ACTA CHEMICA SCANDINAVICA 23 (1969) 3426–3434

Enzyme Activity Calculations by Curvilinear Regression Analysis. A Program for a Computing Desk Top Calculator, Used in the Study of Cytochrome c Reductase and Hyaluronidase

ESKIL HULTIN,[a] GISELA LILJEQVIST,[b] GUNNAR LUNDBLAD,[c] SVEN PALÉUS[b] and GUNNAR STÄHL[d]

Institutionen för Toxikologi, Statens Medicinska Forskningsråd, Karolinska Institutet, S-104 01 Stockholm, Sweden

Numerical calculations of reaction rates, particularly extrapolation of enzymic activities to zero time, are done with a curvilinear regression analysis program for a computing desk top calculator (Olivetti Programma). This program is written to include warnings against inadvertent use of misread data and inadvertent use of linear regression when curvilinear regression is needed. The curvilinearity of the regression is tested statistically, and the standard deviations of extrapolated enzyme activities are calculated. A coordinate transformation can be included, and this is exemplified with a modification enabling direct calculation of enzyme activities from viscosimetric measurements.

Semilogarithmic plots for the determination of cytochrome c reductase activity fit straight lines fairly well but not exactly, as shown with graphs from experiments and with theoretical deductions in two recent publications by some of the present authors.[1,2] The enzyme activities have been recalculated numerically by curvilinear regression analysis, and this article serves in part as an appendix to the papers mentioned.

It is easily discovered from a diagram if one of several plotted points deviates from the series of measurements so much that a misreading probably was done; the significance of the deviation may be tested in doubtful cases.[3]

[a] Department of Chemistry, University of Stockholm, Sweden.
[b] Biochemical Department, Nobel Medical Institute, Karolinska Institutet, S-104 01 Stockholm, Sweden.
[c] Department of Chemistry, Statens Bakteriologiska Laboratorium, S-105 21 Stockholm, Sweden.
[d] Olivetti Svenska AB, Box 40166, S-103 44 Stockholm, Sweden.

Likewise, it is easily noticed from a graph if the experimental data fit a curvilinear regression line unmistakably better then a straight line; if so, one may draw a tangent rather than a linear regression line as the basis for reading the result.

Circumstances are different when linear regression equations are calculated numerically. A significant deviation of the series of measurements from linearity, or of one of the points from the series, would not be obvious and therefore might be easily overlooked. Thus, uncritical calculation of linear regression equations may lead to inaccurate or incorrect results.

However, numeric regression analysis according to the method of least squares has one valuable advantage over graphic evaluations: standard deviations can be calculated. Furthermore, no matter who does the numerical calculation, the result will be the same. If a computing desk top calculator is used in the laboratory, the whole numeric calculation takes even less time than a graphic evaluation.

A computing desk top calculator is useful in the laboratory for solving relatively uncomplicated problems for it eliminates card punching as well as waiting periods for available time, two time-consuming factors often associated with use of a regular computer.

Because of this we have worked out a curvilinear regression program for a computing desk top calculator. It is obviously a matter of routine to write a corresponding program for a regular computer, but the rather limited memory size of a desk top computer causes difficulties in the programming, and this may be one reason why such programs, although very useful, have not been worked out previously, as far as we know.

One main purpose of our program is to warn against the inadvertent use of misread data and against the inadvertent use of linear regression when a curvilinear regression is necessary. Consequently, our program is written for the calculation of the inclination of the tangent to a parabola at zero abscissa value, and the standard deviation of this inclination.[3,4]

EXPERIMENTAL

Cytochrome c reductase. Spectrophotometric measurements of cytochrome c concentrations at various times in a reaction mixture comprising the Keilin and Hartree beef heart succinate-cytochrome c reductase preparation previously described were used.[1,2] (Correction to Ref. 1: the cyanide concentration was 7×10^{-3} mM, and the enzyme preparation concentration (Fig. 1) was 0.070 mg per ml.)

Viscosimetric determination of hyaluronidase activity. Measurements on a staphylococcus toxin, 12/63 No. 57, prepared and studied by Tirunarayanan and Lundblad[7] were used for the present work.

