	186	198	182	$181.0 = R_6$

Col. Mean =	172.8 $= C_1$	179.7 $= C_2$	177.2 $= C_3$	178.0 $= C_4$	174.8 $= C_5$	173.5 $= C_6$

Table 5.1.2.-III. — *Means of the treatments*

	Standard S			Preparation T		
	S_1	S_2	S_3	T_1	T_2	T_3
Mean	158.67	176.50	194.50	156.17	174.67	195.50
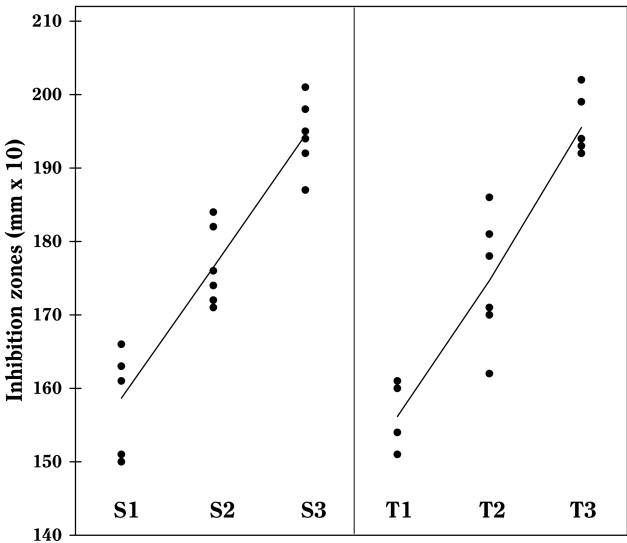
Table 5.1.2.-IV. — *Analysis of variance*

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-ratio	Probability
Preparations	1	11.1111	11.1111		
Regression	1	8475.0417	8475.0417	408.1	0.000
Non-parallelism	1	18.3750	18.3750	0.885	0.358
Non-linearity	2	5.4722	2.7361	0.132	0.877
Treatments	5	8510			
Rows	5	412	82.40	3.968	0.012
Columns	5	218.6667	43.73	2.106	0.107
Residual error	20	415.3333	20.7667		
Total	35	9556			

The formulae in Section 3.2.5 give:

— for the common slope:

$$b = \frac{3 \times (35.833 + 39.333)}{\ln(1.5) \times 6 \times 2} = 46.346$$

— the ln(potency ratio) is:

$$M'_T = \frac{526.333 - 529.667}{3 \times 46.346} = -0.023974$$

$$C = \frac{8475.0417}{8475.0417 - 20.7667 \times 2.086^2} = 1.0108$$

$$V = \frac{8475.0417}{46.346^2 \times 3 \times 6} = 0.2192$$

— and ln(confidence limits) are:

$$\frac{1.0108 \times (-0.0240) \pm}{\sqrt{0.0108 \times \left(1.0108 \times (-0.0240)^2 + 2 \times 0.2192\right)}}$$
$$= -0.02423 \pm 0.06878$$

The potency ratio is found by taking the antilogarithms, resulting in 0.9763 with 95 per cent confidence limits from 0.9112-1.0456.

A correction factor of $\dfrac{4855 \times 25.2/24.5}{5600 \times 21.4/23.95} = 0.99799$ is necessary because the dilutions were not exactly equipotent on the basis of the assumed potency. Multiplying by this correction factor and the assumed potency of 5600 IU/mg yields a potency of 5456 IU/mg with 95 per cent confidence limits from 5092 to 5843 IU/mg.

5.1.3. FOUR-DOSE RANDOMISED BLOCK DESIGN
Antibiotic turbidimetric assay

This assay is designed to assign a potency in international units per vial. The standard has an assigned potency of 670 IU/mg. The test preparation has an assumed potency of 20 000 IU/vial. On the basis of this information the stock solutions are prepared as follows. 16.7 mg of the standard is dissolved in 25 ml solvent and the contents of one vial of the test preparation are dissolved in 40 ml solvent. The final solutions are prepared by first diluting to 1/40 and further using a dilution ratio of 1.5. The tubes are placed in a water-bath in a randomised block arrangement (see Section 8.5). The responses are listed in Table 5.1.3.-I.

Inspection of Figure 5.1.3.-I gives no rise to doubt the validity of the assumptions of normality and homogeneity of variance of the data. The standard deviation of S_3 is somewhat high but is no reason for concern.

Figure 5.1.2.-I.

The formulae in Tables 3.2.3.-I and 3.2.3.-II lead to:

$$P_S = 719.4 \qquad L_S = -229.1$$

$$P_T = 687.6 \qquad L_T = -222$$

$$H_P = \frac{5}{4} = 1.25 \qquad H_L = \frac{60}{60} = 1$$

The analysis of variance is constructed with the formulae in Tables 3.2.3.-III and 3.2.3.-IV. The result is shown in Table 5.1.3.-II.

Table 5.1.3.-I. — *Absorbances of the suspensions (× 1000)*

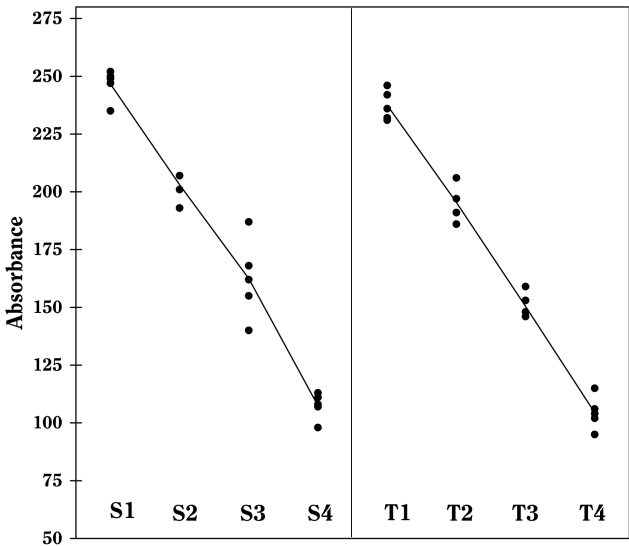| Block	Standard S				Preparation T				Mean
	S_1	S_2	S_3	S_4	T_1	T_2	T_3	T_4	
1	252	207	168	113	242	206	146	115	181.1
2	249	201	187	107	236	197	153	102	179.0
3	247	193	162	111	246	197	148	104	176.0
4	250	207	155	108	231	191	159	106	175.9
5	235	207	140	98	232	186	146	95	167.4
Mean	246.6	203.0	162.4	107.4	237.4	195.4	150.4	104.4	

Figure 5.1.3.-I.

Table 5.1.3.-II. — *Analysis of variance*

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F*-ratio	Proba-bility
Preparations	1	632.025	632.025		
Regression	1	101 745.6	101 745.6	1887.1	0.000
Non-parallelism	1	25.205	25.205	0.467	0.500
Non-linearity	4	259.14	64.785	1.202	0.332
Treatments	7	102 662			
Blocks	4	876.75	219.188	4.065	0.010
Residual error	28	1509.65	53.916		
Total	39	105 048.4			

A significant difference is found between the blocks. This indicates the increased precision achieved by using a randomised block design. A highly significant regression

and no significant departure from parallelism and linearity confirms that the assay is satisfactory for potency calculations. The formulae in Section 3.2.5 give:

— for the common slope:

$$b = \frac{1 \times (-229.1 - 222)}{\ln(1.5) \times 5 \times 2} = -111.255$$

— the ln(potency ratio) is:

$$M_T' = \frac{687.6 - 719.4}{4 \times (-111.255)} = 0.071457$$

$$C = \frac{101\,745.6}{101\,745.6 - 53.916 \times 2.048^2} = 1.00223$$

$$V = \frac{101\,745.6}{(-111.255)^2 \times 4 \times 5} = 0.4110$$

— and ln(confidence limits) are:

$$1.00223 \times 0.0715 \pm$$
$$\sqrt{0.00223 \times (1.00223 \times 0.0715^2 + 2 \times 0.4110)}$$
$$= 0.07162 \pm 0.04293$$

The potency ratio is found by taking the antilogarithms, resulting in 1.0741 with 95 per cent confidence limits from 1.0291 to 1.1214. A correction factor of $\frac{670 \times 16.7/25}{20\,000 \times 1/40} = 0.89512$ is necessary because the dilutions were not exactly equipotent on the basis of the assumed potency. Multiplying by this correction factor and the assumed potency of 20 000 IU/vial yields a potency of 19 228 IU/vial with 95 per cent confidence limits from 18 423-20 075 IU/vial.

5.1.4. FIVE-DOSE MULTIPLE ASSAY WITH COMPLETELY RANDOMISED DESIGN
An in-vitro assay of three hepatitis B vaccines against a standard

3 independent two-fold dilution series of 5 dilutions were prepared from each of the vaccines. After some additional steps in the assay procedure, absorbances were measured. They are shown in Table 5.1.4.-I.

Table 5.1.4.-I. — *Optical densities*

Dilution	Standard S			Preparation T		
1:16 000	0.043	0.045	0.051	0.097	0.097	0.094
1:8000	0.093	0.099	0.082	0.167	0.157	0.178
1:4000	0.159	0.154	0.166	0.327	0.355	0.345
1:2000	0.283	0.295	0.362	0.501	0.665	0.576
1:1000	0.514	0.531	0.545	1.140	1.386	1.051

Dilution	Preparation U			Preparation V		
1:16 000	0.086	0.071	0.073	0.082	0.082	0.086
1:8000	0.127	0.146	0.133	0.145	0.144	0.173
1:4000	0.277	0.268	0.269	0.318	0.306	0.316
1:2000	0.586	0.489	0.546	0.552	0.551	0.624
1:1000	0.957	0.866	1.045	1.037	1.039	1.068

The logarithms of the optical densities are known to have a linear relationship with the logarithms of the doses. The mean responses of the ln-transformed optical densities are listed in Table 5.1.4.-II. No unusual features are discovered in a graphical presentation of the data.

