A TECHNIQUE FOR THE EVALUATION OF EMULSION STABILITY

By Witold Saski and Marvin H. Malone

From the College of Pharmacy, University of Nebraska, Lincoln, Nebraska, and the College of Pharmacy, University of New Mexico, Albuquerque, New Mexico

An economical sampling and statistical treatment has been developed for screening the stability of experimental fat emulsion formulations. Coalescence or progressive increase in mean globule diameter with passage of time was interpreted to indicate physical instability. The test technique attempts to efficiently extract reliable data from a relative minimum of man hours and materials per test formulation. The counting and measuring of 100–200 globules per determination was considered to be adequate.

THE search for a satisfactory method of determining the physical stability of emulsions has been of interest for many years. Sumner[1] and Becher[2] have correlated the literature on this subject. Where time is important in the evaluation of stability of a series of experimental emulsion formulations, and the cost of technical assistance must be kept nominal, no single procedure has been available that would accurately screen a number of formulations in a short time. In an attempt to formulate an intravenous emulsion of vegetable oil in an aqueous solution of glucose, a method that would be both reliable and practical was sought. Using the coalescence, or lack of it, of the fatty globules as a criterion of emulsion stability, a limited number of the dispersed globules were counted and their diameters measured at definite time intervals and a simple statistical analysis promptly applied to the data as they accumulated.

In the past it has been recommended that a large number of globules, 2,000 or more, should be counted for each determination. Such a procedure is extremely lengthy and tedious. This study set out to ascertain whether or not a much smaller number of globules could be counted without impairing the validity of the evaluation of emulsion stability.

EXPERIMENTAL

Preparation of the Emulsions

To illustrate the evaluation procedure, an emulsion of the following composition was used (Formulation 48): sesame oil, 25 per cent; polyoxyethylene sorbitan trioleate (Tween 85, Atlas), 1 per cent; 5 per cent glucose in water for injection, to 1,000 ml. Sesame oil was pre-sterilised first by heating in an oven for 90 minutes at 170°. Glucose was then dissolved in water for injection and polyoxyethylene sorbitan trioleate and sesame oil added. This mixture was run through the Eppenbach Colloid Mill at 0·005 inches clearance for 5 minutes and then autoclaved at 121·2° with a gauge pressure of 15·3 pounds (30 pounds, absolute) for 15 minutes.

Technique of the Measurements

A drop of the undiluted emulsion was placed on a microscope slide with a glass stirring rod and a small amount of dry amaranth added. After uniform distribution of the colour throughout the sample, giving a contrast between the internal and external phases of the emulsion, the slide, covered with a cover slip, was placed beneath the microscope. A few minutes for drying out was allowed. A Spencer research microscope was used with a Filar Micrometer Eyepiece, Bausch and Lomb catalogue number 31–16–50. This micrometer is equipped with a movable cross hair, the traverse of which can be measured by rotating a drum marked off in 100 subdivisions. By using the 8 mm. ($\times 20$ inital magnification) Spencer objective and the $\times 12\cdot 5$ ocular of the micrometer attachment, a final magnification of $\times 250$ is obtained.

After bringing the field into focus, an area was chosen at random between the two static lines of the micrometer and the globules were counted and their diameters measured. Accurate measurements were made by revolving the drum. For example, if upon calibration with the 8 mm. objective, one division of the micrometer was equivalent to $42\cdot5\,\mu$ and the reading on the drum was 5, then $42\cdot5 \times 0\cdot05$ or $2\cdot125\,\mu$ was the diameter of the globule.

Sampling and Statistical Evaluation

Figure 1 illustrates the sampling technique practiced upon the formulation. This work outline was used to extract the maximum amount of information from the work hours available. This procedure enables one to test the stability of one lot over a period of 3 weeks, to test the reproducibility of four lots manufactured at various intervals of time, as well as to test the reliability of the two technicians against one another.

Slide samples should be prepared and coded by a third person to insure unbiased recording. The sample coding scheme used consistently in the figures and Table I is an identification system provided for this paper and was not used to identify samples during the collection of the data. Sample identification by means of tables of random numbers is recommended. The short statistical treatment used for our daily data involved the use of an automatic calculator that could accumulate products of multiplication and a mimeographed work sheet that was prepared from the techniques outlined by Bliss and Calhoun[3]. This work sheet and summary of results is shown with an illustrative test calculation in Figure 2. This work form provides for a test of significance between the means of the two groups (row 1), an F test for comparable group variances (rows 2, 3), calculation of group mean diameters, standard deviation of group data, and standard error calculation for the group means. The format is arranged to provide means for checking the accuracy of calculations. If desired, the expected range of globule size present per sample ($P = 0\cdot05$) can be calculated as the mean diameter $\pm 2 \times$ standard deviation. In calculating the ratio of the group variances or the F test, the larger variance ($1\cdot8006$) is always in the numerator; the critical value ($P = 0\cdot05$)

for this determination can be obtained from a 2·5 per cent point table for the F distribution.

