THE EVALUATION OF THE BACTERICIDAL ACTIVITY OF ETHYLENE GLYCOL AND SOME OF ITS MONOALKYL ETHERS AGAINST *BACTERIUM COLI*

PART VI

BY H. BERRY AND I. MICHAELS

From the Department of Pharmaceutics, School of Pharmacy, University of London

Received 17th October, 1947

PREVIOUS communications in this series[1,2,3,4] have illustrated how the disinfection data of *Bact. coli* could be transformed into a relationship suitable for statistical treatment; this was achieved by plotting percentage survivors as probits against log. survivor time. Although it had been shown that the section of the regression between probits 4 and 6 was not strictly linear, critical analysis had indicated[5] that there was no significant difference between the slopes of regressions obtained from different concentrations of the disinfectant (ethylene glycol) and of regressions obtained from the same concentration. It had also been argued that a disinfectant-organism reaction has a characteristic probit-log time regression which should remain constant under standard experimental conditions.

In the present paper, the regression coefficients of higher members of the homologous series (the monoalkyl ethers from methyl to hexyl) have been determined, together with the confidence limits of the estimations. Physical constants of these compounds have been published in Appendix 1 of Part I[1] of this series of papers. Preliminary experiments with the monoamyl ether indicated that it was so sparingly soluble in water that even the saturated solution showed very little bactericidal activity. Investigations with this compound were therefore discontinued.

DISINFECTION STUDIES IN THE ETHYLENE GLYCOL MONOALKYL ETHERS AT 20°C.

Probit-log time regressions.

Experimental part.—Concentrations of the ethers were prepared and their disinfectant activities tested against *Bact. coli* by means of the standardised technique[1]. In most instances four tests were carried out at each concentration.

Results and calculations.—Probit-log. time regressions were calculated for each test. Summaries of the terms necessary to calculate the mean slopes and for the calculation of the error mean square at each concentration are submitted in Tables IA to VD. Those for ethylene glycol are to be found in Part V[5] of this series of papers.

Combined data from the calculations of the probit-log time regressions for ethylene glycol and its monoalkyl ethers at 20°C.

Table VI presents a summary of the massed statistical data from the calculations of probit-log. time regressions for concentrations of ethylene glycol and its monoalkyl ethers at 20°C. From it has been calculated the mean slope ($\bar{\bar{b}}$ = 1·2040) and the sum of squares for the joint regression (224·508804).

The analysis of variance of the massed regressions is presented in Table VII. The mean square for the residual in y (0·044293) has been used as denominator to calculate the variance ratios between the variation in regression between compounds, variation in regression between concentrations, and variation in regression between tests; these have been included in Table VIII. Owing to the large number of degrees of freedom

TABLE IA

SUMMARY OF TERMS FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS OF DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOMETHYL ETHER AT 20°C.

Observation	Concentrations of ethylene glycol monomethyl ether				
	42·5 per cent.	45·0 per cent.	47·5 per cent.	50·0 per cent.	52·5 per cent.
$S[(x-\bar{x})(y-\bar{y})]$	1·994529	3·189758	4·013346	3·873828	2·923893
$S(x-\bar{x})^2$	4·170328	4·148876	5·797243	4·490993	2·307026
$S(y-\bar{y})^2$	1·485401	3·219608	4·575887	3·857329	3·972279
N	11	11	13	13	9
SS for individual regressions ...	1·075110	2·522687	3·167978	3·585350	3·763250
b	0·478267	0·768825	0·692285	0·862577	1·267386
SS pool	0·953917	2·452364	2·778380	3·341476	3·705702

TABLE IB

CALCULATION OF THE ERROR MEAN SQUARE OF REGRESSIONS FROM CONCENTRATIONS OF ETHYLENE GLYCOL MONOMETHYL ETHER AT 20°C.

Concentration	Item	Sum of squares	N	Mean square
42·5 per cent.	Common regression ...	0·953917	1	0·953917
	Variation in regression ...	0·121183	3	0·040394
	Total	1·075100	4	
	Residual in y	0·410291	11	0·037299
45·0 per cent.	Common regression ...	0·452354	1	0·452354
	Variation in regression ...	0·070323	3	0·023441
	Total	2·522687	4	
	Residual in y	0·696921	11	0·063356
47·5 per cent.	Common regression ...	2·778380	1	2·778380
	Variation in regression ...	0·389598	3	0·129866
	Total	3·167978	4	
	Residual in y	1·407909	13	0·108301
50·0 per cent.	Common regression ...	3·341476	1	3·341476
	Variation in regression ...	0·243874	3	0·081291
	Total	3·585350	4	
	Residual in y	0·273979	13	0·021075
52·5 per cent.	Common regression ...	3·705702	1	3·705702
	Variation in regression ...	0·057548	3	0·019183
	Total	3·763250	4	
	Residual in y	0·209029	9	0·023225

TABLE Ic

SUMMARY OF STATISTICAL DATA FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS FOR CONCENTRATIONS OF ETHYLENE GLYCOL MONOMETHYL ETHER AT 20°C.

Concentration	Residual in y		Variation in b		SS pooled b	$S[(x-)\bar{x}(y-\bar{y})]$	$S(x(-\bar{x})^2$
	SS	N	SS	N			
42·5 per cent. ...	0·410291	11	0·121183	3	0·953917	1·994529	4·170328
45·0 ,, ,, ...	0·696921	11	0·070323	3	2·452364	3·189758	4·148876
47·5 ,, ,, ...	1·407909	13	0·389598	3	2·778380	4·013346	5·797243
50·0 ,, ,, ...	0·273979	13	0·243874	3	3·341476	3·873838	4·490993
42·5 ,, ,, ...	0·209029	9	0·057548	3	3·705702	2·923893	2·307026
Totals	2·998129	57	0·882526	15	13·231839	15·995354	20·914466

$$\bar{b} = \frac{15 \cdot 995345}{20 \cdot 914466} = 0 \cdot 764799$$

$$SS \text{ for joint regression} = \frac{(15 \cdot 995354)^2}{20 \cdot 914466} = 12 \cdot 233224$$

TABLE Id

MEAN SQUARES OF THE VARIATIONS IN THE PROBIT-LOG TIME REGRESSIONS FROM DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOMETHYL ETHER AT 20°C.