COMPUTER PROGRAMS AND CALCULATIONS

An Olivetti Programma 101 computing desk top calculator was used.[5,6] Programs are stored on magnetic cards. We have worked out a curvilinear regression analysis program for general use which is stored sequentially on 4 cards. Additionally, we have written accessory programs with which our main program may be occasionally supplemented for optional calculations. Further-

more, for the program part stored on the first card we have prepared a substitute which is intended for the calculation of enzymic activity from viscosimetric measurements and which includes a coordinate transformation; it is described in subsequent sections of this article. Finally, we have made an expanded version of the program part stored on the last card in the sequential series.

A print out of our programs is reproduced in Tables 1 and 2.

Card No. 1 contains the part of the program which is used when experimental data are introduced. The decimal point of the original measurements may thereby have to be shifted because of the limited capacity of the computer.

Card Nos. 2 and 3 are used for the subsequent internal computations.

Card No. 4 contains a program for a print out of the inclination of the tangent at zero abscissa value, and its standard deviation, the variance ratio F for the curvilinearity of the regression line, the standard deviation of the points of measurements about the curvilinear regression line, and finally, the enzyme activity and its standard deviation.

Table 1. A print-out of the program on Cards Nos. 1, 2, 3, and 4, and on Card 1 D.

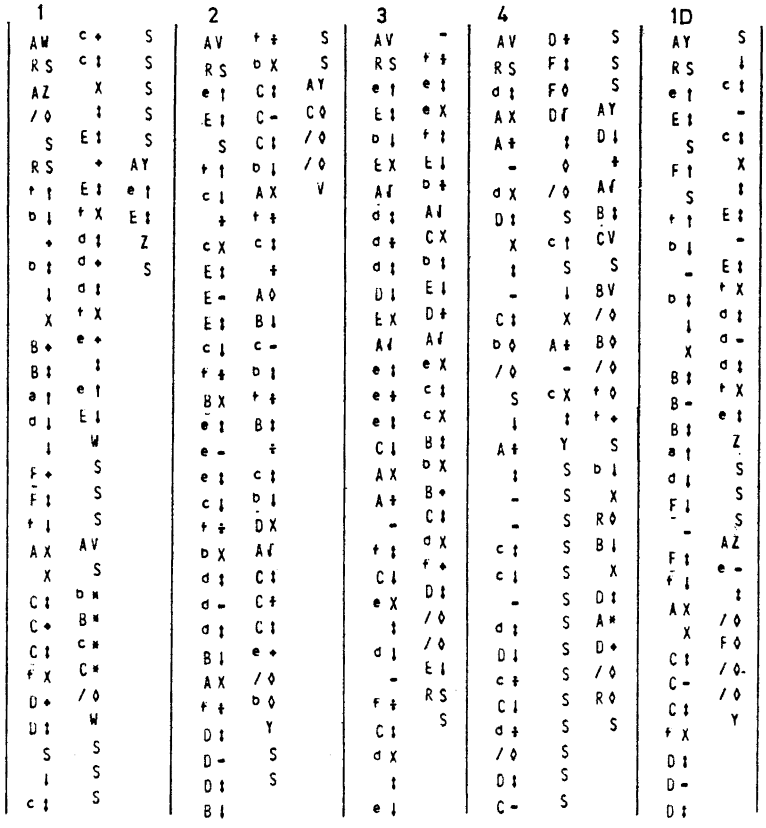
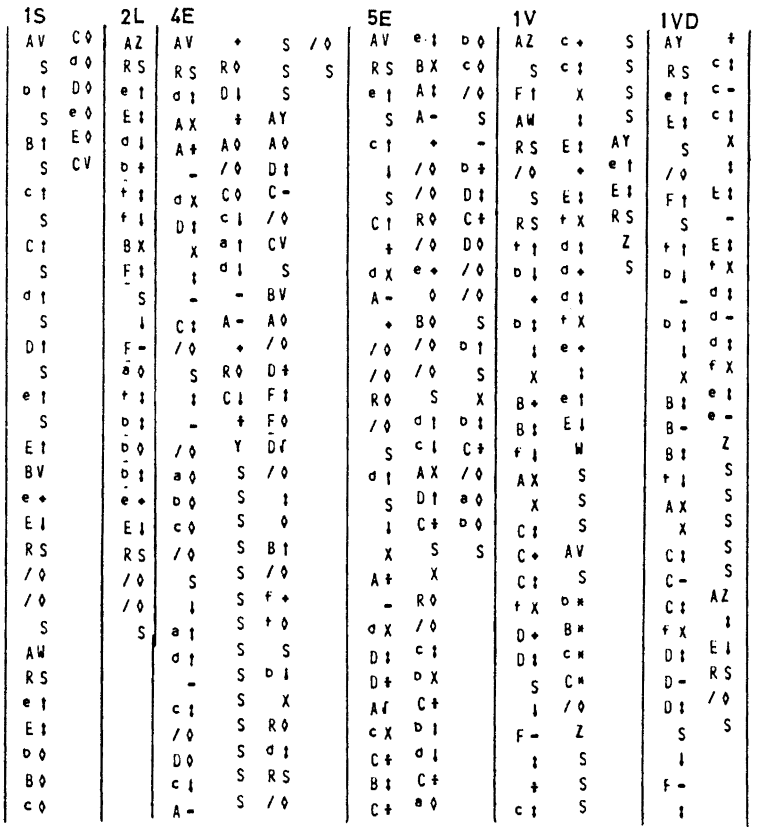