The sampling design and the statistical format were chosen to qualify and concisely summarise each day's results as soon as possible after the data had been recorded. This test design need not be fully completed before any calculation and evaluation may be made; the screening can be stopped after the first week of testing if the results are unfavourable. Manufacturing procedure or formulation or both can then be changed and a new stability screen begun.
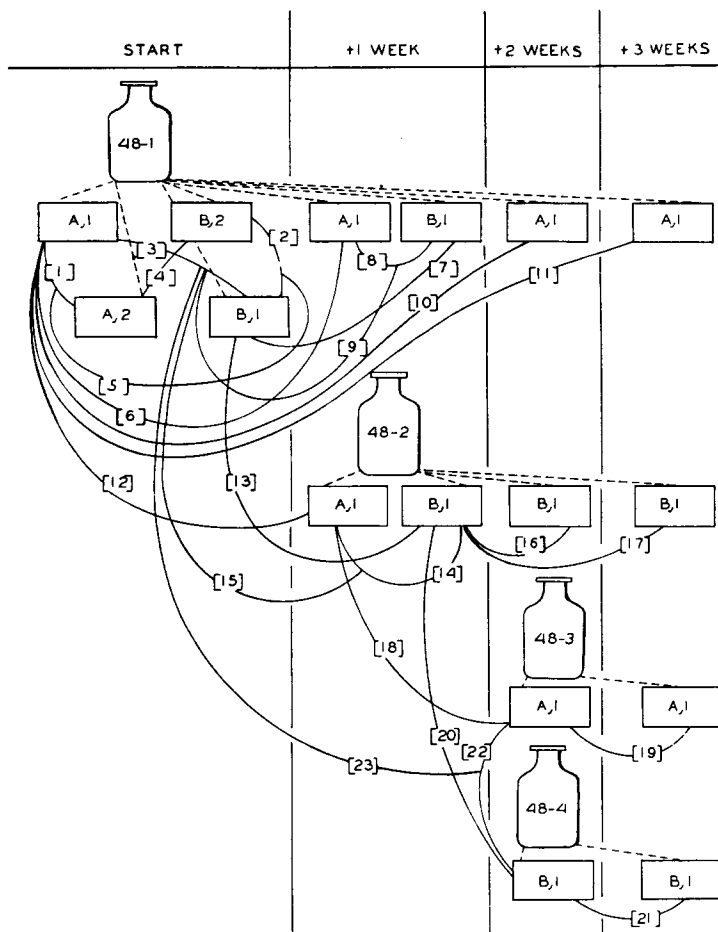

FIG. 1. Sampling scheme and test comparisons made of 4 lots of Formulation 48, with technicians A and B. The figures in brackets indicate the reference numbers for the statistical comparisons summarised in Table I.

After the results of the first day have been calculated, it is necessary to determine the optimum number of globules that should be counted per sample to detect an arbitrary 10 per cent shift in globule diameter.