Item	N	Sum of squares	Mean square
Grand regression	1	12·233224	12·233224
Variation in regression between concentrations ...	4	0·998615	0·249654
Variation in regression within concentrations ...	15	0·882526	0·058835
Residual in y	57	2·998129	0·052599

TABLE IIa

SUMMARY OF TERMS FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS OF DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOETHYL ETHER AT 20°C.

Observation	Concentrations of ethylene glycol monoethyl ether				
	25·0 per cent.	27·5 per cent.	30·0 per cent.	32·5 per cent.	35·0 per cent.
$S[(x-\bar{x})(y-\bar{y})]$	2·273209	1·570023	5·630061	3·867540	2·473940
$S(x-\bar{x})^2$	2·325423	2·284708	6·834377	3·203024	2·463525
$S(y-\bar{y})^2$	2·348405	1·286192	5·179256	5·533561	2·907313
N	8	7	15	11	11
SS for individual regressions ...	2·233417	1·200121	4·931650	5·002978	2·549408
b	0·977547	0·687188	0·823786	1·207465	1·004228
SS pool	2·223167	1·078900	4·637963	4·669920	2·484399

nvolved it is not possible to use the ordinary statistical tables in order to compute the probabilities ; this difficulty was overcome by calculating z as exemplified by Mather[6] (section 16). The expected values of z for the variance ratios have been calculated from the relationship $z = \frac{1}{2}$ log. (Variance Ratio) ; these have been set out in Table VIII. The calculated values of z have been compared with the theoretical values, to give the

TABLE IIB

CALCULATION OF THE ERROR MEAN SQUARE OF REGRESSIONS FROM CONCENTRATIONS OF ETHYLENE GLYCOL MONOETHYL ETHER AT 20°C.

Concentration	Item	Sum of squares	N	Mean square
25·0 per cent.	Common regression ...	2·222167	1	2·222167
	Variation in regression ...	0·011250	2	0·005625
	Total	2·011250	3	
	Residual in y	0·114988	8	0·014373
27·5 per cent.	Common regression ...	1·078900	1	1·078900
	Variation in regression ...	0·121221	1	0·121221
	Total	1·200121	2	
	Residual in y	0·086071	7	0·012296
30·0 per cent.	Common regression ...	4·637963	1	4·637963
	Variation in regression ...	0·293687	3	0·097896
	Total	4·931650	4	
	Residual in y	0·247606	15	0·016507
32·5 per cent.	Common regression ...	4·669920	1	4·669920
	Variation in regression ...	0·333058	4	0·084265
	Total	5·002978	5	
	Residual in y	0·526583	00	0·047871
35·0 per cent.	Common regression ...	2·484399	1	2·484399
	Variation in regression ...	0·065009	3	0·021669
	Total	2·549408	4	
	Residual in y	0·357905	11	0·032537

TABLE IIC

SUMMARY OF STATISTICAL DATA FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS FOR CONCENTRATIONS OF ETHYLENE GLYCOL MONOETHYL ETHER AT 20°C.

Concentration	Residual in y		Variation in b		SS pooled b	$S[(x-\bar{x})(y-\bar{y})]$	$S(x-\bar{x})^2$
	SS	N	SS	N			
25·0 per cent. ...	0·111498	8	0·011250	2	2·222167	2·273209	2·325423
27·5 „ „ ...	0·086071	7	0·121221	1	1·078900	1·570023	2·284708
30·0 „ „ ...	0·247606	15	0·293687	3	4·637963	5·630061	6·834377
32·5 „ „ ...	0·526583	11	0·333058	4	4·669920	3·867540	3·203024
35·0 „ „ ...	0·357905	11	0·065009	3	2·484399	2·473940	2·463525
Totals	1·333153	52	0·824225	13	15·093349	15·814773	17·111057

$$\bar{b} = \frac{15\cdot814773}{17\cdot111057} = 0\cdot924243$$

$$SS \text{ for joint regression} = \frac{(15\cdot814773)^2}{17\cdot111057} = 14\cdot616692$$

probabilities of the significance between the items selected; these have also been included in Table VIII.

Test of significance of the difference between the mean squares for the variation in regression between concentrations and variation in regression between individual tests.

The formula for high values of N_1 and N_2 is given in Table V (Distribution of z) of the Statistical Tables of Fisher and Yates[7].

$$z_{(5 \text{ per cent.})} = \frac{1 \cdot 6499}{\sqrt{h - 1}} - 0 \cdot 7843 \left(\frac{1}{N_1} - \frac{1}{N_2}\right) \text{approximately,}$$

$$\text{where } \frac{2}{h} = \frac{1}{N_1} + \frac{1}{N_2}$$

$N_1 = 27$ and $N_2 = 172$, therefore $h = 46 \cdot 6734$.

$$\text{Hence } z_{(5 \text{ per cent.})} = \frac{1 \cdot 6449}{\sqrt{46 \cdot 6734 - 1}} - (0 \cdot 7843 \times 0 \cdot 0132) = 0 \cdot 2189$$

The observed value of z (0·082202) is less than that calculated at the 5 per cent. level, hence $P > 0 \cdot 05$ and the difference is not significant.

TABLE IID

MEAN SQUARES OF THE VARIATIONS IN THE PROBIT-LOG TIME REGRESSIONS FROM DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOETHYL ETHER AT 20°C.

Item	N	Sum of squares	Mean square
Grand regression	1	14·616692	14·616692
Variation in regression between concentrations ...	4	0·476657	0·119164
Variation in regression within concentrations ...	13	0·824325	0·062948
Residual in y	52	1·333153	0·025638

TABLE IIIA

SUMMARY OF TERMS FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS OF DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOPROPYL ETHER AT 20°C.