5. General texts

Table 5.1.4.-II. — *Means of the ln-transformed absorbances*

S_1	−3.075	T_1	−2.344	U_1	−2.572	V_1	−2.485
S_2	−2.396	T_2	−1.789	U_2	−2.002	V_2	−1.874
S_3	−1.835	T_3	−1.073	U_3	−1.305	V_3	−1.161
S_4	−1.166	T_4	−0.550	U_4	−0.618	V_4	−0.554
S_5	−0.635	T_5	0.169	U_5	−0.048	V_5	0.047
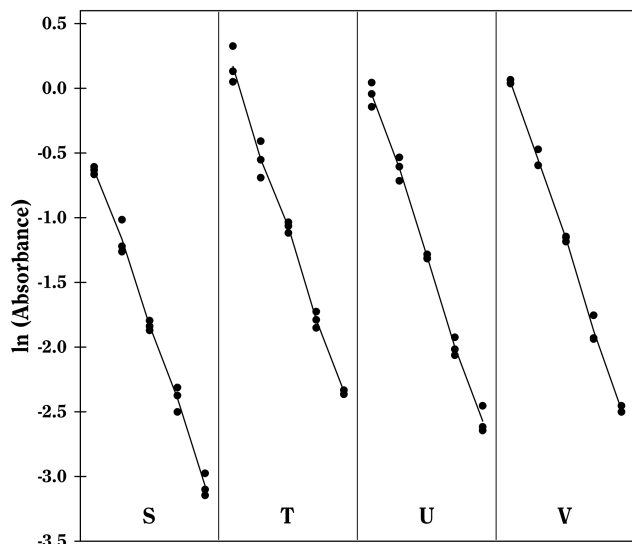
Figure 5.1.4.-I.

The formulae in Tables 3.2.3.-I and 3.2.3.-II give :

P_S	=	−9.108	L_S	=	6.109
P_T	=	−5.586	L_T	=	6.264
P_U	=	−6.544	L_U	=	6.431
P_V	=	−6.027	L_V	=	6.384
H_P	=	$\frac{3}{5} = 0.6$	H_L	=	$\frac{36}{120} = 0.3$

The analysis of variance is completed with the formulae in Tables 3.2.3.-III and 3.2.3.-IV. This is shown in Table 5.1.4.-III.

Table 5.1.4.-III. — *Analysis of variance*

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F*-ratio	Probabil-ity
Preparations	3	4.475	1.492		
Regression	1	47.58	47.58	7126	0.000
Non-parallelism	3	0.0187	0.006	0.933	0.434
Non-linearity	12	0.0742	0.006	0.926	0.531
Treatments	19	52.152			
Residual error	40	0.267	0.0067		
Total	59	52.42			

A highly significant regression and a non-significant departure from parallelism and linearity confirm that the potencies can be safely calculated. The formulae in Section 3.2.5 give :

— for the common slope:

$$ b = \frac{0.3 \times (6.109 + 6.264 + 6.431 + 6.384)}{\ln 2 \times 3 \times 4} = 0.90848 $$

— the ln(potency ratio) for preparation T is :

$$ M'_T = \frac{-5.586 - (-9.108)}{5 \times 0.90848} = 0.7752 $$

$$ C = \frac{47.58}{47.58 - 0.0067 \times 2.021^2} = 1.00057 $$

$$ V = \frac{47.58}{0.9085^2 \times 5 \times 3} = 3.8436 $$

— and ln(confidence limits) for preparation T are :

$$ 1.00057 \times 0.7752 \pm $$
$$ \sqrt{0.00057 \times (1.00057 \times 0.7752^2 + 2 \times 3.8436)} $$
$$ = 0.7756 \pm 0.0689 $$

By taking the antilogarithms a potency ratio of 2.171 is found with 95 per cent confidence limits from 2.027 to 2.327. All samples have an assigned potency of 20 μg protein/ml and so a potency of 43.4 μg protein/ml is found for test preparation T with 95 per cent confidence limits from 40.5-46.5 μg protein/ml.

The same procedure is followed to estimate the potency and confidence interval of the other test preparations. The results are listed in Table 5.1.4.-IV.

Table 5.1.4.-IV. — *Final potency estimates and 95 per cent confidence intervals of the test vaccines (in μg protein/ml)*

	Lower limit	Estimate	Upper limit
Vaccine *T*	40.5	43.4	46.5
Vaccine *U*	32.9	35.2	37.6
Vaccine *V*	36.8	39.4	42.2

5.1.5. TWIN CROSS-OVER DESIGN
Assay of insulin by subcutaneous injection in rabbits

The standard preparation was administered at 1 unit and 2 units per millilitre. Equivalent doses of the unknown preparation were used based on an assumed potency of 40 units per millilitre. The rabbits received subcutaneously 0.5 ml of the appropriate solutions according to the design in Table 5.1.5.-I and responses obtained are shown in Table 5.1.5.-II. The large variance illustrates the variation between rabbits and the need to employ a cross-over design.

Table 5.1.5.-I. — *Arrangements of treatments*

	Group of rabbits			
	1	**2**	**3**	**4**
Day 1	S_1	S_2	T_1	T_2
Day 2	T_2	T_1	S_2	S_1

Table 5.1.5.-II. — *Response y: sum of blood glucose readings (mg/100 ml) at 1 hour and $2\frac{1}{2}$ hours*

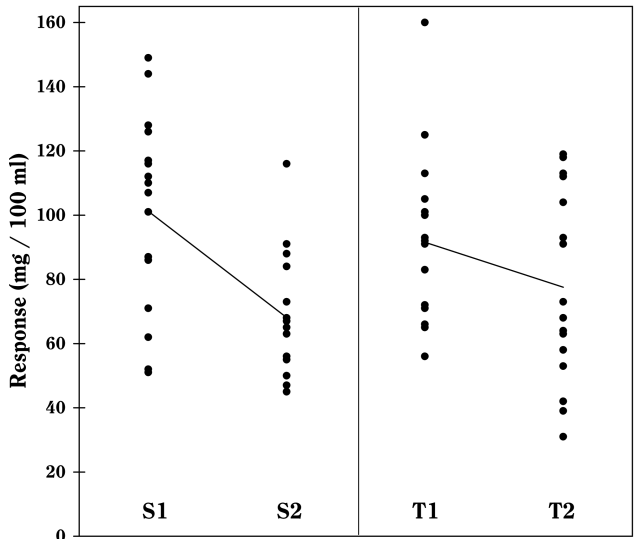|	Group 1		Group 2		Group 3		Group 4	
	S_1	T_2	S_2	T_1	T_1	S_2	T_2	S_1
	112	104	65	72	105	91	118	144
	126	112	116	160	83	67	119	149
	62	58	73	72	125	67	42	51
	86	63	47	93	56	45	64	107
	52	53	88	113	92	84	93	117
	110	113	63	71	101	56	73	128
	116	91	50	65	66	55	39	87
	101	68	55	100	91	68	31	71
Mean	95.6	82.8	69.6	93.3	89.9	66.6	72.4	106.8

Figure 5.1.5.-I.

The analysis of variance is more complicated for this assay than for the other designs given because the component of the sum of squares due to parallelism is not independent of the component due to rabbit differences. Testing of the parallelism of the regression lines involves a second error-mean-square term obtained by subtracting the parallelism component and 2 "interaction" components from the component due to rabbit differences.

3 "interaction" components are present in the analysis of variance due to replication within each group:

days × preparation; days × regression; days × parallelism.

These terms indicate the tendency for the components (preparations, regression and parallelism) to vary from day to day. The corresponding F-ratios thus provide checks on these aspects of assay validity. If the values of F obtained are significantly high, care should be exercised in interpreting the results of the assay and, if possible, the assay should be repeated.

The analysis of variance is constructed by applying the formulae given in Tables 3.2.3.-I to 3.2.3.-III separately for both days and for the pooled set of data. The formulae in Tables 3.2.3.-I and 3.2.3.-II give:

Day 1: P_S =   165.25 L_S = -13

          P_T =   162.25 L_T = -8.75

          H_P = $\dfrac{8}{2} = 4$ H_L = $\dfrac{96}{6} = 16$

Day 2: P_S =   173.38 L_S = -20.06

          P_T =   176.00 L_T = -5.25

          H_P = $\dfrac{8}{2} = 4$ H_L = $\dfrac{96}{6} = 16$

Pooled: P_S =   169.31 L_S = -16.53

          P_T =   169.13 L_T = -7.00

          H_P = $\dfrac{16}{2} = 8$ H_L = $\dfrac{192}{6} = 32$

and with the formulae in Table 3.2.3.-III this leads to:

	Day 1		Day 2		Pooled
SS_{prep} =	18.000	SS_{prep} =	13.781	SS_{prep} =	0.141
SS_{reg} =	3784.5	SS_{reg} =	5125.8	SS_{reg} =	8859.5
SS_{par} =	144.5	SS_{par} =	1755.3	SS_{par} =	1453.5

The interaction terms are found as Day 1 + Day 2 - Pooled.

$$SS_{\text{days}\times\text{prep}} = 31.64$$

$$SS_{\text{days}\times\text{reg}} = 50.77$$

$$SS_{\text{days}\times\text{par}} = 446.27$$

In addition the sum of squares due to day-to-day variation is calculated as:

$$SS_{\text{days}} = \frac{1}{2}N\left(D_1^2 + D_2^2\right) - K = 478.52$$

and the sum of squares due to blocks (the variation between rabbits) as:

$$SS_{\text{block}} = 2\sum_i B_i^2 - K = 39\ 794.7$$

where B_i is the mean response per rabbit.

The analysis of variance can now be completed as shown in Table 5.1.5.-III.

Table 5.1.5.-III. — *Analysis of variance*

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-ratio	Proba-bility
Non-parallelism	1	1453.5	1453.5	1.064	0.311
Days × Prep.	1	31.6	31.6	0.023	0.880
Days × Regr.	1	50.8	50.8	0.037	0.849
Residual error between rabbits	28	38 258.8	1366.4		
Rabbits	31	39 794.7	1283.7		
Preparations	1	0.14	0.14	0.001	0.975
Regression	1	8859.5	8859.5	64.532	0.000
Days	1	478.5	478.5	3.485	0.072
Days × non-par.	1	446.3	446.3	3.251	0.082
Residual error within rabbits	28	3844.1	137.3		
Total	63	53 423.2			

The analysis of variance confirms that the data fulfil the necessary conditions for a satisfactory assay: a highly significant regression, no significant departures from parallelism, and none of the three interaction components is significant.