TABLE I

PROGRESSIVE STATISTICAL EVALUATION OF FORMULATION 48

Fig. 1 ref. no.	Test comparison	Mean globule size, μ ±1 standard deviation	Significance between groups; observed P
1	(48–1)(A,1)(0) vs. (48–1)(A,2)(0)	4·12(195)[a] ± 1·34 vs. 4·07(203) ± 1·34	>0·50
2	(48–1)(B,1)(0) vs. (48–1)(B,2)(0)	4·22(200) ± 1·14 vs. 4·31(200) ± 1·36	0·50–0·25[b]
3	(48–1)(A,1)(0) vs. (48–1)(B,1)(0)	4·12(195) ± 1·34 vs. 4·22 (200) ± 1·14	0·50–0·25[b]
4	(48–1)(A,2)(0) vs. (48–1)(B,2)(0)	4·07(203) ± 1·34 vs. 4·31(200) ± 1·36	0·10–0·05
5	(48–1)(A,1,2)(0) vs. (48–1)(B,1,2)(0)	4·10(398) ± 1·34 vs. 4·26(400) ± 1·26	0·10–0·05
6	(48–1)(A,1)(0) vs. (48–1)(A,1)(7)	4·12(195) ± 1·34 vs. 4·24(175) ± 1·42	0·50–0·25
7	(48–1)(B,1)(0) vs. (48–1)(B,1)(7)	4·22(200) ± 1·14 vs. 4·06(200) ± 1·37	0·25–0·10[b]
8	(48–1)(A,1)(7) vs. (48–1)(B,1)(7)	4·24(175) ± 1·42 vs. 4·06(200) ± 1·37	0·25–0·10[b]
9	(48–1)(A,1:B,1)(0) vs. (48–1)(A,1:B,1)(7)	4·17(395) ± 1·24 vs. 4·14(375) ± 1·39	>0·50
10	(48–1)(A,1)(0) vs. (48–1)(A,1)(14)	4·12(195) ± 1·34 vs. 4·11(104) ± 1·26	>0·50
11	(48–1)(A,1)(0) vs. (48–1)(A,1)(21)	4·12(195) ± 1·34 vs. 4·19(100) ± 1·40	>0·50
12	(48–1)(A,1)(0) vs. (48–2)(A,1)(0)	4·12(195) ± 1·34 vs. 4·09(200) ± 1·29	>0·50
13	(48–1)(B,1)(0) vs. (48–2)(B,1)(0)	4·22(200) ± 1·14 vs. 3·87(201) ± 1·27	0·005–0·001[c]
14	(48–2)(A,1)(0) vs. (48–2)(B,1)(0)	4·09(200) ± 1·29 vs. 3·87(201) ± 1·27	0·10–0·05
15	(48–1)(A,1:B,1)(0) vs. (48–2)(A,1:B,1)(0)	4·17(395) ± 1·24 vs. 3·98(401) ± 1·28	0·05–0·025[c]
16	(48–2)(B,1)(0) vs. (48–2)(B,1)(7)	3·87(201) ± 1·27 vs. 4·01(101) ± 1·12	0·50–0·25
17	(48–2)(B,1)(0) vs. (48–2)(B,1)(14)	3·87(201) ± 1·27 vs. 4·08(100) ± 1·24	0·25–0·10
18	(48–2)(A,1)(0) vs. (48–3)(A,1)(0)	4·09(200) ± 1·29 vs. 4·20(100) ± 1·77	>0·50[b]
19	(48–3)(A,1)(0) vs. (48–3)(A,1)(7)	4·20(100) ± 1·77 vs. 4·09(100) ± 1·49	>0·50
20	(48–2)(B,1)(0) vs. (48–4)(B,1)(0)	3·87(201) ± 1·27 vs. 3·42(100) ± 1·18	0·005–0·001[c]
21	(48–4)(B,1)(0) vs. (48–4)(B,1)(7)	3·42(100) ± 1·18 vs. 3·68(100) ± 0·90	0·10–0·05[b]
22	(48–3)(A,1)(0) vs. (48–4)(B,1)(0)	4·20(100) ± 1·77 vs. 3·42(100) ± 1·18	<0·001[c]
23	(48–1)(A,1:B,1)(0) vs. (48–3)(A,1)(0) + (48–4)(B,1)(0)	4·17(395) ± 1·24 vs. 3·81(200) ± 1·55	0·005–0·001[bc]

[a] Number in parentheses indicates the number of globules counted.

[b] Significant F test calculated, P = 0·05.

[c] Group mean diameters considered to be significantly different.

These simple calculations have been outlined by Snedecor[4]. Thus 1 week later, the number of globules counted can either be increased or decreased at the desire of the investigator. For Formulation 48, the sample size should range between 100–200 globules. Counting more, decreases the efficiency; counting less, decreases accuracy.

An alternative screening procedure is indicated where a full 3 weeks' stability evaluation is routinely required for each test formulation. By restricting the sampling design to only two lots prepared concurrently, with both technicians reading one sample per each lot at the time of preparation, at +1 week, at +2 weeks, and at +3 weeks, a balanced design is produced that can be treated at the end of that period by the classical statistical techniques outlined by Cochran and Cox[5].

DISCUSSION

A formulation that does not fail during the statistical evaluation based on this flexible sampling scheme should then be considered to be worthy of a pharmacological evaluation and further stability testing utilising control chart analysis over a period of months to years. The accumulated results for Formulation 48 are summarised in Table I. The test comparisons are coded as follows: (48–2) (B, 1) (0) translates to (Formulation 48 —lot 2) (technician B made the observations, this is the first reading that he has made on this lot this day) (age of the test emulsion in days— here freshly prepared).

Within the limits of this experimental procedure, Formulation 48 displayed excellent physical stability; however, significant variations were

noted in globule size between the different lots prepared as shown by comparisons 13, 15, 20, 22, and 23 of Table I. This indicates that the reproducibility of lots by this manufacturing procedure may be a problem. The degree of efficiency achieved here with counts of 100–200 globules appears adequate for preliminary stability screening. With Formulation 48, higher counts quantitated variations that were of little practical significance. The sampling techniques used here appear to be both efficient and reliable, yet they require only a relative minimum of labour and materials per test formulation.