Observation	Concentrations of ethylene glycol monopropyl ether				
	7·8 per cent.	9·0 per cent.	10·0 per cent.	11·0 per cent.	12·0 per cent.
$S[(x-\bar{x})(y-\bar{y})]$	4·589090	2·749305	3·868675	2·837293	1·773908
$S(x-\bar{x})^2$	2·620754	1·649256	1·970901	1·710080	0·847386
$S(y-\bar{y})^2$	9·647604	5·729806	8·440537	5·643021	5·642021
N	12	7	10	9	4
SS for individual regressions ...	8·184688	5·574039	7·776921	4·772751	4·732316
b	1·751057	1·666997	1·962897	1·659158	2·093388
SS pool	8·035759	4·583084	7·593809	4·707518	3·713478

The two errors may therefore be pooled (as in Table VII); the combined sum of squares is 19·633745 for 199 degrees of freedom, with a mean square of 0·098662. This figure therefore represents the error mean square of the estimation of the slopes of all the regressions for ethylene glycol and its monoalkyl ethers at 20°C., and is used later in determining the standard errors.

TABLE IIIB

CALCULATION OF THE ERROR MEAN SQUARE OF REGRESSIONS FROM CONCENTRATIONS OF ETHYLENE GLYCOL MONOPROPYL ETHER AT 20°C.

Concentration	Item	Sum of squares	N	Mean square
7·8 per cent.	Common regression ...	8·035759	1	8·035759
	Variation in regression ...	0·148929	3	0·049643
	Total	8·184688	4	
	Residual in y	1·462916	12	0·121909
9·0 per cent.	Common regression ...	4·583084	1	4·583084
	Variation in regression ...	0·990955	3	0·330318
	Total	5·574039	4	
	Residual in y	0·155767	7	0·022252
10·0 per cent.	Common regression ...	7·593809	1	7·593,809
	Variation in regression ...	0·183112	3	0·061037
	Total	7·776921	4	
	Residual in y	0·060270	9	0·007699
11·0 per cent.	Common regression ...	4·707518	1	4·707518
	Variation in regression ...	0·065233	3	0·021744
	Total	4·772751	4	
	Residual in y	0·069270	9	0·007699
12·0 per cent.	Common regression ...	3·713478	1	3·713478
	Variation in regression ...	1·018838	3	0·272946
	Total	4·732316	4	
	Residual in y	1·109705	4	0·277426

TABLE IIIC

SUMMARY OF STATISTICAL DATA FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS FOR CONCENTRATIONS OF ETHYLENE GLYCOL MONOPROPYL ETHER AT 20°C.

Concentration	Residual in y		Variation in b		SS pooled b	$S[(x-\bar{x})(y-\bar{y})]$	$S(x-\bar{x})^2$
	SS	N	SS	N			
7·8 per cent. ...	1·462916	12	0·148929	3	8·035759	4·589090	2·620754
9·0 „ „ ...	0·155767	7	0·990955	3	4·583084	2·749305	1·649256
10·0 „ „ ...	0·663616	10	0·183112	3	7·593809	3·868675	1·970901
11·0 „ „ ...	0·069270	9	0·065233	3	4·707518	2·837293	1·710080
12·0 „ „ ...	1·109705	4	1·018838	3	3·713478	1·773908	0·847386
Totals	3·461274	42	2·407067	15	28·633648	15·818271	8·798377

$$\bar{b} = \frac{15\cdot818271}{8\cdot798377} = 1\cdot797864$$

$$SS \text{ for joint regression} = \frac{(15\cdot818271)^2}{8\cdot798377} = 28\cdot439074$$

Test of significance of the difference between the mean squares for the variation in regression between the different compounds and the residual in y.

Here $N_1 = 5$ and $N_2 = 557$; P may be found by direct consultation of the table showing the distribution of z. The observed value (0·994715) is larger than the theoretical value even at the 0·1 per cent. level (which is about 0·7), hence $P < 0{\cdot}001$.

Test of significance of the difference between the mean squares for the variation in regression between concentrations and the residual in y.

Here $N_1 = 27$ and $N_2 = 557$, therefore $h = 51{\cdot}5034$; from this, $z_{(5\text{ per cent.})} = 0{\cdot}20376$.

The observed value of z (0·268825) is slightly greater than that calculated at the 5 per cent. level and hence $P < 0{\cdot}05$.

$$z_{(1\text{ per cent.})} = \frac{2{\cdot}3263}{\sqrt{h-1{\cdot}4}} - 1{\cdot}235\left(\frac{1}{N_1} - \frac{1}{N_2}\right) \text{ approximately,}$$

$$= \frac{2{\cdot}3263}{\sqrt{51{\cdot}5034} - 1} - 1{\cdot}235\,(0{\cdot}03524) = 0{\cdot}2852.$$

TABLE IIID

MEAN SQUARES OF THE VARIATIONS IN THE PROBIT-LOG TIME REGRESSIONS FROM DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOPROPYL ETHER AT 20°C.

Item	N	Sum of squares	Mean square
Grand regression	1	28·439074	28·439074
Variation in regression between concentrations ...	4	0·194574	0·048644
Variation in regression within concentrations ...	15	2·407067	0·160471
Residual in y	42	3·461274	0·082411

TABLE IVA

SUMMARY OF TERMS FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS OF DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOBUTYL ETHER AT 20°C.

Observation	Concentrations of ethylene glycol monobutyl ether				
	3·5 per cent.	3·75 per cent.	4·0 per cent.	4·25 per cent.	4·5 per cent.
$S[(x-\bar{x})(y-\bar{y})]$	4·059036	2·358281	3·279052	3·267227	2·335002
$S(x-\bar{x})^2$	2·525145	1·856849	2·038721	2·656913	1·714294
$S(y-\bar{y})^2$	7·013784	5·127533	5·590826	5·812414	3·560732
N	10	6	10	13	7
SS for individual regressions ...	6·696465	4·799289	5·470142	4·155352	3·443844
b	1·607447	1·270045	1·608387	1·229708	1·362078
SS pool	6·524684	2·995122	5·273984	4·017735	3·180455

The observed value of z is smaller than that calculated at the 1 per cent. level ; hence $P = 0{\cdot}05$ to $0{\cdot}01$.