Table 2. A print-out of the program on Cards 1 S, 2 L, 4 E, 5 E, 1 V, and 1 VD.



Card 1 D (Delete) is used to delete one or several pairs of measurements which have been introduced with the program of Card No. 1 or 1 S.

Card No. 1 S (Sums) contains two accessory programs to be optionally used: one for a print out of the sums of squares *etc.* available as contents of the registers after the experimental data have been introduced with the program on the first card in the sequential series; and another program which reversely, may be used to introduce the sums; this card, if so used, substitutes Card No. 1 and enables a rapid reintroduction of the data.

Card 2 L (Linear) is optionally introduced after the use of Card No. 2 but before the use of Card No. 3 for a print out of the coefficients of the linear regression equation.

Card No. 4 E (Expanded) contains a program for a print out of the coefficients of the curvilinear regression equation, an analysis of variance regarding the curvilinearity of the regression equation, and the enzyme activity.

Table 3. An example of the calculation of cytochrome c reductase activity using the expanded program: the print-out (the tally paper trimmed, cut, and arranged in consecutive columns). Reference numbers are added for the following instructions for use of the program cards (letters denote the keys to be depressed), and for the following explanations of the print-out. 1) Card No. 1; V. 2) The code number of the experiment; ↑ 3) The factor by which the readings of the independent variable (time) was multiplied before they were entered; ↑ 4) The factor by which the values of the dependent variable were multiplied before they were entered; ↑ 5) Key-board clear key; S. 6) The time (multiplied by its factor); S. 7) The logarithm of the cytochrome c concentration* (multiplied by its factor); S; 6 and 7 are repeated until all experimental points have been introduced. 8) F ◇: the number of experimental points (the decimal wheel was temporarily set to 0). 9) Optionally: a print out of the contents of the registers. Card 1 S; W. 10) Card No. 2; V. 11) The number ref. 8; S. 12) A number ($\bar{y}$) to be reintroduced later if Card 2 L or 4 E is used. 13) and 14) Numbers to be reintroduced later. 15) Optionally: Card 2L; Z. 16) The number ref. 12; S. 17) and 18) The coefficients of the linear regression eqn. $\hat{Y}=a+bX$. 19) Card No. 3; V. 20) Card No. 4 E; V. 21) The number ref. 12; S. 22–24) The coefficients of the curvilinear regression eqn. $\hat{Y}=a+bX+cX^2$. 25) The number ref. 8; S. 26–33) The analysis of variance. 26–28) Linear regression: reduced sum of squares, degrees of freedom d. f., and variance. 29–31) Curvilinear regression: reduced sum of squares, d. f., and variance. 32) Curvilinearity: sums of squares (remainder). 33) The variance ratio. 34) The standard deviation about the curvilinear regression line. 35) A factor by which the inclination of the tangent of the extrapolated line at zero abscissa value (ref. 23) is multiplied to give the extrapolated enzyme activity (here the factor had not been programmed). 36) If the factor printed is to be used: S, if it must be changed: the proper factor is introduced; S. 37) A complete print-out of the extrapolated enzyme activity (the place of the decimal point may be erroneous; *cf.* ref. 41; N.B. the effect of the factors ref. 36, 39, and 40, and the decimal wheel setting). 38) Card No. 5E; V. 39) The number ref. 3; S. 40) The number ref. 4; S. 41) The result: the corrected enzyme activity. 42) The number ref. 13; S. 43) The number ref. 14; S. 44) The standard deviation of the enzyme activity. 45) The corrected standard deviation about the curvilinear regression line. 46) The standard deviation of the inclination of the tangent (*cf.* ref. 51). 47) and 48) The coefficients a and c of the curvilinear regression equation (refs. 22 and 24). 49) A complete print out of the corrected coefficient c (*cf.* ref. 52). 50–52) Corrected coefficients of the eqn. $\hat{Y}=a+bX+cX^2$. 53) The theoretical ordinate value at the beginning of the experiment (here: the logarithm of the cytochrome c concentration in the reaction mixture); S. 54) The time correction constant (decimal point correct), to be added to the original time readings. 55) and 56) The coefficients a and b of the linear regression eqn. (refs. 17–18); S. 57) and 58) The corrected coefficients of the linear regression eqn. $\hat{Y}=a+bX$.