The formulae in Section 3.2.5 give:

— for the common slope:

$$b = \frac{32 \times (-16.53 - 7)}{\ln 2 \times 16 \times 2} = -33.95$$

— the ln(potency ratio) is:

$$M'_T = \frac{169.13 - 169.31}{2 \times (-33.95)} = 0.00276$$

$$C = \frac{8859.5}{8859.5 - 137.3 \times 2.048^2} = 1.0695$$

$$V = \frac{8859.5}{(-33.95)^2 \times 2 \times 16} = 0.2402$$

— and ln(confidence limits) are:

$$1.0695 \times 0.00276 \pm \sqrt{0.0695 \times (1.0695 \times 0.00276^2 + 2 \times 0.2402)}$$
$$= 0.00295 \pm 0.18279$$

By taking the antilogarithms a potency ratio of 1.003 with 95 per cent confidence limits from 0.835 to 1.204 is found. Multiplying by $A_T = 40$ yields a potency of 40.1 units per millilitre with 95 per cent confidence limits from 33.4-48.2 units per millilitre.

5.2. SLOPE-RATIO MODEL

5.2.1. A COMPLETELY RANDOMISED (0,3,3)-DESIGN
An assay of factor VIII

A laboratory carries out a chromogenic assay of factor VIII activity in concentrates. The laboratory has no experience with the type of assay but is trying to make it operational. 3 equivalent dilutions are prepared of both the standard and the test preparation. In addition a blank is prepared, although a linear dose-response relationship is not expected for low doses. 8 replicates of each dilution are prepared, which is more than would be done in a routine assay.

A graphical presentation of the data shows clearly that the dose-response relationship is indeed not linear at low doses. The responses to blanks will therefore not be used in the calculations (further assays are of course needed to justify this decision). The formulae in Tables 3.3.3.1.-I and 3.3.3.1.-II yield.

P_S = 0.6524	P_T = 0.5651
L_S = 1.4693	L_T = 1.2656
a_S = 0.318	a_T = 0.318
b_S = 0.329	b_T = 0.271
G_S = 0.1554	G_T = 0.1156
J_S = $4.17 \cdot 10^{-8}$	J_T = $2.84 \cdot 10^{-6}$

and

$$H_1 = 0.09524 \qquad a' = 0.05298 \qquad K = 1.9764$$

and the analysis of variance is completed with the formulae in Tables 3.3.3.1.-III and 3.3.3.1.-IV.

A highly significant regression and no significant deviations from linearity and intersection indicate that the potency can be calculated.

Slope of standard:

$$b'_S = \frac{6 \times 1.469 - 36 \times 0.0530}{84} = 0.0822$$

Slope of test sample:

$$b'_T = \frac{6 \times 1.266 - 36 \times 0.0530}{84} = 0.0677$$

Formula 3.3.5.1.-3 gives:

$$R = \frac{0.0677}{0.0822} = 0.823$$

$$C = \frac{0.0822^2}{0.0822^2 - 3.86 \cdot 10^{-6} \times 2.018^2 \times 0.0357} = 1.000083$$

$$K' = 0.000083 \times 0.75 = 0.000062$$

and the 95 per cent confidence limits are:

$$0.823 \pm \sqrt{0.000083 \times 1.678 + 0.000062 \times (-1.646)}$$
$$= 0.823 \pm 0.006$$

The potency ratio is thus estimated as 0.823 with 95 per cent confidence limits from 0.817 to 0.829.

Table 5.2.1.-I. — *Absorbances*

	Blank	Standard *S* (in IU/ml)			Preparation *T* (in IU/ml)		
Conc.	*B*	S_1 0.01	S_2 0.02	S_3 0.03	T_1 0.01	T_2 0.02	T_3 0.03
	0.022	0.133	0.215	0.299	0.120	0.188	0.254
	0.024	0.133	0.215	0.299	0.119	0.188	0.253
	0.024	0.131	0.216	0.299	0.118	0.190	0.255
	0.026	0.136	0.218	0.297	0.120	0.190	0.258
	0.023	0.137	0.220	0.297	0.120	0.190	0.257
	0.022	0.136	0.220	0.305	0.121	0.191	0.257
	0.022	0.138	0.219	0.299	0.121	0.191	0.255
	0.023	0.137	0.218	0.302	0.121	0.190	0.254
Mean	0.0235	0.1351	0.2176	0.2996	0.1200	0.1898	0.2554
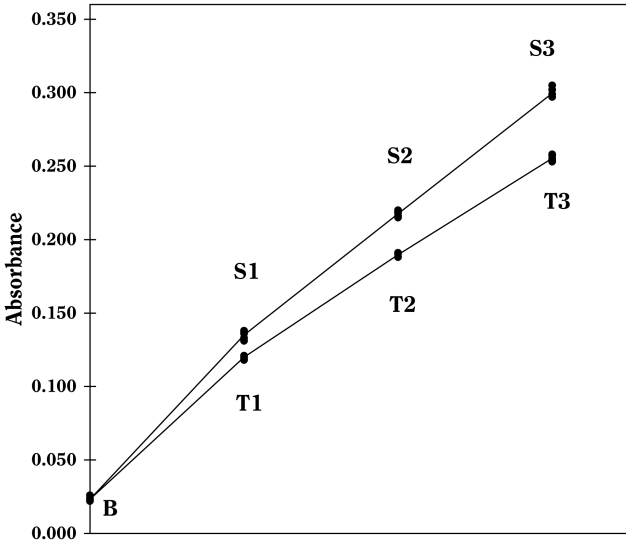
Figure 5.2.1.-I.

Table 5.2.1.-II. — *Analysis of variance*

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F*-ratio	Probabil-ity
Regression	2	0.1917	0.0958	24 850	0.000
Intersection	1	$3 \cdot 10^{-9}$	$3 \cdot 10^{-9}$	$7 \cdot 10^{-4}$	0.978
Non-linearity	2	$2 \cdot 10^{-5}$	$1 \cdot 10^{-5}$	2.984	0.061
Treatments	5	0.1917			
Residual error	42	$1.62 \cdot 10^{-4}$	$3.86 \cdot 10^{-6}$		
Total	47	0.1919			

5.2.2. A COMPLETELY RANDOMISED (0,4,4,4)-DESIGN
An in-vitro assay of influenza vaccines

The haemagglutinin antigen (HA) content of 2 influenza vaccines is determined by single radial immunodiffusion. Both have a labelled potency of 15 µg HA per dose, which is equivalent with a content of 30 µg HA/ml. The standard has an assigned content of 39 µg HA/ml.

Standard and test vaccines are applied in 4 duplicate concentrations which are prepared on the basis of the assigned and the labelled contents. When the equilibrium

between the external and the internal reactant is established, the zone of the annulus precipitation area is measured. The results are shown in Table 5.2.2.-I.

A graphical presentation of the data shows no unusual features. The formulae in Tables 3.3.3.1.-I and 3.3.3.1.-II yield

P_S	= 108.2	P_T	= 103.85	P_U	= 85.8
L_S	= 301.1	L_T	= 292.1	L_U	= 234.1
a_S	= 141.0	a_T	= 116.7	a_U	= 139.8
b_S	= 61.2	b_T	= 64.95	b_U	= 39.2
G_S	= 3114.3	G_T	= 2909.4	G_U	= 1917.3
J_S	= 0.223	J_T	= 2.227	J_U	= 0.083

and

H_I	= 0.0093	a'	= 11.04	K	= 14 785.8

and the analysis of variance is completed with the formulae in Tables 3.3.3.1.-III and 3.3.3.1.-IV. This is shown in Table 5.2.2.-II.

A highly significant regression and no significant deviations from linearity and intersection indicate that the potency can be calculated.

Slope of standard:

$$b'_S = \frac{6 \times 301.1 - 60 \times 11.04}{180} = 6.356$$

Slope of T is:

$$b'_T = \frac{6 \times 292.1 - 60 \times 11.04}{180} = 6.056$$

Slope of U is:

$$b'_U = \frac{6 \times 234.1 - 60 \times 11.04}{180} = 4.123$$

This leads to a potency ratio of 6.056/6.356 = 0.953 for vaccine T and 4.123/6.356 = 0.649 for vaccine U.

$$C = \frac{6.356^2}{6.356^2 - 1.068 \times 2.179^2 \times 0.0444} = 1.0056$$

$$K' = 0.0056 \times 0.625 = 0.0035$$

And the confidence limits are found with formula 3.3.5.1.-4.

For vaccine T:

$$0.955 \pm \sqrt{0.0056 \times 1.913 + 0.0035 \times (-1.913)} = 0.955 \pm 0.063$$

For vaccine U:

$$0.649 \pm \sqrt{0.0056 \times 1.423 + 0.0035 \times (-1.301)} = 0.649 \pm 0.058$$

The HA content in µg/dose can be found by multiplying the potency ratios and confidence limits by the assumed content of 15 µg/dose. The results are given in Table 5.2.2.-III.

Table 5.2.2.-I. — *Zone of precipitation area (mm²)*

| Conc.	Standard S		Preparation T		Preparation U	
(µg/ml)	I	II	I	II	I	II
7.5	18.0	18.0	15.1	16.8	15.4	15.7
15.0	22.8	24.5	23.1	24.2	20.2	18.6
22.5	30.4	30.4	28.9	27.4	24.2	23.1
30.0	35.7	36.6	34.4	37.8	27.4	27.0
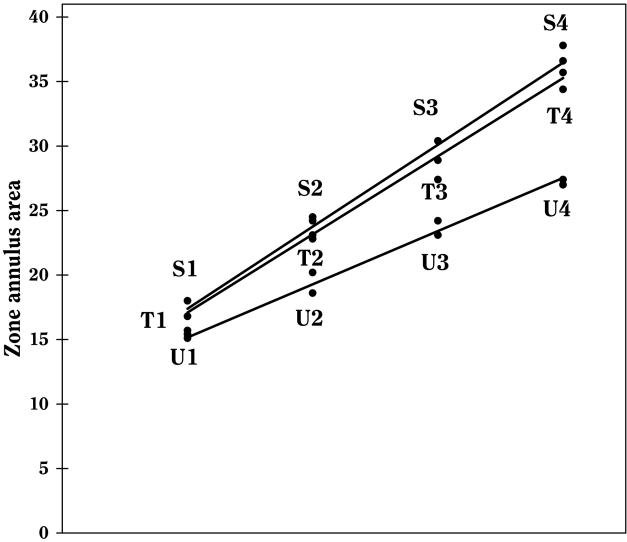
Figure 5.2.2.-I.

