Journal of Pharmaceutical & Biomedical Analysis
Vol. 6, No. 3, pp. 317–319, 1988
Printed in Great Britain

0731–7085/88 $3.00 + 0.00

Short Communication

A basic program to determine outliers in biopharmaceutical analysis using the Nair criterion

GRAHAM S. CLARKE* and MALCOLM L. ROBINSON

E. R. Squibb and Sons, Reeds Lane, Moreton, Wirral, Merseyside, L46 1QW, UK

Keywords: *Outliers; Nair criterion; biopharmaceutical analysis.*

Introduction

In bioanalysis, as in other areas of analysis, difficulties can arise in judging whether to accept or reject individual values from a set of replicate analytical results. There appear to be no definitive procedures described in the literature which are applicable to bioanalytical results; operations systems of individual laboratories range from the purely subjective to sophisticated statistical techniques.

In the authors' laboratory, the analysis of a sample can yield a set of triplicate results, two of which are close and the other an apparent outlier. If this is judged to be so, then the outlier is rejected and the mean of the two remaining values is quoted as the analytical result. If it is judged otherwise, the mean of the three results is reported. However, a problem arises when two operators judge the same data. A finite number of instances will occur where there is disagreement. To overcome this problem, a number of statistical tests were examined. Of these, the Nair criterion was found to be the test that most often gave a result agreeing with the authors' own judgement. However, since this is a time-consuming exercise, ways of reducing the time spent were evaluated.

The application described in the present communication has been written into a Basic program, use of which enables inexperienced investigators to make consistently objective data-handling decisions without the need to consult superiors or colleagues.

Experimental

An external estimate of the standard deviation for the method is required and this is provided by controls that are assayed together with the samples. At least 11 estimates are required for the control data to obtain a t value from the Extreme Studentized Deviate Table. The probability α that the analyst is willing to take of rejecting an observation

* To whom correspondence should be addressed.

that really belongs to the group is chosen. $t\alpha\ (n,\ v)$ is obtained from the Extreme Studentized Deviate Table where n is the number of observations in the sample and v is the number of degrees of freedom for S_v, the standard deviation for the control samples.

The algorithms used in the program are defined thus:

(i) If observations that are too large are considered rejectable, compute

$$t = (X - \bar{X})/Sv$$

(ii) If observations that are too small are considered rejectable, compute

$$t = (\bar{X} - X)/Sv$$

If t is larger than $t\alpha\ (n,\ v)$ the observation must be rejected and the mean of the remaining values quoted. If the observation cannot be rejected, all values must be included in the calculation of the mean.

Program

The program uses Digital Basic Plus and is used in the authors' laboratory on a PDP 11/70 minicomputer, although the program is readily compatible with many micro-computers.

The program requires input of the percentage recovery values for which the standard deviation is calculated. The t value is inputted from the Extreme Studentized Deviate Table, the replicate values of the observation inserted and the suspect observation identified. The program calculates the t value for the experimental data, makes a judgement on whether the suspect observation can be rejected and gives the resultant mean to be quoted. Results are stored in a named file for hardcopy storage.

Example

The program was used to evaluate data generated in a pharmacokinetic study of an experimental formulation. Serum samples taken at 0–6 h were assayed singly as shown in Table 1. Reference to the dosing regimen indicated that the observations at time 0 and 6 h were suspect and these samples were repeated in duplicate. The reassayed observations were tested using the Nair criterion program and, as demonstrated in Fig. 1, the apparent outlier was included in the zero-time observation but rejected in the 6-h observation. The reported values are shown in Table 1.

Discussion

When reviewing rejection criteria it becomes apparent that no completely satisfactory rule can be devised for every situation and that no available criterion is a substitute for the judgement of an experienced investigator [2]. Thus the choice of a rejection criterion must reflect the analyst's knowledge about the nature and frequency of errors. The advantage of using such a procedure is that it allows consistency in handling data in an area where subjectivity is often the rule; if chosen wisely, the procedure will reject only data which would otherwise be rejected by an experienced investigator. The Nair criterion has been chosen in this instance because it allows extreme observations in only one direction to be rejected; that is the situation which is often addressed in

Figure 1
Example of an output file from the Nair criterion program

Nair outlier data

File name: NADOL.001

Standard deviation = 4.8
t value from Natrella handbook = 1.82

Sample identity = 0 h
Input values = 91.5, 94, 100
t value for variables is 1.00694
Outlier cannot be rejected therefore mean value must be given i.e. 95.1667

Sample identity = 6 h
Input values = 387.5, 420.4, 439.1
t value for variables is 5.86805
Outlier can be rejected: mean of other 2 values = 429.75

Table 1
Summary of the treatment and final reported values for drug concentration

Time (h)	Original observation Drug concentration (ng ml^{-1})	Replicate observation Drug concentration (ng ml^{-1}) 1		Reported value Drug concentration (ng ml^{-1})
0	100*	94	91.5	95.2
1½	149	—	—	149
2	169	—	—	169
3	350	—	—	350
4	375	—	—	375
5	413	—	—	413
6	387.5*	420.4	439.1	429.8

* Suspect observation.

biopharmaceutical analysis. The use of the program enables easy use of the rejection criterion, gives the inexperienced operator confidence in reporting results and ensures that the same decisions will be made by all operators.

References

[1] M. G. Natrella, *Experimental Statistics, National Bureau of Standards Handbook* **91**, 17–5 (1963).
[2] E. B. Wilson, Jr, *An Introduction to Scientific Research*, pp. 256–258, McGraw-Hill Book Co., New York, NY, (1952).

[Received for review 27 March 1987; revised manuscript received 8 October 1987]