Exemplification of the use of the accessory cards. 59) A recalculation which may comprise the addition and removal of experimental points is started by direct introduction of the sums of squares etc. (*cf.* the print out ref. 9). Card No. 1 S; V. 60) The numbers printed out with Card no. 1 S (ref. 9) are introduced; S. 61) An experimental point is deleted (here: the last one which appeared to deviate noticeable from the regression line (*cf.* Fig. 2, Ref. 1): Card 1 D; Y. 62) The number ref. 8. 63) and 64) The measurement to be deleted is introduced; S. 65) The number of remaining experimental points. 66) The variance about the curvilinear regression line. The variance quotient (*cf.* ref. 31) $0.001055/0.000426=2.48$ is significant ($0.01>P>0.001$), and so the removal of the point in this extrapolation is suggested. Removal of the next last point could not be justified in this way. 67) An example showing how more experimental points can be introduced after one experimental point has been deleted with Card 1 D: Card No. 1; Y. 68) The number of remaining experimental points is introduced (*cf.* ref. 65); F↑ (N.B.). 69) and 70) Additional experimental points are introduced (*cf.* ref. 6 and 7).

* Extinction value.

Table 3.

```
        V  1
   37   T  2

 0.01   T  3
   10   T  4
        S  5

 0.06   S  6
 8.09   S  7

 0.08   S
 8.05   S

0.095   S
 8.01   S

 0.11   S
 7.97   S

0.125   S
 7.92   S

0.145   S
 7.87   S

 0.17   S
 7.82   S

 0.19   S
 7.77   S

 0.21   S
 7.72   S

 0.23   S
 7.66   S

 0.26   S
 7.61   S

0.285   S
 7.55   S

 0.31   S
 7.49   S

0.335   S
 7.43   S

0.365   S
 7.36   S

0.395   S
 7.28   S

 0.43   S
 7.21   S

 0.46   S
 7.13   S

 0.49   S
 7.05   S

 0.53   S
 6.96   S

 0.57   S
 6.87   S

 0.60   S
 6.77   S

 0.65   S
 6.67   S

 0.70   S
 6.56   S

 0.75   S
 6.44   S

 0.80   S
 6.31   S

 0.87   S
 6.17   S

 0.93   S
 6.02   S

 0.99   S
 5.85   S

 1.06   S
 5.67   S

 1.14   S
 5.47   S

 1.245  S
 5.24   S

 1.36   S
 4.99   S

 1.49   S
 4.70   S

 1.67   S
 4.35   S

 1.87   S
 3.95   S

 2.19   S
 3.47   S

   37   F◊  8

                W  9
 24.160000   D◊
 26.320300   B◊
247.450000   C◊
 37.214576   C◊
137.464500   d◊
 60.672485   D◊
130.903716   e◊
1710.204100   E◊

                V 10
   37           S 11
 6.687837       A◊ 12

 10.544473      D◊ 13
  0.952282      C◊ 14

                Z 15
 6.687837       S 16
 8.181086   a◊  17
-2.286851   D◊  18

                V 19

                V 20

 6.687837       S 21

 8.273020   a◊  22
-2.616969   D◊  23
 0.173787   c◊  24

   37           S 25

 0.154670   D◊  26
   35       R◊  27
 0.004419   A◊  28

 0.035889   C◊  29
   34       R◊  30
 0.001055   A◊  31

 0.118781   A◊  32

112.588625   F◊  33

 0.032480   ◊   34

                f◊  35
-1.535065       S 36
4.017217517985  R◊  37

                V 38
  0.01          S 39
   10           S 40

 0.004017       R◊  41

10.544473       S 42
 0.952282       S 43

 0.000049       R◊  44

 0.003248   ◊   45
 0.000032   B◊  46

 8.273020       S 47
 0.173787       S 48
0.000001737870  R◊  49
```

Card No. 5 E (Expanded) provides correction of the decimal point of the main print out from Card No. 4 E, the standard deviation of the enzyme activity and the time correction constant.