Test of significance of the difference between the mean squares for the variation in regression between tests and the residual in y.

$$z_{(0\cdot 1 \text{ per cent.})} = \frac{3\cdot 0902}{\sqrt{h-2\cdot 1}} - 1\cdot 925\left(\frac{1}{N_1} - \frac{1}{N_2}\right) \text{ approximately.}$$

Here, $N_1 = 172$ and $N_2 = 557$, therefore $h = 262\cdot 837$.

$$\text{Hence } z_{(0\cdot 1 \text{ per cent.})} = \frac{3\cdot 0902}{\sqrt{262\cdot 837 - 2\cdot 1}} - 1\cdot 925\,(0\cdot 004) = 0\cdot 1838$$

The observed value of z ($0\cdot 186797$) is greater than that calculated at the $0\cdot 1$ per cent. level, hence $P < 0\cdot 001$.

Test of significance of the difference between the mean squares for the variation in regression between the pooled error (i.e. variation between concentrations + between tests) and the residual in y.

Here $N_1 = 199$ and $N_2 = 557$, hence $h = 292{\cdot}24$.

$$\text{Hence } z_{(0\cdot 1 \text{ per cent.})} = \frac{3\cdot 0902}{\sqrt{292\cdot 24 - 2\cdot 1}} - 1\cdot 925\,(0\cdot 0032) = 0\cdot 1749$$

The observed value of z ($0\cdot 20021$) is greater than that calculated at the $0\cdot 1$ per cent. level, hence $P < 0\cdot 001$.

TABLE IVB

CALCULATION OF THE ERROR MEAN SQUARE OF REGRESSIONS FROM CONCENTRATIONS OF ETHYLENE GLYCOL MONOBUTYL ETHER AT 20°C.

Concentration	Item	Sum of squares	N	Mean square
3·50 per cent.	Common regression ...	6·524684	1	6·524684
	Variation in regression ...	0·071781	3	0·057260
	Total	6·696465	4	
	Residual in y	0·317319	10	0·031732
3·75 per cent.	Common regression ...	2·995122	1	2·995122
	Variation in regression ...	1·804167	3	0·601389
	Total	4·799289	4	
	Residual in y	0·328244	6	0·054707
4·00 per cent.	Common regression ...	5·273984	1	5·273984
	Variation in regression ...	0·196158	3	0·653860
	Total	5·470142	4	
	Residual in y	0·120684	10	0·012068
4·25 per cent.	Common regression ...	4·017735	1	4·017735
	Variation in regression ...	0·137617	3	0·045872
	Total	4·155352	4	
	Residual in y	1·657062	13	0·127466
4·50 per cent.	Common regression ...	3·180455	1	3·180455
	Variation in regression ...	0·263389	3	0·087796
	Total	3·443844	4	
	Residual in y	0·116888	7	0·016697

TABLE IVc

SUMMARY OF STATISTICAL DATA FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS FOR CONCENTRATIONS OF ETHYLENE GLYCOL MONOBUTYL ETHER AT 20°C.

Concentration	Residual in y		Variation in b		SS pooled b	$S[(x-\bar{x})(y-\bar{y})]$	$S(x-\bar{x})^2$
	SS	N	SS	N			
3·50 per cent. ...	0·317319	10	0·171781	3	6·524684	4·059036	2·525145
3·75 „ „ ...	0·328244	6	1·804167	3	2·995122	2·358281	1·856849
4·00 „ „ ...	0·120684	10	0·196158	3	5·273984	3·279052	2·038721
4·25 „ „ ...	1·657062	13	0·137617	3	4·017735	3·267227	2·656913
4·50 „ „ ...	0·116888	7	0·263389	3	3·180455	2·335002	1·714294
Totals	2·540197	46	2·573112	15	21·991980	15·298598	10·791922

$$\bar{b} = \frac{15 \cdot 298598}{10 \cdot 791922} = 1 \cdot 417597$$

$$SS \text{ for joint regression} = \frac{(15 \cdot 298598)^2}{10 \cdot 791922} = 21 \cdot 687249$$

TABLE IVd

MEAN SQUARES OF THE VARIATIONS IN THE PROBIT-LOG TIME REGRESSIONS FROM DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOBUTYL ETHER AT 20°C.

Item	N	Sum of squares	Mean square
Grand regression, ...	1	21·687249	21·687249
Variation in regression between concentrations ...	4	0·304731	0·076183
Variation in regression within concentrations ...	15	2·573112	0·171541
Residual in y	46	2·540197	0·055222

TABLE Va

SUMMARY OF TERMS FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS OF DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOHEXYL ETHER AT 20°C.

Observation	Concentrations of ethylene glycol monohexyl ether					
	0·400 per cent.	0·425 per cent.	0·450 per cent.	0·475 per cent.	0·500 per cent.	Miscellaneous
$S[(x-)\bar{x}(y-\bar{y})]$ ...	6·410363	4·731915	2·950592	4·051499	4·047545	2·493439
$S(x-\bar{x})^2$	3·410920	2·432214	1·872617	2·945145	2·889779	1·494108
$S(y-\bar{y})^2$	12·666664	10·214115	4·703732	5·951287	6·096003	4·714120
N	17	12	6	10	15	15
SS for individual regressions ...	12·099071	9·629824	4·659579	5·735731	5·866990	4·496297
b	1·879365	1·945518	1·575652	1·375653	1·400642	1·668848
SS pool	12·047411	9·206024	4·649105	5·573459	5·669160	4·161170

TABLE Vb

CALCULATION OF THE ERROR MEAN SQUARE OF REGRESSIONS FROM CONCENTRATIONS OF ETHYLENE GLYCOL MONOHEXYL ETHER AT 20°C.

Concentration	Item	Sum of squares	N	Mean square
0·400 per cent.	Common regression ...	12·047411	1	12·047411
	Variation in regression ...	0·051660	4	0·012915
	Total	12·099071	5	
	Residual in y	0·567593	17	0·033388
0·425 per cent.	Common regression ...	9·206024	1	9·206024
	Variation in regression ...	0·423800	4	0·105950
	Total	9·629824	5	
	Residual in y	0·584291	12	0·048524
0·450 per cent.	Common regression ...	4·649105	1	4·649105
	Variation in regression ...	0·010474	1	0·010474
	Total	4·659579	2	
	Residual in y	0·044153	6	0·007356
0·475 per cent.	Common regression ...	5·573459	1	5·573459
	Variation in regression ...	0·162272	3	0·054091
	Total	5·735731	4	
	Residual in y	0·215556	10	0·021556
0·500 per cent.	Common regression ...	5·669160	1	5·669160
	Variation in regression ...	0·197830	4	0·049458
	Total	5·866990	5	
	Residual in y	0·229013	15	0·015268
Miscellaneous	Common regression ...	4·161170	1	4·161170
	Variation in regression ...	0·335127	5	0·067025
	Total	4·496297	6	
	Residual in y	0·217823	13	0·016756

TABLE Vc

SUMMARY OF STATISTICAL DATA FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS FOR CONCENTRATIONS OF ETHYLENE GLYCOL MONOHEXYL ETHER AT 20°C.

