PHARMACEUTICAL
TECHNOLOGY®

Ruggedness and Robustness

with Designed Experiments

by Lynn D. Torbeck

 AN ADVANSTAR PUBLICATION Printed in U.S.A.

Assay Validation:

Ruggedness and Robustness with Designed Experiments

Ruggedness and robustness are two similar analytical parameters of test methods that are either required or suggested for validation by *USP 23* and the tripartite ICH texts on validation of analytical procedures. This article describes an easy-to-use procedure to collect and analyze data to meet these requirements.

Lynn D. Torbeck

Mr. Lynn Torbeck *is president of Torbeck Associates, 7812 Kenneth Street, Skokie, IL 60076-3506, tel. (847) 677-8070, fax (847) 677-8073.*

Companies are not required by GMPs to validate analytical methods to USP or ICH standards, but these standards do form a benchmark for comparison (1–3). Many laboratories may look to these documents for guidance in validation of analytical methods. Two of the analytical parameters required or referenced are ruggedness and robustness.

RUGGEDNESS AND ROBUSTNESS

USP defines ruggedness as "the degree of reproducibility of test results obtained by the analysis of the same samples under a variety of normal test conditions, such as different laboratories, different analysts, different instruments, different lots of reagents, different elapsed assay times, different assay temperatures, different days, etc." Notice that the factors mentioned are external to the written method. The measurement of ruggedness is usually expressed as "a lack of influence on test results." It is required for all four USP categories of methods: I, II quantitative, II limit, and III. Note that ruggedness does not appear in the ICH document.

Robustness is defined by both USP and ICH as "a measure of its capacity to remain unaffected by small but deliberate vari-

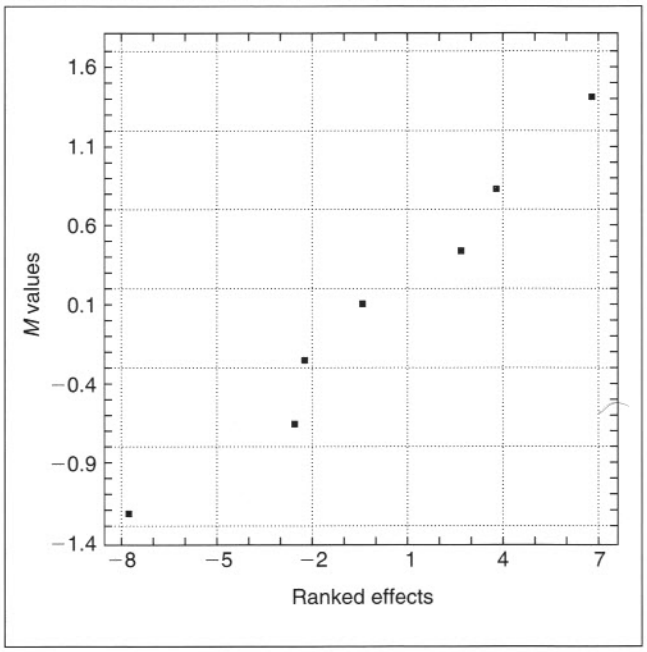

Figure 1: *Normal probability plot of effects.*

Table I: An eight-run design.

	A	B	C	D	E	F	G	Test Result
1	+	+	+	−	+	−	−	98.13
2	−	+	+	+	−	+	−	110.57
3	−	−	+	+	+	−	+	107.83
4	+	−	−	+	+	+	−	95.87
5	−	+	−	−	+	+	+	97.23
6	+	−	+	−	−	+	+	91.07
7	+	+	−	+	−	−	+	99.33
8	−	−	−	−	−	−	−	99.97

Table II: Ranked effects and means.

Factors	Ranked Effects	*M* Values
A	−7.80	−1.35
F	−2.63	−0.76
G	−2.27	−0.35
E	−0.47	0
B	+2.63	+0.35
C	+3.80	+0.76
D	+6.80	+1.35

ations in method parameters and provides an indication of its reliability during normal usage." Here the factors are internal to the written method. As with ruggedness, robustness is measured as a lack of influence on test results. In both USP and ICH, robustness is defined but not required.