Table 5.2.2.-II. — *Analysis of variance*

Source of variation	Degrees of freedom	Sum of squares	Mean square	*F*-ratio	Probability
Regression	3	1087.7	362.6	339.5	0.000
Intersection	2	3.474	1.737	1.626	0.237
Non-linearity	6	5.066	0.844	0.791	0.594
Treatments	11	1096.2			
Residual error	12	12.815	1.068		
Total	23	1109.0			

Table 5.2.2.-III. — *Estimates of HA content* (µg/dose)

	Lower limit	Estimate	Upper limit
Vaccin T	13.4	14.3	15.3
Vaccin U	8.9	9.7	10.6

5.3. QUANTAL RESPONSES

5.3.1. PROBIT ANALYSIS OF A TEST PREPARATION AGAINST A REFERENCE

An in-vivo assay of a diphtheria vaccine

A diphtheria vaccine (assumed potency 140 IU/vial) is assayed against a standard (assigned potency 132 IU/vial). On the basis of this information, equivalent doses are prepared and randomly administered to groups of guinea-pigs. After a given period, the animals are challenged with diphtheria toxin and the number of surviving animals recorded as shown in Table 5.3.1.-I.

Table 5.3.1.-I. — *Raw data from a diphtheria assay in guinea-pigs*

Standard (S) Assigned potency 132 IU/vial			Test preparation (T) Assumed potency 140 IU/vial		
dose (IU/ml)	challenged	protected	dose (I.U./ml)	challenged	protected
1.0	12	0	1.0	11	0
1.6	12	3	1.6	12	4
2.5	12	6	2.5	11	8
4.0	11	10	4.0	11	10

The observations are transferred to the first working table and the subsequent columns are computed as described in Section 4.2.1. Table 5.3.1.-II shows the first cycle of this procedure.

The sums of the last 6 columns are then calculated per preparation and transferred to the second working table (see Table 5.3.1.-III). The results in the other columns are found with formulae 4.2.1.-4 to 4.2.1.-10. This yields a common slope b of 1.655.

The values for Y in the first working table are now replaced by $a + bx$ and a second cycle is carried out (see Table 5.3.1.-IV).

The cycle is repeated until the difference between 2 consecutive cycles has become small. The second working table should then appear as shown in Table 5.3.1.-V.

Linearity is tested as described in Section 4.2.2. The χ^2-value with 4 degrees of freedom is $0.851 + 1.070 = 1.921$ representing a p-value of 0.750 which is not significant.

Since there are no significant deviations from linearity, the test for parallelism can be carried out as described in the same section. The χ^2-value with 1 degree of freedom is

$$(16.71 + 17.27) - \frac{14.15^2}{5.89} = 0.001$$ representing a p-value of 0.974 which is not significant.

The ln(potency ratio) can now be estimated as described in Section 4.2.3.

$$M_T' = \frac{-1.721 - (-2.050)}{2.401} = 0.137$$

Further:

$$C = \frac{2.401^2 \times 5.893}{2.401^2 \times 5.893 - 1^2 \times 1.960^2} = 1.127$$

$$V = \frac{1}{18.37} + \frac{1}{17.96} = 0.110$$

So ln confidence limits are:

$$0.155 - 0.013 \pm \sqrt{0.127 \left(0.649 + 1.127 \times 0.036^2\right)} = 0.142 \pm 0.288$$

The potency and confidence limits can now be found by taking the antilogarithms and multiplying these by the assumed potency of 140 IU/vial. This yields an estimate of 160.6 IU/vial with 95 per cent confidence limits from 121.0-215.2 IU/vial.

Table 5.3.1.-II. — *First working table in the first cycle*

Vac-cine	Dose	n	r	x	p	Y	Φ	Z	y	w	wx	wy	wx^2	wy^2	wxy
S	1.0	12	0	0.000	0.000	0	0.5	0.399	−1.253	7.64	0.00	−9.57	0.00	12.00	0.00
	1.6	12	3	0.470	0.250	0	0.5	0.399	−0.627	7.64	3.59	−4.79	1.69	3.00	−2.25
	2.5	12	6	0.916	0.500	0	0.5	0.399	0.000	7.64	7.00	0.00	6.41	0.00	0.00
	4.0	11	10	1.386	0.909	0	0.5	0.399	1.025	7.00	9.71	7.18	13.46	7.36	9.95
T	1.0	11	0	0.000	0.000	0	0.5	0.399	−1.253	7.00	0.00	−8.78	0.00	11.00	0.00
	1.6	12	4	0.470	0.333	0	0.5	0.399	−0.418	7.64	3.59	−3.19	1.69	1.33	−1.50
	2.5	11	8	0.916	0.727	0	0.5	0.399	0.570	7.00	6.42	3.99	5.88	2.27	3.66
	4.0	11	10	1.386	0.909	0	0.5	0.399	1.025	7.00	9.71	7.18	13.46	7.36	9.95

Table 5.3.1.-III. — *Second working table in the first cycle*

Vaccine	Σw	Σwx	Σwy	Σwx^2	Σwy^2	Σwxy	S_{xx}	S_{xy}	S_{yy}	$\overline{x}$	$\overline{y}$	a
S	29.92	20.30	−7.18	21.56	22.36	7.70	7.79	12.58	20.64	0.68	−0.24	−1.36
T	28.65	19.72	−0.80	21.03	21.97	12.11	7.46	12.66	21.95	0.69	−0.03	−1.17

Table 5.3.1.-IV. — *First working table in the second cycle*

Vac-cine	Dose	n	r	x	p	Y	Φ	Z	y	w	wx	wy	wx^2	wy^2	wxy
S	1.0	12	0	0.000	0.000	−1.36	0.086	0.158	−1.911	3.77	0.00	−7.21	0.00	13.79	0.00
	1.6	12	3	0.470	0.250	−0.58	0.279	0.336	−0.672	6.74	3.17	−4.53	1.49	3.04	−2.13
	2.5	12	6	0.916	0.500	0.15	0.561	0.394	−0.001	7.57	6.94	−0.01	6.36	0.00	−0.01
	4.0	11	10	1.386	0.909	0.93	0.824	0.258	1.260	5.07	7.03	6.39	9.75	8.05	8.86
T	1.0	11	0	0.000	0.000	−1.17	0.122	0.202	−1.769	4.20	0.00	−7.43	0.00	13.14	0.00
	1.6	12	4	0.470	0.333	−0.39	0.349	0.370	−0.430	7.23	3.40	−3.11	1.60	1.34	−1.46
	2.5	11	8	0.916	0.727	0.35	0.637	0.375	0.591	6.70	6.14	3.96	5.62	2.34	3.63
	4.0	11	10	1.386	0.909	1.13	0.870	0.211	1.311	4.35	6.03	5.70	8.36	7.48	7.90

Table 5.3.1.-V. — *Second working table after sufficient cycles*

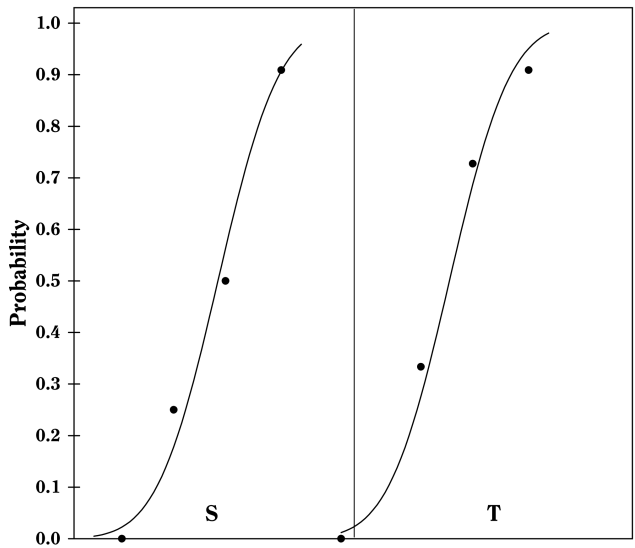| Vaccine	Σw	Σwx	Σwy	Σwx^2	Σwy^2	Σwxy	S_{xx}	S_{xy}	S_{yy}	$\overline{x}$	$\overline{y}$	a
S	18.37	14.80	−2.14	14.85	17.81	5.28	2.93	7.00	17.56	0.81	−0.12	−2.05
T	17.96	12.64	−0.55	11.86	18.35	6.76	2.96	7.15	18.34	0.70	−0.03	−1.72

Figure 5.3.1.-I.

5.3.2. LOGIT ANALYSIS AND OTHER TYPES OF ANALYSES OF A TEST PREPARATION AGAINST A REFERENCE

Results will be given for the situation where the logit method and other "classical" methods of this family are applied to the data in Section 5.3.1. This should be regarded as an exercise rather than an alternative to the probit method in this specific case. Another shape of the curve may be adopted only if this is supported by experimental or theoretical evidence.

Table 5.3.2.-I. — *Results by using alternative curves*

	Logit	Gompit	Angle[(*)]
Φ	$\dfrac{1}{1 + e^{-Y}}$	$1 - e^{-e^Y}$	$\dfrac{1}{2}\sin\ Y + \dfrac{1}{2}$
Z	$\dfrac{e^{-Y}}{(1 + e^{-Y})^2}$	$e^{Y - e^Y}$	$\dfrac{1}{2}\cos\ Y$
slope b	4.101	2.590	1.717
χ^2 lin	2.15	3.56	1.50
χ^2 par	0.0066	0.168	0.0010
Potency	162.9	158.3	155.8
Lower limit	121.1	118.7	122.6
Upper limit	221.1	213.3	200.7

$(*)\begin{cases} \text{If } Y < -\dfrac{1}{2}\pi \text{ then } \Phi = 0 \text{ and } Z = 0 \\ \text{If } Y > \dfrac{1}{2}\pi \text{ then } \Phi = 1 \text{ and } Z = 0 \end{cases}$

5.3.3. THE ED_{50} DETERMINATION OF A SUBSTANCE USING THE PROBIT METHOD

An in-vitro assay of oral poliomyelitis vaccine

In an ED_{50} assay of oral poliomyelitis vaccine with 10 different dilutions in 8 replicates of 50 µl on an ELISA-plate, results were obtained as shown in Table 5.3.3.-I.

The observations are transferred to the first working table and the subsequent columns are computed as described in Section 4.2.1. Table 5.3.3.-II shows the first cycle of this procedure.