Concentration	Residual in y		Variation in b		SS pooled b	$S[(x-\bar{x})(y-\bar{y})]$	$S(x-\bar{x})^2$
	SS	N	SS	N			
0·400 per cent. ...	0·567593	17	0·051660	4	12·047411	6·410363	3·410920
0·425 „ „ ...	0·584291	12	0·423800	4	9·206024	4·731915	2·432214
0·450 „ „ ...	0·044153	6	0·010474	1	4·649105	2·950592	1·872617
0·475 „ „ ...	0·215556	10	0·162272	3	5·573459	4·051449	2·945145
0·500 „ „ ...	0·229013	15	0·197830	4	5·669160	4·047545	2·889779
Miscellaneous ...	0·217823	13	0·335127	5	4·161170	2·493439	1·494108
Totals	1·858429	73	1·181163	21	41·306329	24·685353	15·044783

$$\bar{b} = \frac{24·685353}{15·044783} = 1·640792$$

$$SS \text{ for joint regression} = \frac{(24·685353)^2}{15·044783} = 40·503519$$

TABLE VD

MEAN SQUARES OF THE VARIATIONS IN THE PROBIT-LOG TIME REGRESSIONS FROM DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL MONOHEXYL ETHER AT 20°C.

Item	N	Sum of squares	Mean square
Grand regression	1	40·503519	40·503519
Variation in regression between concentrations ...	5	0·802810	0·160562
Variation in regression within concentrations ...	21	1·181163	0·056246
Residual in y	73	1·858429	0·025458

INFERENCES FROM THE ANALYSIS

The analysis of variance indicates that there is a significant large variation in the regressions between the different compounds, i.e. each substance has its characteristic regression coefficient which differs significantly from the average of the series. The analysis also shows that the variations in the regressions between the concentrations of the substances are of the same order as the variation between the individual tests at a particular concentration. Hence the regressions of different concentrations of the same substance may be taken as parallel.

Calculation of the standard errors of the probit-log time regression coefficients at 20°C.

The variance of b is given by the formula $V_b = \frac{V_y}{S(\bar{x}-x)^2}$ (Mather[6], section 32). The standard error, s_b, of the regression coefficient is $\sqrt{V_b}$. V_y is the error mean square from the pooled error of the variation in regression between concentrations and tests, and is 0·098662 (Table VII). The standard errors of the mean regression coefficients of all the compounds for experiments conducted at 20°C. have been computed and set out in Table IX. The ratio of the regression coefficients to their standard errors in all cases is seen to be large thereby indicating that b has been estimated satisfactorily.

Calculation of the confidence limits of the probit-log time regressions.

It is useful to present graphically the limits of error of the regression lines at a fixed probability level, (P = 0·05 has been chosen), so that the advantages of an increased number of experiments from which to calculate the mean, can be appreciated more readily.

(i) The first stage in this calculation is to determine the sum of the squares for the deviations of $\bar{y}$ (the mean value of the probits in an experiment) from $\bar{\bar{y}}$ (the mean value of $\bar{y}$ for the several tests on the same concentration of the disinfectant). It is necessary to compute this figure for every concentration used for all the compounds. Table X shows the calculations for the monomethyl ether ; the calculations for the remaining compounds are precisely the same. The final figures for all compounds are included in Table XI.

(ii) The next stage is to calculate the empirical variance of $\bar{\bar{y}}$ (the mean of the mean probits). The sum of the squares of the deviations of y is divided by N (the number of degrees of freedom, i.e. one less than the number of experiments), to give the mean square. Thus in the case of 42·5 per cent. monomethyl ether (Table XI) the mean square will be $\frac{0 \cdot 066713}{3} = 0 \cdot 022238$. The grand total of all the sum of squares of the deviations is divided by the total number of degrees of freedom to give the general or pooled mean square. Table XI sets out these figures from which it is seen that the average mean square is $\frac{6 \cdot 43660}{165} = 0 \cdot 039010$.

The variance of $\bar{\bar{y}}$ at a particular concentration is obtained by dividing the general mean square (0·039010) by the number of experiments performed at that concentration; in the instance cited above it will be $\frac{0 \cdot 022238}{4}$, i.e., 0·009753. Hence the more tests performed at a particular concentration the smaller will be the value of $V_{\bar{\bar{y}}}$; thus at 75 per cent. ethylene glycol 32 experiments were used to calculate the mean and here $V_{\bar{\bar{y}}} = 0 \cdot 001219$, whereas at 27·5 per cent. monomethyl ether only 2 experiments were performed and its $V_{\bar{\bar{y}}} = 0 \cdot 019505$.

TABLE VI

SUMMARY OF MASSED STATISTICAL DATA FROM CALCULATIONS OF PROBIT-LOG TIME REGRESSIONS FOR CONCENTRATIONS OF ETHYLENE GLYCOL AND ITS MONOALKYL ETHERS AT 20°C.