Card No. 1 V (Viscosimetric) has a program to be used instead of that on Card No. 1 for the introduction of data when viscosimetric readings needing a coordinate transformation are used for the calculation of enzymic activity.

Card 1 VD (Viscosimetric Delete) is used correspondingly to Card 1 Delete.

Calculations

An example of an expanded print out of calculations of extrapolated enzymic activity is reproduced as Table 3; it is a numerical recalculation of the cytochrome c reductase activity read from the graph shown as Fig. 2 in a publication already referred to here.[1] Table 4 consists of an example of a print out of the calculation of hyaluronidase activity from an experiment published as Fig. 2 in a recent article.[7] In both examples the curvilinearity of the regression line is highly significant, as indicated by the high F value.

Our calculation of enzyme activities, extrapolated to zero time, and their standard deviations, have been done with the programs shown in Tables 1 and 2 within a period which is only about a couple of minutes longer than the time needed for entering the measurements into the computing calculator through its key-board.

RESULTS

The cytochrome c reductase activity in a reaction mixture described in Figs. 1—4 of an earlier article[1] was recalculated as shown in Table 3 here. In the reaction mixture, the enzymic activity and its standard deviation (34 degrees of freedom) were 0.00387 ± 0.00004 mg ml^{-1} sec^{-1}. The concentration of the enzyme preparation was 0.070 mg ml^{-1}, and consequently its activity was 0.0553 ± 0.0006 mg sec^{-1} mg^{-1}.

Table 4. An example of the calculation of hyaluronidase activity from viscosimetric measurements. 1) Card No. 1 V; V. 2) The code number for the experiment; ↑. 3) Keyboard clear key; S. 4) The flow time for water; S. 5) The time for the reading; S. 6) The flow time for the reaction mixture; S; 5 and 6 are repeated until all experimental points have been introduced. 7) Card No. 2; V. 8) The number of experimental points; S. 9) A number ($\bar{y}$) to be reintroduced later if Card 2 L or 4 E is used. 10) and 11) Numbers to be reintroduced later. 12) Card No. 3; V. 13) Card No. 4; V. 14) The inclination of the tangent at zero abscissa value; *i.e.*, the coefficient b in the curvilinear regression eqn. $\hat{Y} = a + bX + cX^2$. 15) The number ref. 8; S. 16) The variance ratio F for curvilinearity. 17) The standard deviation about the curvilinear regression line. 18) and 19) The numbers ref. 10 and 11; S. 20) The standard deviation of the inclination of the tangent (*cf.* ref. 14). 21) A factor by which the inclination of the tangent (ref. 14) is multiplied to give the extrapolated enzyme activity (here the factor had not been programmed). 22) If the factor printed is to be used: S; if it must be changed: the proper factor is introduced; S. 23) The result: the extrapolated enzyme activity. 24) The standard deviation of the enzyme activity. 25) The removal of an experimental point: Card 1 VD; Y. 26) The flow time for water; S. 27) and 28) The numbers to be removed; S. 29) An example showing how more experimental points can be introduced after an experimental point has been deleted with Card 1 VD: Card No. 1 V; Y. 30) The flow time for water; S. 31) and 32) Additional experimental points are introduced (*cf.* refs. 5 and 6).