Table 5.3.3.-I. — *Dilutions (10^x µl of the undiluted vaccine)*

−3.5	−4.0	−4.5	−5.0	−5.5	−6.0	−6.5	−7.0	−7.5	−8.0
+	+	+	+	−	−	−	−	−	−
+	+	+	+	−	−	−	−	−	−
+	+	−	−	−	−	−	−	−	−
+	+	+	+	−	−	−	−	−	−
+	+	+	−	−	−	−	−	−	−
+	+	+	+	+	−	−	−	−	−
+	+	+	+	+	−	+	−	−	−
+	+	+	+	−	+	−	−	−	−

The sums of the last 6 columns are calculated and transferred to the second working table (see Table 5.3.3.-III). The results in the other columns are found with formulae 4.2.1.-4 to 4.2.1.-10. This yields a common slope b of -0.295.

The values for Y in the first working table are now replaced by $a + bx$ and a second cycle is carried out. The cycle is repeated until the difference between 2 consecutive cycles has become small. The second working table should then appear as shown in Table 5.3.3.-IV.

Linearity is tested as described in Section 4.2.2. The χ^2-value with 8 degrees of freedom is 2.711 representing a *p*-value of 0.951 which is not significant.

The potency ratio can now be estimated as described in Section 4.5.
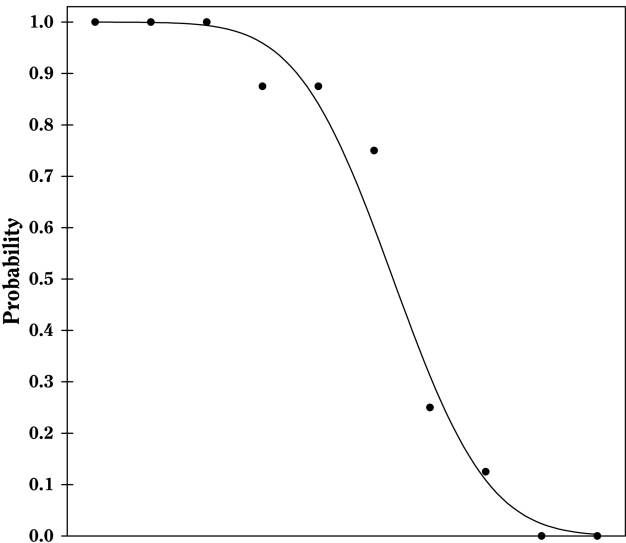
Figure 5.3.3.-I.

$$M'_T = \frac{-(-7.931)}{-0.646} = -12.273$$

Further:

$$C = \frac{(-0.646)^2 \times 55.883}{(-0.646)^2 \times 55.883 - 1^2 \times 1.960^2} = 1.197$$

$$V = \frac{1}{19.39} = 0.052$$

So ln confidence limits are:

$$-14.692 - (-2.420) \pm \sqrt{0.197 \times (2.882 + 1.197 \times 0.009^2)}$$

$$= -12.272 \pm 0.754$$

This estimate is still expressed in terms of the ln(dilutions). In order to obtain estimates expressed in ln(ED_{50})/ml the values are transformed to $-M'_T + \ln\left(\dfrac{1000}{50}\right)$.

5. General texts

Since it has become common use to express the potency of this type of vaccine in terms of $\log_{10}(ED_{50})$/ml, the results have to be divided by $\ln(10)$. The potency is thus estimated as 6.63 $\log_{10}(ED_{50})$/ml with 95 per cent confidence limits from 6.30 to 6.96 $\log_{10}(ED_{50})$/ml.

5.4. EXTENDED SIGMOID DOSE-RESPONSE CURVES

5.4.1. FOUR-PARAMETER LOGISTIC CURVE ANALYSIS
A serological assay of tetanus sera

As already stated in Section 3.4, this example is intended to illustrate a "possible" way to analyse the data presented, but not necessarily to reflect the "only" or the "most appropriate" way. Many other approaches can be found in the literature, but in most cases they should not yield dramatically different outcomes. A short discussion of alternative approaches and other statistical considerations is given in Section 7.5.

A guinea-pig antiserum is assayed against a standard serum (0.4 IU/ml) using an enzyme-linked immunosorbent assay technique (ELISA). 10 two-fold dilutions of each serum were applied on a 96-well ELISA plate. Each dilution was applied twice. The observed responses are listed in Table 5.4.1.-I.

For this example, it will be assumed that the laboratory has validated conditions 1 to 3 in Section 3.1.1 when the assay was being developed for routine use. In addition, the laboratory has validated that the upper limit and lower limit of the samples can be assumed to be equal.

No unusual features are discovered in a graphical representation. A least squares method of a suitable computer program is used to fit the parameters of the logistic function, assuming that the residual error terms are independent and identically distributed normal random variables. In this case, 3 parameters (α, β and δ) are needed to describe the common slope-factor and the common lower and upper asymptotes. 2 additional parameters (γ_S and γ_T) are needed to describe the horizontal location of the 2 curves.

The following estimates of the parameters are returned by the program:

$$\alpha = 3.196 \qquad \gamma_S = -4.307$$
$$\beta = 1.125 \qquad \gamma_T = -4.684$$
$$\delta = 0.145$$

In addition, the estimated residual variance (s^2) is returned as 0.001429 with 20 degrees of freedom (within-treatments variation).

In order to obtain confidence limits, and also to check for parallelism and linearity, the observed responses (u) are linearised and submitted to a weighted parallel-line analysis by the program. This procedure is very similar to that described in Section 4.2 for probit analysis with the following modifications:

$$Y = \beta(x - \gamma) \qquad y = Y + \frac{\left(\dfrac{u-\delta}{\alpha-\delta}\right) - \Phi}{Z}$$

$$\Phi = \frac{1}{1 + e^{-Y}} \qquad w = \frac{Z^2(\alpha-\delta)^2}{s^2}$$

$$Z = \frac{e^{-Y}}{(1 + e^{-Y})^2}$$

The resulting weighted analysis of variance of the transformed responses (y) using weights (w) is as follows:

There are no significant deviations from parallelism and linearity and thus the assay is satisfactory for potency calculations. If the condition of equal upper and lower asymptotes is not fulfilled, significant deviations from linearity and/or parallelism are likely to occur because the tests for linearity and parallelism reflect the goodness of fit of the complete four-parameter model. The residual error in the analysis of variance is always equal to 1 as a result of the transformation. However, a heterogeneity factor (analogous to that for the probit model) can be computed.

Table 5.3.3.-II. — *First working table in the first cycle*

Vaccine	Dose	n	r	x	p	Y	Φ	Z	y	w	wx	wy	wx^2	wy^2	wxy
T	$10^{-3.5}$	8	0	-8.06	0.000	0.00	0.5	0.399	-1.253	5.09	-41.04	-6.38	330.8	8.00	51.4
	$10^{-4.0}$	8	0	-9.21	0.000	0.00	0.5	0.399	-1.253	5.09	-46.91	-6.38	432.0	8.00	58.8
	$10^{-4.5}$	8	1	-10.36	0.125	0.00	0.5	0.399	-0.940	5.09	-52.77	-4.79	546.8	4.50	49.6
	$10^{-5.0}$	8	2	-11.51	0.250	0.00	0.5	0.399	-0.627	5.09	-58.63	-3.19	675.1	2.00	36.7
	$10^{-5.5}$	8	6	-12.66	0.750	0.00	0.5	0.399	0.627	5.09	-64.50	3.19	816.8	2.00	-40.4
	$10^{-6.0}$	8	7	-13.82	0.875	0.00	0.5	0.399	0.940	5.09	-70.36	4.79	972.1	4.50	-66.1
	$10^{-6.5}$	8	7	-14.97	0.875	0.00	0.5	0.399	0.940	5.09	-76.23	4.79	1140.8	4.50	-71.7
	$10^{-7.0}$	8	8	-16.12	1.000	0.00	0.5	0.399	1.253	5.09	-82.09	6.38	1323.1	8.00	-102.9
	$10^{-7.5}$	8	8	-17.27	1.000	0.00	0.5	0.399	1.253	5.09	-87.95	6.38	1518.9	8.00	-110.2
	$10^{-8.0}$	8	8	-18.42	1.000	0.00	0.5	0.399	1.253	5.09	-93.82	6.38	1728.2	8.00	-117.6

Table 5.3.3.-III. — *Second working table in the first cycle*

Vaccine	Σw	Σwx	Σwy	Σwx^2	Σwy^2	Σwxy	S_{xx}	S_{xy}	S_{yy}	$\overline{x}$	$\overline{y}$	a
T	50.93	-674.3	11.17	9484.6	57.50	-312.32	556.92	-164.43	55.05	-13.24	0.219	-3.690

Table 5.3.3.-IV. — *Second working table after sufficient cycles*

Vaccine	Σw	Σwx	Σwy	Σwx^2	Σwy^2	Σwxy	S_{xx}	S_{xy}	S_{yy}	$\overline{x}$	$\overline{y}$	a
T	19.39	-238.2	0.11	2981.1	26.05	-37.45	55.88	-36.11	26.05	-12.28	0.006	-7.931

The relative potency of the test preparation can be obtained as the antilogarithm of $\gamma_S - \gamma_T$. Multiplying by the assigned potency of the standard yields an estimate of $1.459 \times 0.4 = 0.584$ IU/ml. Formula 4.2.3.-2 gives 95 per cent confidence limits from 0.557-0.612 IU/ml.

Table 5.4.1.-I. — *Observed responses*

Standard *S*			Preparation to be examined *T*		
Dil.	Obs. 1	Obs. 2	Dil.	Obs. 1	Obs. 2
1/10	2.912	2.917	1/10	3.017	2.987
1/20	2.579	2.654	1/20	2.801	2.808
1/40	2.130	2.212	1/40	2.401	2.450
1/80	1.651	1.638	1/80	1.918	1.963
1/160	1.073	0.973	1/160	1.364	1.299
1/320	0.585	0.666	1/320	0.861	0.854
1/640	0.463	0.356	1/640	0.497	0.496
1/1280	0.266	0.234	1/1280	0.340	0.344
1/2560	0.228	0.197	1/2560	0.242	0.217
1/5120	0.176	0.215	1/5120	0.178	0.125

Table 5.4.1.-II — *Weighted analysis of variance*

Source of variation	Degrees of freedom	Chi-square	Probability
Preparations	1	0.529653	0.467
Regression	1	6599.51	0.000
Non-parallelism	1	0.0458738	0.830
Non-linearity	16	8.89337	0.918
Treatments	19	6608.98	0.000
Residual error	20	20.0000	
Total	39	6628.98	

6. COMBINATION OF ASSAY RESULTS

6.1. INTRODUCTION

Replication of independent assays and combination of their results is often needed to fulfil the requirements of the European Pharmacopoeia. The question then arises as to whether it is appropriate to combine the results of such assays and if so in what way.

