DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 10(8&9), 1449-1461 (1984)

GANG STABILITY PROGRAMS FOR MARKETED BATCHES

Kenneth Thompson

Manager, Chemical Analytical Services
Quality Control
Ayerst Laboratories Inc.
1025 Laurentain Blvd.
Ville St. Laurent
Quebec, Canada H4R-1J6

In any pharmaceutical company the number of product/packaging combinations can be large and, with few exceptions, all of these have to be monitored for stability. In addition, stability testing is required because of product or packaging changes stemming from product improvement goals, new compendial requirements, cost reduction programs, new sources or grades of raw materials or packaging materials, marketing decisions, and the introduction of new products. Thus a stability program can become large, complex and costly, and it is necessary, periodically to reassess the adequacy and relevancy of the program.

The stability program for marketed batches of Ayerst Canada's Quality Control Department includes over 200 products, many in more than one packaging system. For these,

 0363-9045/84/1008-1449$3.50/0

about 2000 batches are on stability and, on the average five parameters are monitored for each batch. This results in about 10,000 stability tests each year or 200 each week and, as there are about 250 working days in the year, an annual review of all products requires that one product be reviewed almost every working day.

In 1979 the stability section reassessed the adequacy of its stability program and introduced a significant modification. Details of the old and new programs are provided below.

OLD STABILITY PROGRAM

Up until early 1979 the routine program for marketed batches was operated according to the industrial norm, i.e. batches on stability were tested at fixed intervals (e.g. 6, 12, 24, 36, 48, 60 months) from the date of manufacture, generally in isolation from other batches of the same product or product family. This program, which we now refer to as the Fixed Interval program, required complicated operational logistics, for example;

- Each week, test forms had to be prepared for many different products, rarely for more than one batch of a given product at a time. This was required several times each year for each product, depending on the number of batches on stability.

- Each week, stability samples from many different products had to be retrieved from the retention area and distributed to the laboratories.
- The testing laboratories were locked into specified future retest dates which were not always convenient when they arrived if, for example, work load from other sources was high.
- If stability testing was delayed for one or two months, the age of the batch at the actual time of testing did not coincide with any of the fixed interval ages that the computer was programmed to accept. Whenever this occured it introduced some distortion into the data base as, for example, data obtained on a batch for which the testing was delayed from 12 months until 14 months had to be entered as being tested at 12 months.
- Each week, the data resulting from the testing of many different types of samples had to be reviewed by the stability section, entered into the computer and filed.

NEW STABILITY PROGRAM

Under this program, once each year all stability batches of a product are gang-tested on a selected date, irrespective of their age at the time of testing. This program, which we now refer to as the Selected Date program, has greatly simplified operational logistics as:

- All test forms required for a product for the year are prepared at the same time.
- All stability samples for a product are retrieved from the retention area and sent to the laboratories only once each year.
- The stability test scheduling is flexible and can be changed if other work load is high.
- The actual age of the batch at the time of testing can be entered into the computer.
- The years test results for a product are received simultaneously by the stability section, which provides a stimulus and a discipline for the prompt review of data and the regular issuance of reports.

COMPARISON OF THE PROGRAMS

Data Distribution

With the Fixed Interval program, on a typical test occasion only one new result became available (Fig. 1a). Consequently, the comparison that could be made with old data was very limited, for example, was the new result within specifications?, was it greatly different from the old data?

With the Selected Date program, on a typical test occasion many new results become available at the same time (Fig. 1b). The new data set can be readily compared with the old data for slope, intercept, data variation, trends and

aberrant results. The typical new data set shown in Fig. 1b indicates, even without further processing, that the new data confirm the stability pattern of the old data.

Data Variation

The Selected Date program has resulted in reduced data variation. This is because with gang-testing, the reagents, standards, instrument conditions, analyst proficiency, etc. are constant for all the batches being tested at the annual testing time.

The data variation of the two programs for niacinamide in a multivitamin preparation is compared in Fig. 2. The zone between the upper and lower tolerance bounds (i.e. the zone within which 95% of individual assays should fall) is about 30% smaller for the data gathered under the Selected Date program, than the same zone for the Fixed Interval program. Similarly, the zone between the upper and lower confidence bounds (i.e. the zone within which the regression line should fall, with 95% confidence) is smaller for the Selected Date program than for the Fixed Interval program.

Reduced data variation has been found to be typical of the Selected Date program, as is shown in Fig. 3, which compares the relative spread between the tolerance bounds, and between the confidence bounds for data gathered under the two programs on a series of vitamins in a multivitamin product. In the example shown in Fig. 3, the reduced variation was

found for automated chemical methods for niacinamide, riboflavin, pyridoxine and thiamine, for a manual chemical method for vitamin C and for a microbiological method for vitamin B_{12}.

OPERATION OF THE SELECTED DATE PROGRAM

The following are the essential steps in the operation of the Selected Date Stability program;

- Each year, select batches of each product to go onto the stability program.
- Once a year or if necessary more frequently, select a date compatible with laboratory work load and data needs, for the stability testing of a product.
- On that date simultaneously test, selected stability batches of the product irrespective of their age.
- Superimpose the results on a plot of the regression line and confidence and tolerance bounds of the older data. When possible express the data as percent claim, as this facilitates comparisons between different strengths of the same product.
- Visually assess the new data set for bias. If there is an obvious positive or negative bias as shown in Figs. 4a and 4b, look for a source of systematic error such as an incorrectly prepared reagent, a deficient instrument condition, a new or inexperienced analyst etc. A skewed relationship between the new and old data as shown on

Fig. 4c, could indicate that more recent batches are less stable than older batches, and further investigation is necessary. To date, skewness has been observed only once in over 2,000 data sets examined. For 95% of new data sets no obvious bias was detected, and data distribution resembled that shown in Fig. 4d. Even if bias is not obvious, it is still necessary to test for it statistically. Do this with sign tests (two sided binomial tests of $p = 0.5$) at the 0.10 level of significance. Two tests are used. With the first test, compare the proportion of the new data set above the upper 95% confidence bound for the historical regression line, with that below the lower 95% confidence bound. With the second test, compare the proportion of the new data set above the regression line for the historical data, with that below the regression line. From binomial tables assess the significance of these proportions. If the new data set passes both sign tests add it to the data base.