Compound	Range of Concentrations Investigated	Variation in b: Residual in y SS	N	Between Tests SS	N	Between Concentrations SS	N	Grand Regression Pooled SS	$S[(x-\bar{x})(y-\bar{y})]$	$S(x-\bar{x})^2$
Ethylene glycol ...	72·5 to 90·0 per cent.	12·479859	287	8·210628	93	0·727637	6	118·868425	98·850144	82·703083
Monomethyl ether ...	42·5 to 52·5 ,,	2·998129	57	0·882526	15	0·998615	4	12·233224	15·995354	20·914466
Monoethyl ether ...	25·0 to 35·0 ,,	1·333153	52	0·824225	13	0·476657	4	14·616692	15·814773	17·111057
Monopropyl ether ...	7·8 to 12·0 ,,	3·461274	42	2·407067	15	0·194574	4	28·439074	15·818171	8·798377
Monobutyl ether ...	3·5 to 4·5 ,,	2·540197	46	2·573112	15	0·304731	4	21·687249	15·298598	10·791922
Monohexyl ether ...	0·4 to 0·5 ,,	1·858429	73	1·181163	21	0·802810	5	40·503519	24·685353	15·044783
Totals ...	...	24·671041	557	16·078721	172	3·505024	27	236·348183	186·462493	154·863688

$\bar{\bar{b}} = \frac{186 \cdot 462493}{154 \cdot 863688} = 1 \cdot 204043$ $\qquad$ SS for joint regression $= \frac{(186 \cdot 462493)^2}{154 \cdot 863688} = 224 \cdot 508804$

The confidence limits will therefore vary with the number of tests performed for a particular concentration. Figure 1 has been constructed to indicate the confidence limits which are to be expected at a probability level of $P = 0{\cdot}05$ when 32 tests are used to calculate the mean (as for

TABLE VII

ANALYSIS OF VARIANCE OF MASSED REGRESSIONS FOR DISINFECTION OF *BACT. COLI* BY CONCENTRATIONS OF ETHYLENE GLYCOL AND ITS MONOALKYL ETHERS AT 20°C.

Item	N	SS	Mean square	Variance Ratio	Probability
Massed regression	1	224·508804	224·508804	(See Table VIII)	
Variation in regression between compounds	5	11·839379	2·367876		
Variation in regression between concentrations	27	3·505024	0·129816		
Variation in regression between tests	172	16·078721	0·093481		
Residual in y	557	24·671041	0·044293		
Pooled error	199	19·633745	0·098662		

TABLE VIII

CALCULATION OF THE z'S FOR THE ITEMS IN THE ANALYSIS OF VARIANCE OF THE MASSED REGRESSIONS IN TABLE VII

N_1/N_2	Variance ratio (V.R.)	$\log_{10}$V.R.	$\log_e$ V.R. = $\log_{10}$ V.R. x 1·15129	$z = \frac{1}{2}\log_e$ V.R.	Probability
27/172	1·388689	0·1428	0·16440	0·08220	<0·05
5/557	53·459373	1·7280	1·98943	0·99472	<0·001
27/557	2·930847	0·4670	0·53765	0·26883	0·05 to 0·01
172/557	2·110514	0·3245	0·37359	0·18680	<0·001
199/557	2·227485	0·3478	0·40042	0·20021	<0·001

TABLE IX

THE PROBIT-LOG TIME REGRESSION COEFFICIENTS WITH THEIR STANDARD ERRORS, OF THE REACTION BETWEEN *BACT. COLI* AND ETHYLENE GLYCOL AND ITS MONOALKYL ETHERS AT 20°C.

Compound	b	N	$S(x-\bar{x})^2$	V_y	$s_b = \frac{V_y}{S(x-\bar{x})^2}$	Ratio of b to s_b
Ethylene glycol	1·202511	93	82·203083	0·098662	0·03465	35
Monomethyl ether... ...	0·764799	15	20·914466		0·06871	11
Monoethyl ether	0·924243	13	17·111057		0·07595	12
Monopropyl ether	1·797864	15	8·798377		0·10570	17
Mobobutyl ether	1·417597	15	10·791922		0·09563	15
Monohexyl ether	1·640792	21	15·044783		0·08098	20

75 per cent. ethylene glycol), and when only 4 tests are used (as for 0·425 per cent. monohexyl ether). The calculations involved are as follows :—

(*a*) *Confidence limits for* 32 *experiments with* 75 *per cent ethylene glycol at* 20°C.

Confidence limits at $P = 0{\cdot}05 = \pm\, s_y \times 1{\cdot}96$ (where $1{\cdot}96$ = the value of the normal deviate at $P = 0{\cdot}05$).

Now $s_y = \sqrt{V_y}$ where $V_y = V_{mean} + V_b\,(x - \bar{x})^2$

But V_y = variance of y

and V_{mean} = variance of $\bar{\bar{y}}$ (i.e. $V_{\bar{\bar{y}}}$ in Table II)

$= 0{\cdot}001219,$

and $V_b = \dfrac{V}{S\,(x - \bar{x})^2}$

where V = error mean square = $0{\cdot}098662$ (Table VII)

and $S\,(x - \bar{x})^2 = 32{\cdot}035396$ (Table VI, Part V[5])

Hence $V_b = \dfrac{0{\cdot}098662}{32{\cdot}035396} = 0{\cdot}003080$

x = abscissa at which value of V_y is to be determined.

$\bar{\bar{x}}$ = mean value of x (i.e. log. time), in the 32 experiments.

$= \dfrac{67{\cdot}608}{32} = 2{\cdot}115.$

The values of V_y at values of x are obtained by substituting in the equation $V_y = V_{mean} + V_b\,(x - \bar{\bar{x}})^2$.

For example, at $x = 1$, $V_y = 0{\cdot}001219 + 0{\cdot}003080\,(1 - 2{\cdot}115)^2$

$= 0{\cdot}005048$

$s_y = \sqrt{0{\cdot}005048} = 0{\cdot}07105$

TABLE X

CALCULATION OF THE SUM OF SQUARES FOR DEVIATIONS OF THE MEAN PROBIT ($\bar{y}$), FOR CONCENTRATIONS OF ETHYLENE GLYCOL MONOMETHYL ETHER AT 20°C.