2 assays may be regarded as mutually independent when the execution of either does not affect the probabilities of the possible outcomes of the other. This implies that the random errors in all essential factors influencing the result (for example, dilutions of the standard and of the preparation to be examined, the sensitivity of the biological indicator) in one assay must be independent of the corresponding random errors in the other one. Assays on successive days using the original and retained dilutions of the standard therefore are not independent assays.

There are several methods for combining the results of independent assays, the most theoretically acceptable being the most difficult to apply. 3 simple, approximate methods are described below; others may be used provided the necessary conditions are fulfilled.

Before potencies from assays based on the parallel-line or probit model are combined they must be expressed in logarithms; potencies derived from assays based on the slope-ratio model are used as such. As the former models are more common than those based on the slope-ratio model, the symbol *M* denoting ln potency is used in the formulae in this section; by reading *R* (slope-ratio) for *M*, the analyst may use the same formulae for potencies derived from assays based on the slope-ratio model. All estimates of potency must be corrected for the potency assigned to each preparation to be examined before they are combined.

6.2. WEIGHTED COMBINATION OF ASSAY RESULTS

This method can be used provided the following conditions are fulfilled:

1) the potency estimates are derived from independent assays;

2) for each assay *C* is close to 1 (say less than 1.1);

3) the number of degrees of freedom of the individual residual errors is not smaller than 6, but preferably larger than 15;

4) the individual potency estimates form a homogeneous set (see Section 6.2.2).

When these conditions are not fulfilled this method cannot be applied. The method described in Section 6.3 may then be used to obtain the best estimate of the mean potency to be adopted in further assays as an assumed potency.

6.2.1. CALCULATION OF WEIGHTING COEFFICIENTS

It is assumed that the results of each of the n' assays have been analysed to give n' values of *M* with associated confidence limits. For each assay the logarithmic confidence interval *L* is obtained by subtracting the lower limit from the upper. A weight *W* for each value of *M* is calculated from equation 6.2.1.-1, where *t* has the same value as that used in the calculation of confidence limits.

$$W = \frac{4t^2}{L^2} \tag{6.2.1.-1}$$

6.2.2. HOMOGENEITY OF POTENCY ESTIMATES

By squaring the deviation of each value of *M* from the weighted mean, multiplying by the appropriate weight and summing over all assays, a statistic is obtained which is approximately distributed as χ^2 (see Table 8.3) and which may be used to test the homogeneity of a set of ln potency estimates:

$$\chi^2 \approx \sum_{n'} W \left(M - \overline{M} \right)^2 \quad \text{where} \quad \overline{M} = \frac{\sum WM}{\sum W} \tag{6.2.2.-1}$$

If the calculated χ^2 is smaller than the tabulated value corresponding to $(n' - 1)$ degrees of freedom the potencies are homogeneous and the mean potency and limits obtained in Section 6.2.3 will be meaningful.

If the calculated value of this statistic is greater than the tabulated value, the potencies are heterogeneous. This means that the variation between individual estimates of *M* is greater than would have been predicted from the estimates of the confidence limits, i.e. that there is a significant variability between the assays. Under these circumstances condition 4 is not fulfilled and the equations in Section 6.2.3 are no longer applicable. Instead, the formulae in Section 6.2.4 may be used.

6.2.3. CALCULATION OF THE WEIGHTED MEAN AND CONFIDENCE LIMITS

The products *WM* are formed for each assay and their sum divided by the total weight for all assays to give the logarithm of the weighted mean potency.

$$\overline{M} = \frac{\sum WM}{\sum W} \tag{6.2.3.-1}$$

The standard error of the ln (mean potency) is taken to be the square root of the reciprocal of the total weight:

$$s_{\overline{M}} = \sqrt{\frac{1}{\sum W}} \qquad (6.2.3.\text{-}2)$$

and approximate confidence limits are obtained from the antilogarithms of the value given by

$$\overline{M} \pm t \times s_{\overline{M}} \qquad (6.2.3.\text{-}3)$$

where the number of degrees of freedom of t equals the sum of the number of degrees of freedom for the error mean squares in the individual assays.

6.2.4. WEIGHTED MEAN AND CONFIDENCE LIMITS BASED ON THE INTRA- AND INTER-ASSAY VARIATION

When results of several repeated assays are combined, the (χ^2-value may be significant. The observed variation is then considered to have two components:

— the intra-assay variation $s_M^2 = 1/W$,

— the inter-assay variation $s_{\overline{M}}^2 = \dfrac{\sum \left(M - \overline{M}\right)^2}{n'\left(n' - 1\right)}$

where $\overline{M}$ is the unweighted mean. The former varies from assay to assay whereas the latter is common to all M.

For each M a weighting coefficient is then calculated as:

$$W' = \frac{1}{s_M^2 + s_{\overline{M}}^2}$$

which replaces W in Section 6.2.3. where t is taken to be approximately 2.

6.3. UNWEIGHTED COMBINATION OF ASSAY RESULTS

To combine the n' estimates of M from n' assays in the simplest way, the mean is calculated and an estimate of its standard deviation is obtained by calculating:

$$s_{\overline{M}}^2 = \frac{\sum \left(M - \overline{M}\right)^2}{n'\left(n' - 1\right)} \qquad (6.3.\text{-}1)$$

and the limits are:

$$\overline{M} \pm t s_{\overline{M}} \qquad (6.3.\text{-}2)$$

where t has $(n' - 1)$ degrees of freedom. The number n' of estimates of M is usually small, and hence the value of t is quite large.

6.4. EXAMPLE OF A WEIGHTED MEAN POTENCY WITH CONFIDENCE LIMITS

Table 6.4.-I lists 6 independent potency estimates of the same preparation together with their 95 per cent confidence limits and the number of degrees of freedom of their error variances. Conditions 1, 2 and 3 in Section 6.2. are met. The ln potencies and the weights are calculated as described in Section 6.2.

Table 6.4.-I. — *Potency estimates and confidence intervals of 6 independent assays*

Potency estimate (I.U./vial)	Lower limit (I.U./vial)	Upper limit (I.U./vial)	Degrees of freedom	ln potency M	Weight W
18 367	17 755	19 002	20	9.8183	3777.7
18 003	17 415	18 610	20	9.7983	3951.5
18 064	17 319	18 838	20	9.8017	2462.5
17 832	17 253	18 429	20	9.7887	4003.0
18 635	17 959	19 339	20	9.8328	3175.6
18 269	17 722	18 834	20	9.8130	4699.5

Homogeneity of potency estimates is assessed with formula 6.2.2.-1 which gives a χ^2 of 4.42 with 5 degrees of freedom. This is not significant ($p = 0.49$) and thus all conditions are met.

A weighted mean potency is calculated with formula 6.2.3.-1 which yields 9.8085.

Formula 6.2.3.-2 gives a standard deviation of 0.00673 and approximate 95 per cent confidence limits of 9.7951 and 9.8218 are calculated with formula 6.2.3.-3 where t has 120 degrees of freedom.

By taking the antilogarithms a potency of 18 187 IU/vial is found with 95 per cent confidence limits from 17 946-18 431 IU/vial.

7. BEYOND THIS ANNEX

It is impossible to give a comprehensive treatise of statistical methods in a pharmacopoeial text. However, the methods described in this annex should suffice for most pharmacopoeial purposes. This section tries to give a more abstract survey of alternative or more general methods that have been developed. The interested reader is encouraged to further explore the existing literature in this area. The use of more specialised statistical methods should, in any case, be left to qualified personnel.

7.1. GENERAL LINEAR MODELS

The methods given in this annex can be described in terms of general linear models (or generalised linear models to include the probit and logit methods). The principle is to define a linear structure matrix X (or design matrix) in which each row represents an observation and each column a linear effect (preparation, block, column, dose). For example: the Latin square design in example 5.1.2 would involve a matrix with 36 rows and 13 columns. 1 column for each of the preparations, 1 column for the doses, 5 columns for each block except the first, and 5 columns for each row except the first. All columns, except the one for doses, are filled with 0 or 1 depending on whether or not the observation relates to the effect. A vector Y is filled with the (transformed) observations. The effects are estimated with the formula $(X'X)^{-1}X'Y$ from which the potency estimate m can easily be derived as a ratio of relevant effects. Confidence intervals are calculated from Fieller's theorem:

$$m_L, m_U = \frac{\left[m - \frac{g v_{12}}{v_{22}} \pm \frac{ts}{b}\sqrt{v_{11} - 2mv_{12} + m^2 v_{22} - g\left(v_{11} - \frac{v_{12}^2}{v_{22}}\right)} \right]}{(1 - g)}$$

where $g = \dfrac{t^2 s^2 v_{22}}{b^2}$

See the information section on general monographs (cover pages)

and v_{11}, v_{22}, v_{12} represent the variance multipliers for the numerator, the denominator and their covariance multiplier respectively. These are taken directly from $(X'X)^{-1}$ or indirectly by noting that:

$$\text{Var}(a_1 - a_2) = \text{Var}(a_1) + \text{Var}(a_2) - 2\text{Cov}(a_1, a_2)$$

$$\text{and Cov}(a_1 - a_2, b) = \text{Cov}(a_1, b) - \text{Cov}(a_2, b)$$

A full analysis of variance in which all components are partitioned is slightly more complicated as it involves a renewed definition of X with more columns to relax the assumptions of parallelism and linearity, after which the linear hypotheses can be tested. For assays depending upon quantal responses the linear effects (intercepts a_S, a_T etc. and the common slope b are found by maximising the sum over treatments of $n \ln \Phi(a_i + bx) + (n - r)\ln(1 - \Phi(a_i + bx))$ where x is the ln(dose), Φ denotes the shape of the distribution and $i \in \{S, T, ...\}$.