If bias is present to any appreciable degree, identify and if possible eliminate the source of bias. If this is not possible, retest, not employing gang testing, but assaying each batch separately over a period of time.

To date we have observed bias in only 5% of data sets obtained under this program. Occasionally, although no bias is detected, an individual result may be aber-

rant. Retest and, if the aberrance is confirmed, investigate to see if the batch has an atypical stability pattern.

- Pool the unbiased data with the historical data base and calculate individual batch slopes and intercepts.
- Use an F test to determine if there are statistically significant differences between batches in either slope or intercept. Each time a variable such as individual slopes or individual intercepts is to be eliminated from a stability model, a comparison of the precision of the model with the variable, to the precision of the model without the variable must be made by use of the F statistic. A significant F statistic indicates that the given variable has a significant effect on the estimated variable and should remain in the model. Conversely, an insignificant F statistic indicates that the given variable can be eliminated.

In applying the F statistic to a stability model, first assume the active ingredient concentration in a product can be attributed to both sample age and to individual batches. Estimate the active ingredient concentration by the linear model:

$$Y_i = \alpha_i + \beta_i X \qquad (1)$$

where:

Y_i is active ingredient concentration of batch i;

X is the age;

α_i is the intercept of batch i;

β_i is the slope of batch i;

Assume different slopes (β_i) and intercepts (α_i) for each batch. Should the F test indicate that batch to batch variation of slopes is not significant, assume a commone slope (β) for all batches and reduce model (1) to:

$$Y_i = \alpha_i + \beta X \quad (2)$$

Should the F test indicate that batch to batch variation of intercepts is not significant, assume a common intercept (α) for all batches and reduce model (2) to:

$$Y_i = \alpha + \beta X \quad (3)$$

TYPES OF STABILITY PATTERNS

The stability data will indicate one of the following patterns depending on which mathematical model is applicable:

Different Intercepts and Different Slopes

To date, about 5% of data bases have shown significant batch to batch variation in both slope and intercept. Pro-

ducts having this type of stability pattern have a regression analysis similar to that shown on Fig. 5.

Different Intercepts and Common Slope

To date, about 35% of data bases have shown batch to batch variation in intercept but a common slope. Products having this type of stability pattern have a regression analysis similar to that shown on Fig. 6.

Common Intercept and Common Slope

To date, about 60% of data bases have shown no batch to batch variation in either slope or intercept. Products having this type of stability pattern have a regression analysis similar to that shown in Fig. 7.

Batch to batch variation in intercept can be expected in some instances, and is usually attributable to different lots of raw materials, errors in raw material assay, manufacturing variation between batches, variable extractability of the active ingredient from the finished product etc.

Batch to batch variation in slope is more important and warrants further investigation, especially if the most recent batches are less stable. We always review individual batch slopes, even if the F test does not indicate significant variation in slope. Variation in slope is usually attributable to changes in the source or grade of raw materials, active or inactive, changes in the manufacturing process or changes in the packaging materials or systems. Whenever a

common slope is found for a data base including data from batches with and without a product change, it is a good indication that the change had no significant effect on the product stability.

The estimated shelf-life and any overage modifications are calculated using the lower bound which gives the required level of assurance of conformance to specifications, i.e. the confidence or tolerance bound for either, the average of the product, individual batches, or individual assays.

RESERVE SAMPLE REVIEW

In addition to testing stability samples, we also examine reserve samples of various ages to determine if there are any changes in appearance indicative of stability problems. Fig. 8 shows a form suitable for recording the results of such a review.

STABILITY REPORT

When all testing is completed we issue an in depth report, commenting on:

- The estimated conformance periods for all active ingredients;
- Any adjustment in overage, which may be required;
- The overage required to compensate for stability loss and/or manufacturing loss;
- The conformance periods of other attributes such as pH, disintegration, dissolution, moisture, etc.;

- The specificity of the analytical methods;
- The adequacy of the present stability schedule;
- The findings of the reserve sample review.

Once the stability of a product is well characterized we limit future reports to a one-page summary of the type shown on Fig. 9, unless the product or its stability have significantly changed.

APPLICABILITY

Gang-testing under the Selected Date program is applicable to any well characterized product, irrespective of whether it is stable or not. For new or revised products, where there are less data and less assurance of conformance, we schedule stability testing more frequently to build up the data base more rapidly. If we discern an unexpected stability trend for an established product, we test more extensively.

EXPERIENCE WITH THE PROGRAM

To date we have issued over 500 reviews on 200 products. These reviews have enabled use to reduce over 100 overages with concomitant savings in active raw materials, without compromising product quality. In addition, gang-testing has reduced our stability testing costs by about 30% and the corresponding stability section work load.

ACKNOWLEDGEMENTS

I wish to thank the many people of the Ayerst Quality Control Department who have encouraged, aided and guided me.

The following in particular should be mentioned: Mr. W.W. Charnutzky, Director of Quality Control; Mrs. C. Oliver, Supervisor of Stability; Ms. H. Prastalo, Data Analyst; Mrs. L. Gauthier, Statistician; Dr. J.R. Soucy, Director of Quality Assurance; and Dr. A.J. Smith and Mr. D.J. Harrison of Ayerst U.S.