	Concentrations of ethylene glycol monomethyl ether									
	42·5 per cent.		45·0 per cent.		47·5 per cent.		50·0 per cent.		52·5 per cent.	
	Expt. No.	Mean Probit $\bar{y}$	Expt. No.	Mean Probit $\bar{y}$	Expt. No.	Mean Probit $\bar{y}$	Expt. No.	Mean Probit $\bar{y}$	Expt. No.	Mean Probit $\bar{y}$
	208a	4·414	208d	4·720	209c	4·797	209d	5·092	210g	5·853
	209f	4·503	209e	4·516	211e	4·891	210f	5·591	211g	5·870
	210c	4·759	210d	5·064	213e	5·126	211f	5·423	212e	5·169
	211c	4·617	211d	5·119	214c	5·070	212d	4·807	213d	5·183
$S(\bar{y})$		18·293		19·419		19·884		20·913		22·075
No. of expts.		4		4		4		4		4
$\bar{\bar{y}}$		4·573		4·855		4·971		5·228		5·519
$S(\bar{y})^2$	83·725175		94·520913		98·913866		109·703923		122·296559	
$\frac{S^2(\bar{y})}{n}$	83·658462		94·274390		98·843364		109·338392		121·826406	
$S(\bar{y}-\bar{\bar{y}})^2 = S(\bar{y})^2 - \frac{S^2(\bar{y})}{n} = SS$	0·066713		0·246523		0·070532		0·365531		0·470153	

Confidence limits (at $P = 0 \cdot 05$) $= 0 \cdot 07105 \times 1 \cdot 96 = \pm 0 \cdot 1396$ probits.

Confidence limits (at $P = 0 \cdot 05$) have been calculated for a number of values of x and the results are set out in Table XII.

TABLE XI

THE EMPIRICAL VARIANCE OF THE INDIVIDUAL MEANS FROM THEIR MEAN PROBIT OF EXPERIMENTS WITH CONCENTRATIONS OF ETHYLENE GLYCOL AND ITS MONOALKYL ETHERS AT 20°C.

Compound	Concentration per cent.	SS	N	Mean square	$V_{\bar{y}}$
Ethylene glycol	72·5	0·318245	11	0·028931	0·003251
	75·0	0·476395	31	0·015368	0·001219
	77·5	0·411183	14	0·029370	0·002601
	80·0	0·406955	9	0·045217	0·003901
	82·5	0·307976	8	0·038497	0·004334
	85·0	0·238682	9	0·026520	0·003901
	90·0	0·617642	9	0·068627	0·003901
Monomethyl ether	42·5	0·066713	3	0·022238	0·009753
	45·0	0·246523	3	0·082174	0·009753
	47·5	0·070532	3	0·023511	0·009753
	50·0	0·365531	3	0·121877	0·009753
	52·5	0·470513	3	0·156838	0·009753
Monoethyl ether	25·0	0·141523	2	0·070762	0·013003
	27·5	0·000420	1	0·000420	0·019505
	30·0	0·086265	3	0·028755	0·009753
	32·5	0·316301	4	0·079075	0·007802
	35·0	0·133905	3	0·044635	0·009753
Monopropyl ether	7·8	0·110310	3	0·036770	0·009753
	9·0	0·174145	3	0·058048	0·009753
	10·0	0·076907	3	0·025636	0·009753
	11·0	0·051941	3	0·017314	0·009735
	12·0	0·120228	3	0·040076	0·009753
Monobutyl ether	3·50	0·216277	3	0·072092	0·009753
	3·75	0·411050	3	0·137017	0·009753
	4·00	0·146367	3	0·048789	0·009753
	4·25	0·033001	3	0·011000	0·009753
	4·50	0·078645	3	0·026215	0·009753
Monohexyl ether	0·375	0·000061	1	0·000061	0·019503
	0·400	0·038307	4	0·009577	0·007802
	0·425	0·089235	3	0·029745	0·009753
	0·450	0·016928	1	0·016928	0·019503
	0·475	0·149173	3	0·049724	0·009753
	0·500	0·048775	4	0·012194	0·007802
Totals	—	6·436660	165	0·039010	—

TABLE XII

CONFIDENCE LIMITS AT $P = 0{\cdot}05$ FOR VALUES OF x FOR 32 EXPERIMENTS WITH 75 PER CENT. ETHYLENE GLYCOL, AND 4 EXPERIMENTS WITH 0·425 PER CENT. MONOHEXYL ETHER, AT 20°C.

Value of x	Confidence Limits for 32 experiments (75 per cent. ethylene glycol)	Confidence Limits for 4 experiments (0·425 per cent. monohexyl ether)
1·0	± 0·1393 probits	± 0·4214 probits
1·2	0·1208 ,,	0·3130 ,,
1·4	0·1017 ,,	0·2903 ,,
1·6	0·0884 ,,	0·2376 ,,
1·8	0·0765 ,,	0·2021 ,,
2·0	0·0696 ,,	0·1946 ,,
2·2	0·0690 ,,	0·2178 ,,
2·4	0·0727 ,,	0·2632 ,,
2·6	0·0864 ,,	0·3191 ,,
2·8	0·1012 ,,	0·3848 ,,

(b) Confidence limits for 4 experiments with 0·425 per cent. monohexyl ether at 20°C.

$$V_b = \frac{V}{S(x-\bar{x})^2} \quad \text{where } V = 0{\cdot}098662 \text{ (Table VII.)}$$

$$\text{and } S(x-\bar{x})^2 = 2{\cdot}432214 \text{ (Table Va.)}$$

$$= \frac{0{\cdot}098662}{2{\cdot}432214} = 0{\cdot}040565$$

$$S(\bar{x}) = 7{\cdot}791, \text{ therefore } \bar{\bar{x}} = \frac{7{\cdot}791}{4} = 1{\cdot}948$$

The values of V_y at values of x are obtained by substituting in the equation $V_y = V_{mean} + V_b\,(x\text{-}\bar{x})^2$, where $V_{mean} = V_{\bar{\bar{y}}} = 0{\cdot}009753$ (Table XI.)

For example: at $x = 1$, $V_y = 0{\cdot}009753 + 0{\cdot}040565\,(1000 + 1{\cdot}948)^2 = 0{\cdot}046209$; $s_y = \sqrt{0{\cdot}046209} = \pm\,0{\cdot}215$.

Confidence limits (at $P = 0{\cdot}05$) have been calculated for a number of values of x. The results are set out in Table XII.

Figure 1 has been constructed (from the results in Table XII) to show graphically the limits of error (a) when 32 experiments are used to compute the mean and (b) when 4 experiments are used.