7.2. HETEROGENEITY OF VARIANCE

Heterogeneity of variance cannot always be solved by simply transforming the responses. A possible way to cope with this problem is to perform a weighted linear regression. In order to obtain an unbiased estimate, the weight of the observations is taken to be proportional to the reciprocal of the error variances. Since the true error variance is not always known, an iterative reweighted linear procedure may be followed. However, the calculation of the confidence interval involves new problems.

7.3. OUTLIERS AND ROBUST METHODS

The method of least squares described in this annex has the disadvantage of being very sensitive to outliers. A clear outlier may completely corrupt the calculations. This problem is often remedied by discarding the outlying result from the dataset. This policy can lead to arbitrary rejection of data and is not always without danger. It is not easy to give a general guideline on how to decide whether or not a specific observation is an outlier and it is for this reason that many robust methods have been developed. These methods are less sensitive to outliers because they give less weight to observations that are far away from the predicted value. New problems usually arise in computing confidence intervals or defining a satisfactory function to be minimised.

7.4. CORRELATED ERRORS

Absolute randomisation is not always feasible or very undesirable from a practical point of view. Thus, subsequent doses within a dilution series often exhibit correlated errors leading to confidence limits that are far too narrow. Some methods have been developed that take account of this autocorrelation effect.

7.5. EXTENDED NON-LINEAR DOSE-RESPONSE CURVES

Analysis of extended non-linear dose-response curves raises a number of statistical questions which require consideration, and for which professional advice is recommended. Some of these are indicated below.

1) An example using the four-parameter logistic function has been shown. However, models based on functions giving other sigmoid curves may also be used. Models incorporating additional asymmetry parameters have been suggested.

2) Heterogeneity of variance is common when responses cover a wide range. If the analysis ignores the heterogeneity, interpretation of results may not be correct and estimates may be biased. Use of the reciprocal of the error variances as weights is unlikely to be reliable with limited numbers of replicates. It may be appropriate to estimate a function which relates variance to mean response.

3) The statistical curve-fitting procedures may give different estimates depending on assumptions made about the homogeneity of the variance and on the range of responses used.

4) In principle, equality of upper and lower response limits for the different preparations included in an assay can be directly tested in each assay. However, interpretation of the results of these tests may not be straightforward. The tests for linearity and parallelism given by the simplified method of analysis (Example 5.4.1) indirectly incorporate tests for equality and accuracy of upper and lower limits.

5) Many assays include "controls" which are intended to identify the upper and/or lower response limits. However, these values may not be consistent with the statistically fitted upper and lower response limits based on the extended dose-response curve.

6) The simplified method of analysis given in Example 5.4.1 provides approximate confidence intervals. Other methods may also be used, for example intervals based on lack-of-fit of the completely specified model. For typical assay data, with responses covering the complete range for each preparation tested, all methods give similar results.

8. TABLES AND GENERATING PROCEDURES

The tables in this section list the critical values for the most frequently occurring numbers of degrees of freedom. If a critical value is not listed, reference should be made to more extensive tables. Many computer programs include statistical functions and their use is recommended instead of the tables in this section. Alternatively, the generating procedures given below each table can be used to compute the probability corresponding to a given statistic and number of degrees of freedom.

8.1. THE *F*-DISTRIBUTION

If an observed value is higher than the value in Table 8.1.-I, it is considered to be significant (upper lines, $p = 0.05$) and highly significant (lower lines, $p = 0.01$). $df1$ is the number of degrees of freedom of the numerator and $df2$ is the number of degrees of freedom of the denominator.

Generating procedure. Let F be the *F*-ratio and `df1` and `df2` as described above. Let `pi` = π = 3.14159265358979... The procedure in Table 8.1.-II will then generate the *p*-value.

8.2. THE *t*-DISTRIBUTION

If an observed value is higher than the value in Table 8.2.-I, it is considered to be significant ($p = 0.05$) and highly significant ($p = 0.01$).

Generating procedures. The *p*-value for a given *t* with df degrees of freedom can be found with the procedures in Section 8.1 where F = t^2, `df1` = 1 and `df2` = df.

The *t*-value ($p = 0.05$) for a given number of degrees of freedom df can be found with the procedure in Table 8.2.-II which should be accurate up to 6 decimal places.

Table 8.1.-I − *Critical values of the F-distribution*

df1 → df2 ↓	1	2	3	4	5	6	8	10	12	15	20	∞
10	4.965	4.103	3.708	3.478	3.326	3.217	3.072	2.978	2.913	2.845	2.774	2.538
	10.044	7.559	6.552	5.994	5.636	5.386	5.057	4.849	4.706	4.558	4.405	3.909
12	4.747	3.885	3.490	3.259	3.106	2.996	2.849	2.753	2.687	2.617	2.544	2.296
	9.330	6.927	5.953	5.412	5.064	4.821	4.499	4.296	4.155	4.010	3.858	3.361
15	4.543	3.682	3.287	3.056	2.901	2.790	2.641	2.544	2.475	2.403	2.328	2.066
	8.683	6.359	5.417	4.893	4.556	4.318	4.004	3.805	3.666	3.522	3.372	2.868
20	4.351	3.493	3.098	2.866	2.711	2.599	2.447	2.348	2.278	2.203	2.124	1.843
	8.096	5.849	4.938	4.431	4.103	3.871	3.564	3.368	3.231	3.088	2.938	2.421
25	4.242	3.385	2.991	2.759	2.603	2.490	2.337	2.236	2.165	2.089	2.007	1.711
	7.770	5.568	4.675	4.177	3.855	3.627	3.324	3.129	2.993	2.850	2.699	2.169
30	4.171	3.316	2.922	2.690	2.534	2.421	2.266	2.165	2.092	2.015	1.932	1.622
	7.562	5.390	4.510	4.018	3.699	3.473	3.173	2.979	2.843	2.700	2.549	2.006
50	4.034	3.183	2.790	2.557	2.400	2.286	2.130	2.026	1.952	1.871	1.784	1.438
	7.171	5.057	4.199	3.720	3.408	3.186	2.890	2.698	2.563	2.419	2.265	1.683
∞	3.841	2.996	2.605	2.372	2.214	2.099	1.938	1.831	1.752	1.666	1.571	1.000
	6.635	4.605	3.782	3.319	3.017	2.802	2.511	2.321	2.185	2.039	1.878	1.000

Table 8.1.-II — *Generating procedure for the F-distribution*

If df1 is even	If df1 is odd and df2 is even	If df1 and df2 are odd
`x=df1/(df1+df2/F)`	`x=df2/(df2+df1*F)`	`x=atn(sqr(df1*F/df2))`
`s=1`	`s=1`	`cs=cos(x)`
`t=1`	`t=1`	`sn=sin(x)`
`for i=2 to (df1-2) step 2`	`for i=2 to (df2-2) step 2`	`x=x/2`
`t=t*x*(df2+i-2)/i`	`t=t*x*(df1+i-2)/i`	`s=0`
`s=s+t`	`s=s+t`	`t=sn*cs/2`
`next i`	`next i`	`v=0`
`p=s*(1-x)^(df2/2)`	`p=1-s*(1-x)^(df1/2)`	`w=1`
		`for i=2 to (df2-1) step 2`
		`s=s+t`
		`t=t*i/(i+1)*cs*cs`
		`next i`
		`for i=1 to (df1-2) step 2`
		`v=v+w`
		`w=w*(df2+i)/(i+2)*sn*sn`
		`next i`
		`p=1+(t*df2*v-x-s)/pi*4`

Table 8.2.-I — *Critical values of the t-distribution*

df	$p = 0.05$	$p = 0.01$	df	$p = 0.05$	$p = 0.01$
1	12.706	63.656	22	2.074	2.819
2	4.303	9.925	24	2.064	2.797
3	3.182	5.841	26	2.056	2.779
4	2.776	4.604	28	2.048	2.763
5	2.571	4.032	30	2.042	2.750
6	2.447	3.707	35	2.030	2.724
7	2.365	3.499	40	2.021	2.704
8	2.306	3.355	45	2.014	2.690
9	2.262	3.250	50	2.009	2.678
10	2.228	3.169	60	2.000	2.660
12	2.179	3.055	70	1.994	2.648
14	2.145	2.977	80	1.990	2.639
16	2.120	2.921	90	1.987	2.632
18	2.101	2.878	100	1.984	2.626
20	2.086	2.845	∞	1.960	2.576

Table 8.2.-II — *Generating procedure for the t-distribution*

```
t =  1.959964+
     2.37228/df+
     2.82202/df^2+
     2.56449/df^3+
     1.51956/df^4+
     1.02579/df^5+
     0.44210/df^7
```

8.3. THE χ^2-DISTRIBUTION

Table 8.3.-I — *Critical values of the χ^2-distribution*

df	$p = 0.05$	$p = 0.01$	df	$p = 0.05$	$p = 0.01$
1	3.841	6.635	11	19.675	24.725
2	5.991	9.210	12	21.026	26.217
3	7.815	11.345	13	22.362	27.688
4	9.488	13.277	14	23.685	29.141
5	11.070	15.086	15	24.996	30.578
6	12.592	16.812	16	26.296	32.000
7	14.067	18.475	20	31.410	37.566
8	15.507	20.090	25	37.652	44.314
9	16.919	21.666	30	43.773	50.892
10	18.307	23.209	40	55.758	63.691

If an observed value is higher than the value in Table 8.3.-I, it is considered to be significant ($p = 0.05$) or highly significant ($p = 0.01$).

Generating procedure. Let `X2` be the χ^2-value and `df` as described above. The procedure in Table 8.3.-II will then generate the *p*-value.

Table 8.3.-II — *Generating procedure for the χ^2-distribution*

If df is even	If df is odd
`s=0`	`x=sqr(x2)`
`t=exp(-x2/2)`	`s=0`
`for i=2 to df step 2`	`t=x*exp(-x2/2)/sqr(pi/2)`
`s=s+t`	`for i=3 to df step 2`
`t=t*x2/i`	`s=s+t`
`next i`	`t=t*x2/i`
`p=1-s`	`next i`
	`p=1-s-2*phi(x)`

In this procedure `phi` is the cumulative standard normal distribution function Φ (see Section 8.4).