Table 3. Continued

```
 0·827302  a◊ 50
-0·002616  b◊ 51
 0·000001  c◊ 52

      0·824   S 53
-1·262232  D◊ 54

 8·181086   S 55
-2·286851   S 56

 0·818108  a◊ 57
-0·002286  b◊ 58

              V 59
     24·16    S 60
   26·3203    S
    247·45    S
 37·214576    S
  137·4645    S
 60·672485    S
130·903716    S
1710·2041     S

              Y 61
        37    S 62
      2·19    S 63
      3·47    S 64

36·000000  F◊ 65

              V 10
        36    S
 6·777222  A◊

 8·116398  b◊
 0·956838  C◊

              V 19

              V 20

 6·777222     S
```

```
 8·251910  a◊
-2·517875  b◊
 0·103552  c◊

        36    S

 0·036439  D◊
        34  R◊
 0·001071  A◊

 0·014067  C◊
        33  R◊
 0·000426  A◊ 66

 0·022372  A◊

52·516431  F◊

 0·020639   ◊

              f◊
-1·535065    S
3·865101786875 R◊

              V 38
      0·01   S
        10   S

 0·003865  R◊

 8·116398   S
 0·956838   S

 0·000036  R◊

 0·002063   ◊
 0·000024  B◊

              Y 67

        36  F† 68
      2·19    S 69
      3·47    S 70

        37  F◊  8
```

Table 4.

```
              V 1
      1       † 2
              S 3

 25·8         S 4

  0·5         S 5
 47·4         S 6

 1·75         S
 40·1         S

 2·75         S
 37·2         S

 3·75         S
 35·2         S

 4·75         S
 34·0         S

 5·75         S
 33·2         S

 6·75         S
 32·4         S

 7·75         S
 32·0         S

  8·5         S
 31·6         S

  9·5         S
 31·2         S

 10·5         S
 31·0         S

 11·5         S
 30·8         S

 12·5         S
 30·5         S
```

```
              V  7
       13     S  8
3·657433   A◊  9

175·201924  b◊ 10
 0·968614  C◊ 11

              V 12

              V 13
 0·523046  b◊ 14

       13     S 15

216·666880  F◊ 16
 0·039471   ◊ 17

175·201924   S 18
 0·968614    S 19

 0·011958  B◊ 20

              f◊ 21
     1500     S 22
784·569000  R◊ 23

17·937000  R◊ 24

              Y 25
      25·8    S 26

      12·5    S 27
      30·5    S 28

              Y 29
      25·8    S 30

      12·5    S 31
      30·5    S 32
```

The hyaluronidase activity of the staphylococcus toxin solution,[7] and the corresponding standard deviation (10 degrees of freedom) were calculated as shown in Table 4 here: 785 ± 18 H.U./ml.

DISCUSSION

A routine use of a program for linear regression would obviously have caused errors in the results from both examples, whereas the present program largely eliminates the need for special attention to such risks, *i.e.*, if a second degree equation fits the data satisfactorily. However, one must still pay attention to the risk of errors caused by significantly deviating points. The standard deviation of the experimental points about the curvilinear regression line is printed out to provide a screening for deviating points; a serious misreading would cause the standard deviation to be larger than usual.

It must be emphasized that the standard deviation calculated here for the enzymic activity only concerns the variability from one reading to another within the same experiment. It was noticed during the course of this investigation that other sources of variation have a considerable influence on the reproducibility of the enzyme activity determinations.

The use of our curvilinear regression program apparently is not limited to studies of enzyme reaction rates, and it can, of course, be translated for use with other computing desk-top calculators of sufficient size.

Acknowledgement. We wish to thank Professor Bo Holmstedt for the use of an Olivetti Programma computing desk-top calculator which was bought for funds obtained from *Statens medicinska forskningråd*, grant No. B 69—40X—2549—01.

REFERENCES

1. Paléus, S., Tota, B., Hultin, E. and Liljeqvist, G. *Acta Chem. Scand.* **23** (1969) 1.
2. Hultin, E., Paléus, S., Tota, B. and Liljeqvist, G. *Acta Chem. Scand.* **23** (1969) 3417.
3. Dixon, W. J. and Massey, F. J., Jr. *Introduction to statistical analysis*, 2nd. Ed., McGraw, New York 1957.
4. Snedecor, G. W. *Statistical methods applied to experiments in agriculture and biology*, 5th Ed., Iowa State College Press, Ames, Iowa 1956.
5. Olivetti *General Manual. Programma 101 Electronic Desk-top Computer*. Olivetti, Milano 1967.
6. Olivetti *Programming Manual. Programma 101 Electronic Desk-top Computer*.
7. Tirunarayanan, M. G. and Lundblad, G. *Acta Pathol. Microbiol. Scand.* **73** (1968) 211.

Received March 31, 1969.