Group 1 Identification: (48–1)(A,1)(0)^a				vs.	Group 2 Identification: (48–1)(A,2)(0)			
Globule size, μ	Mid-range, n	Observed frequency, f	nf	n²f	n	f	nf	n²f
0–1	0·5	0	0	0	0·5	0	0	0
1–2	1·5	3	4·5	6·75	1·5	3	4·5	6·75
2–3	2·5	26	65·0	162·50	2·5	35	87·5	218·75
3–4	3·5	86	301·0	1053·50	3·5	80	280·0	980·00
4–5	4·5	39	175·5	789·75	4·5	42	189·0	850·50
5–6	5·5	19	104·5	574·75	5·5	24	132·0	726·00
6–7	6·5	13	84·5	549·25	6·5	11	71·5	464·75
7–8	7·5	8	60·0	450·00	7·5	6	45·0	337·50
8–9	8·5	1	8·5	72·25	8·5	2	17·0	144·50
9–10	9·5	0	0	0	9·5	0	0	0
10–11	10·5	0	0	0	10·5	0	0	0
Column totals:		195	803·5	3658·75		203	826·5	3728·75
Designations:		N_1	$S(x_1)$	$S(x_1^2)$		N_2	$S(x_2)$	$S(x_2^2)$
Mean diameter:	$S(x_1)/N_1 = 4·12$				$S(x_2)/N_2 = 4·07$			

Row	Term	D.F.^{b}	Sum of squares	Mean squares	F	Observed P
1	Between groups	1	$[x^2]g = (C_1 + C_2) - C = 0·240$	$A = [x^2]g/1 = 0·240$	$A/s^2 = 0·133$	$> 0·50$ $(<3·84)^c$
2	Within group 1	$N_1 - 1 = 194$	$[x_1^2] = S(x_1^2) - C_1 = 347·918$	$s_1^2 = [x_1^2]/N_1 - 1 = 1·7934$	$s_1^2/s_2^2 = 1·0039$	$> 0·50$ $(< 1·20)^c$
3	Within group 2	$N_2 - 1 = 202$	$[x_2^2] = S(x_2^2) - C_2 = 363·714$	$s_2^2 = [x_2^2]/N_2 - 1 = 1·8006$		
4	Within groups	$N - 2 = 396$	$[x^2] - [x^2]g = [x_1^2] + [x_2^2] = 711·632$	$s^2 = [x^2]-[x^2]g/N-2 = 1·7970$		
5	Total	$N - 1 = 397$	$[x^2] = S(x^2) - C = 711·872$	$C_1 = T_1^2/N_1 = 3310·832$		
6	Corr.	1	$C = T^2/N = 6675·628$	$C_2 = T_2^2/N_2 = 3365·036$		

$T = T_1 + T_2 = S(x_1) + S(x_2) = 803·5 + 826·5 = 1630·0$
$S(x^2) = S(x_1^2) + S(x_2^2) = 3658·75 + 3728·75 = 7387·50$
Calculations check: $[x_1^2] + [x_2^2] + [x^2]g = [x^2] = 711·872$
Standard deviation: $\sqrt{s_1^2} = s_1 = 1·3392$; $\sqrt{s_2^2} = s_2 = 1·3418$
Standard error of the mean: $s_1/\sqrt{N_1} = 0·0959$; $s_2/\sqrt{N_2} = 0·0942$

$N_1 + N_2 = N = 398$

^{a} Groups are coded for speed and clarity of identification; translation: (Formulation 48, lot 1) (globules counted by person A; this is his first test sample at this time) (age of the emulsion in days; here the emulsion is freshly prepared).
^{b} D.F. designates degrees of freedom.
^{c} Figure in parenthesis is the critical value, $P = 0·05$.

FIG. 2. Concise work sheet for statistical treatment with illustration for use.

REFERENCES

1. Sumner, *Clayton's The Theory of Emulsions and Their Technical Treatment*, 5th Edn, J. A. Churchill Ltd., London, 1954.

2. Becher, *Emulsions: Theory and Practice*, Reinhold Publishing Corp., New York, 1957.
3. Bliss and Calhoun, *Outline of Biometry*, Yale Co-op Corp., New Haven, Conn., 1954, p. 87.
4. Snedecor, *Statistical Methods*, 5th Edn, Iowa State College Press, Ames, Iowa, 1956, p. 275.
5. Cochran and Cox, *Experimental Designs*, John Wiley and Sons, Inc., New York, 1950, p. 91.