8.4. THE Φ-DISTRIBUTION (THE CUMULATIVE STANDARD NORMAL DISTRIBUTION)

Table 8.4.-I — *Values of the Φ-distribution*

x	Φ	x	Φ	x	Φ
0.00	0.500	1.00	0.841	2.00	0.977
0.05	0.520	1.05	0.853	2.05	0.980
0.10	0.540	1.10	0.864	2.10	0.982
0.15	0.560	1.15	0.875	2.15	0.984
0.20	0.579	1.20	0.885	2.20	0.986
0.25	0.599	1.25	0.894	2.25	0.988
0.30	0.618	1.30	0.903	2.30	0.989
0.35	0.637	1.35	0.911	2.35	0.991
0.40	0.655	1.40	0.919	2.40	0.992
0.45	0.674	1.45	0.926	2.45	0.993
0.50	0.691	1.50	0.933	2.50	0.994
0.55	0.709	1.55	0.939	2.55	0.995
0.60	0.726	1.60	0.945	2.60	0.995
0.65	0.742	1.65	0.951	2.65	0.996
0.70	0.758	1.70	0.955	2.70	0.997
0.75	0.773	1.75	0.960	2.75	0.997
0.80	0.788	1.80	0.964	2.80	0.997
0.85	0.802	1.85	0.968	2.85	0.998
0.90	0.816	1.90	0.971	2.90	0.998
0.95	0.829	1.95	0.974	2.95	0.998

The Φ-value for negative *x* is found from table 8.4.-I as $1 - \Phi(-x)$.

Generating procedure: Let `x` be the *x*-value. The procedure in Table 8.4.-II will generate the corresponding Φ-value if $0 \le x \le 8.15$. If *x* is greater than 8.15 the Φ-value can be set to 1. If *x* is negative, the formula given above can be used.

This procedure assumes that the computer can represent about 15 decimal places. If less digits or more digits can be represented, the procedure needs some trivial modifications.

Table 8.4.-II — *Generating procedure for the Φ-distribution*

```
s=0
t=x
i=1
repeat
s=s+t
i=i+2
t=t*x*x/i
until t<1E-16
phi=0.5+s*exp(-x*x/2)/sqr(2*pi)
```

8.5. RANDOM PERMUTATIONS

Random permutations are needed in randomised block designs. The following algorithm shows how the built-in random generator of a computer can be used to create random permutations of *N* treatments.

Step 1. Write the *N* possible treatments down in a row.

Step 2. Obtain a random integer *r* such that $1 \le r \le N$.

Step 3. Exchange the *r*-th treatment with the *N*-th treatment in the row.

Step 4. Let $N = N - 1$ and repeat steps 2 to 4 until $N = 1$.

An example with 6 treatments will illustrate this algorithm.

1.	$N = 6$	S_1	S_2	S_3	T_1	T_2	T_3
2.	$r = 2$		$\rightarrow$				$\leftarrow$
3.		S_1	T_3	S_3	T_1	T_2	S_2
4.	$N = 5$						
2.	$r = 4$				$\rightarrow$	$\leftarrow$	
3.		S_1	T_3	S_3	T_2	T_1	S_2
4.	$N = 4$						
2.	$r = 4$				$\downarrow$		
3.		S_1	T_3	S_3	T_2	T_1	S_2
4.	$N = 3$						
2.	$r = 1$	$\rightarrow$		$\leftarrow$			
3.		S_3	T_3	S_1	T_2	T_1	S_2
4.	$N = 2$						
2.	$r = 1$	$\rightarrow$	$\leftarrow$				
3.		T_3	S_3	S_1	T_2	T_1	S_2
4.	$N = 1$						

8.6. LATIN SQUARES

The following example shows how 3 independent permutations can be used to obtain a Latin square.

1). Generate a random permutation of the *N* possible treatments (see Section 8.5):

T_3	S_3	S_1	T_2	T_1	S_2

2). A simple Latin square can now be constructed by "rotating" this permutation to the right. This can be done as follows. Write the permutation found in step 1 down on the first row. The second row consists of the same permutation, but with all treatments shifted to the right. The rightmost

treatment is put on the empty place at the left. This is repeated for all the rows until all the treatments appear once in each column:

T_3	S_3	S_1	T_2	T_1	S_2
S_2	T_3	S_3	S_1	T_2	T_1
T_1	S_2	T_3	S_3	S_1	T_2
T_2	T_1	S_2	T_3	S_3	S_1
S_1	T_2	T_1	S_2	T_3	S_3
S_3	S_1	T_2	T_1	S_2	T_3

3). Generate 2 independent random permutations of the figures 1 to N:
— one for the rows:

2	3	6	1	4	5

— and one for the columns:

3	4	6	2	5	1

4). The Latin square can now be found by sorting the rows and columns of the simple Latin square according to the 2 permutations for the rows and columns:

	3	4	6	2	5	1
2	T_3	S_3	S_1	T_2	T_1	S_2
3	S_2	T_3	S_3	S_1	T_2	T_1
6	T_1	S_2	T_3	S_3	S_1	T_2
1	T_2	T_1	S_2	T_3	S_3	S_1
4	S_1	T_2	T_1	S_2	T_3	S_3
5	S_3	S_1	T_2	T_1	S_2	T_3

$\downarrow$

	1	2	3	4	5	6
1	S_1	T_3	T_2	T_1	S_3	S_2
2	S_2	T_2	T_3	S_3	T_1	S_1
3	T_1	S_1	S_2	T_3	T_2	S_3
4	S_3	S_2	S_1	T_2	T_3	T_1
5	T_3	T_1	S_3	S_1	S_2	T_2
6	T_2	S_3	T_1	S_2	S_1	T_3

9. GLOSSARY OF SYMBOLS

Symbol	Definition
a	Intersection of linear regression of responses on dose or ln(dose)
b	Slope of linear regression of responses on dose or on ln(dose)
d	Number of dose levels for each preparation (excluding the blank in slope-ratio assays)
e	Base of natural logarithms (= 2.71828182845905...)
g	Statistic used in Fieller's theorem: $g = \dfrac{C-1}{C}$
h	Number of preparations in an assay, including the standard preparation
m	Potency estimate obtained as a ratio of effects in general linear models
n	Number of replicates for each treatment
p	Probability of a given statistic being larger than the observed value. Also used as the ratio r/n in probit analysis
r	The number of responding units per treatment group in assays depending upon quantal responses
s	Estimate of standard deviation ($= \sqrt{s^2}$)
s^2	Estimate of residual variance given by error mean square in analysis of variance
t	Student's statistic (Table 8.2.)
u	Observed response in four-parameter analysis
v_{11}, v_{12}, v_{22}	(Co)variance multipliers for numerator and denominator of ratio m in Fieller's theorem
w	Weighting coefficient
x	The ln(dose)
y	Individual response or transformed response
A	Assumed potencies of test preparations when making up doses
B	Mean response to blanks in slope-ratio assays
C	Statistic used in the calculation of confidence intervals: $C = \dfrac{1}{1-g}$
$C_1, ..., C_n$	Mean response to each column of a Latin square design
D_1, D_2	Mean response on time 1 or time 2 in the twin cross-over design
F	Ratio of 2 independent estimates of variance following an F-distribution (Table 8.1.)
$G_S, G_T, ...$	Treatment values used in the analysis of variance for slope-ratio assays
H_P, H_L	Multipliers used in the analysis of variance for parallel-line assays
H_B, H_I	Multipliers used in the analysis of variance for slope-ratio assays

Symbol	Definition
I	In parallel-line assays, the ln of the ratio between adjacent doses. In slope-ratio assays, the interval between adjacent doses
$J_S, J_T ...$	Linearity values used in the analysis of variance for slope-ratio assays
K	Correction term used in the calculation of sums of squares in the analysis of variance
L	Width of confidence interval in logarithms
$L_S, L_T ...$	Linear contrasts of standard and test preparations
M'	ln potency ratio of a given test preparation
N	Total number of treatments in the assay ($= dh$)
$P_S, P_T ...$	Sum of standard and test preparations
R	Estimated potency of a given test preparation
R'	Potency ratio of a given test preparation
$R_1 ... , R_n$	Mean response in each of rows 1 to n of a Latin square design, or in each block of a randomised block design
S	Standard preparation
$S_1, ... , S_d$	Mean response to the lowest dose 1 up to the highest dose d of the standard preparation S
SS	Sum of squares due to a given source of variation
$T, U, V, ...$	Test preparations
$T_1, ... , T_d$	Mean response to the lowest dose 1 up to the highest dose d of test preparation T
V	Variance coefficient in the calculation of confidence limits
W	Weighting factor in combination of assay results
X	Linear structure or design matrix used in general linear models

Symbol	Definition
Y	Vector representing the (transformed) responses in general linear models
Z	The first derivative of Φ
α	Upper asymptote of the ln(dose)-response curve in four-parameter analysis
β	Slope-factor of the ln(dose)-response curve in four-parameter analysis
γ	The ln(dose) giving 50 per cent response in the four-parameter analysis
δ	Lower asymptote of the ln(dose)-response curve in four-parameter analysis
π	3.141592653589793238...
Φ	Cumulative standard normal distribution function (Table 8.4.)
χ^2	Chi-square statistic (Table 8.3.)

10. LITERATURE

This section lists some recommended literature for further study.

Finney, D.J. (1971). *Probit Analysis*, 3rd Ed. Cambridge University Press, Cambridge.

Nelder, J.A. & Wedderburn, R.W.M. (1972). Generalized linear models, *Journal of the Royal Statistical Society, Series* A 135, 370-384.

DeLean, A., Munson, P.J., and Rodbard, D. (1978). *Simultaneous analysis of families of sigmoidal curves: Application to bioassay, radioligand assay, and physiological dose-response curves*, Am. J. Physiol. 235(2): E97-E102.

Finney, D.J. (1978). *Statistical Method in Biological Assay*, 3rd Ed. Griffin, London.

Sokal, R.R. & Rohlf, F.R. (1981). *Biometry: Principles and Practice of Statistics in Biological Research*, 2nd Ed. W.H. Freemann & CO, New York.

Peace, K.E. (1988). *Biopharmaceutical Statistics for Drug Development*, Marcel Dekker Inc., New York/Basel.

Bowerman, B.L. & O'Connell, R.T. (1990). *Linear Statistical Models an Applied Approach*, 2nd Ed. PWS-KENT Publishing Company, Boston.

Govindarajulu, Z. (2001). *Statistical Techniques in Bioassay*, 2nd revised and enlarged edition, Karger, New York.