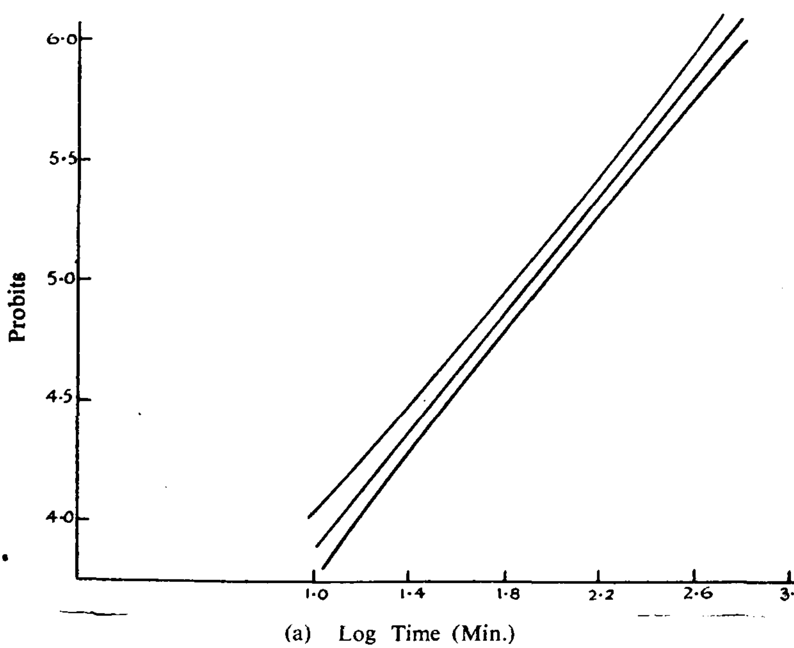

FIG. 1 (*a*).—Confidence limits (at $P = 0{\cdot}05$) of probit-log time regressions calculated from 32 expts. of 75 per cent. ethylene glycol at 20°C.

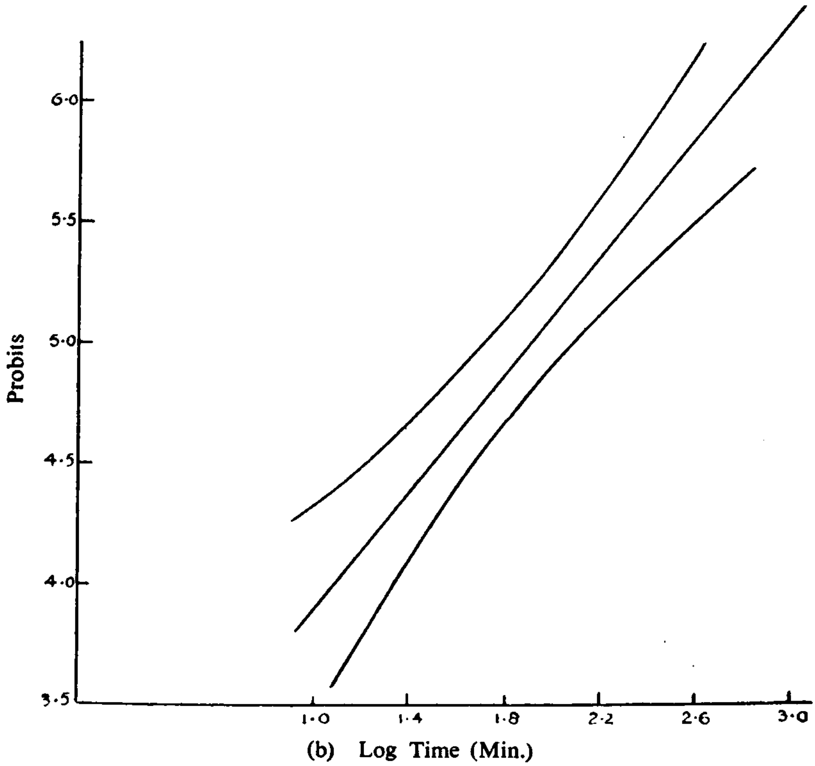

FIG. 1 (*b*).—Confidence limits (at $P = 0{\cdot}05$) of probit-log time regressions calculated from 4 expts. of 0·425 per cent. ethylene glycol monohexyl ether at 20°C.

SUMMARY

1. The course of the disinfection (at 20°C) between *Bact. coli* and several concentrations of the following ethers of ethylene glycol has been investigated: monomethyl, monoethyl, monopropyl, monobutyl and monohexyl. Several experiments were conducted at every concentration and probit-log. time regressions calculated for all experiments.

2. For every concentration of a substance the sum of squares for the common regression and for the variation in regression was calculated; the error mean square of the regression was also computed.

3. The data for every concentration of each compound have been pooled, and a mean regression has been calculated for each compound.

4. The statistical data from all the calculations for the terms of the regressions for every concentration of the compounds (at 20°C) have been massed and an analysis of variance carried out.

5. The probabilities for the differences between the mean squares of the items in the analysis of variance have been deduced by means of the *z* distribution.

6. No significant difference could be shown between the variation in regression between concentrations and between tests; these two errors

have been pooled in order to establish the error mean square for all the estimations performed.

7. The probit-log. time regression coefficient for every compound has been compared with its standard error; in all cases the ratio was large thereby indicating that b had been estimated satisfactorily.

8. Confidence limits (at $P=0{\cdot}05$) of the probit-log. time regressions have been calculated and a diagram constructed to show the increased precision obtained when many tests are performed at the same concentration.

REFERENCES

1. Berry and Michaels, *Quart. J. Pharm. Pharmacol.*, 1947, **20,** 331.
2. Berry and Michaels, *ibid.*, 1947, **20,** 348.
3. Berry and Michaels, *ibid.*, 1947, **20,** 527.
4. Berry and Michaels, *ibid.*, 1948, **21,** 24.
5. Berry and Michaels, *ibid.*, 1948, **21,** 503.
6. Mather, *Statistical Analysis in Biology,* 2nd ed., 1946, London, Methuen and Co., Ltd.
7. Fisher and Yates, *Statistical Tables for Biological Agricultural, and Medical Research,* 2nd ed., 1943, Edinburgh: Oliver and Boyd.