Given their similarity, ruggedness and robustness can be evaluated with the same set of activities. Although ruggedness and robustness may be tested by varying one factor at a time and holding all other factors fixed, it is much more cost-effective, efficient, and informative to use a multifactor-designed (matrix) experiment. The savings in time and materials are dramatically greater.

DESIGNED EXPERIMENTS

Design of experiments has a long history in biology, chemistry, pharmaceuticals, and medical devices (4). Most double-blind clinical trials as required by FDA are statistically designed experiments. Designed experiments are the most logical, rational, and scientific way known for collecting data.

For assay validation, designed experiments can demonstrate that method and environmental factors do not influence the results. They may also identify conditions that need to be more closely controlled or investigated.

PLACKETT–BURMAN DESIGNS

Many types of designed experiments can be used to accomplish this objective. The most commonly used ones are Plackett–Burman designs of 8 or 12 runs (5–12). An 8- or 12-run design (matrix) will permit up to 7 or 11 factors to be investigated. Designs with as many as 100 runs exist, but usually 11 factors in 12 runs are more than enough. These designs have two levels, values, or versions per factor. They vary the combinations of levels in a very specific and highly symmetric way. Because of the symmetry of the designs, each factor effect can be estimated using all of the data collected. For an 8-run (8 data values), 7-factor design, this is the equivalent of doing 56 one-factor-at-a-time data points. For a 12-run, 11-factor design, each of the 11 factors is estimated using all 12 data values or the equivalent of doing 132 one-factor-at-a-time data points. This is the source of their efficiency.

THE PROCEDURE

- Define the responses, the dependent variables, to be measured by the method. More than one variable can be measured in the same experiment. This is a major advantage.
- List the factors, conditions, or independent variables to be investigated. Assign the factors to the columns labeled A–G. Not all columns must be used. For ruggedness, these are factors external to the written method. For robustness, the factors are internal to the written method. They can be combined.
- Define a low and high value, level, or version for each factor. Often these are the specifications for the factor. For discrete factors, such as analyst or equipment, high and low may be an arbitrary designation.
- Replace the ± signs in the design shown with the high (+) and low (−) values.
- Conduct the experimental runs, if possible, in random order.
- With the data collected, calculate the factor effects for each response by attaching the signs in the columns of the design to the data collected, adding and dividing by half the number of runs. For example, in Table I, G effect = (−98.13 −110.57 +107.83 −95.87 +97.23 +91.07 +99.33 −99.97)/4. G effect = −2.27.

Table III: A 12-run design.

	A	B	C	D	E	F	G	H	I	J	K
1	+	+	−	+	+	+	−	−	−	+	−
2	−	+	+	−	+	+	+	−	−	−	+
3	+	−	+	+	−	+	+	+	−	−	−
4	−	+	−	+	+	−	+	+	+	−	−
5	−	−	+	−	+	+	−	+	+	+	−
6	−	−	−	+	−	+	+	−	+	+	+
7	+	−	−	−	+	−	+	+	−	+	+
8	+	+	−	−	−	+	−	+	+	−	+
9	+	+	+	−	−	−	+	−	+	+	−
10	−	+	+	+	−	−	−	+	−	+	+
11	+	−	+	+	+	−	−	−	+	−	+
12	−	−	−	−	−	−	−	−	−	−	−

Table IV: *M* values and ranked effects for a 12-run design.

Factors	Ranked Effects	*M* Values
A	_____	−1.59
F	_____	−1.06
G	_____	−0.73
E	_____	−0.46
B	_____	−0.22
C	_____	0.00
D	_____	+0.22
E	_____	+0.46
F	_____	+0.73
G	_____	+1.06
H	_____	+1.59

- Rank the effects from the smallest to the largest as in Table II.
- Plot the effects on linear–linear graph paper by hand or with a spreadsheet plot (see Figure 1). Place the effects on the horizontal axis and the values in column *M* on the vertical axis. This is a normal probability plot.

The values in column *M* in Table II are the means of the order statistics (13) for a sample size of seven. Note that for a given design, the values for *M* will always remain the same and can be used for any method validation.

INTERPRETATION

The above example illustrates the approach. The factors A, B, C, ... etc., may be temperature, stirring time, dilution levels, additions of materials, instruments, analysts, vendors, batches, or equilibration times. If the method is rugged, the calculated factor effects will be normally distributed random noise, and the normal probability plot will form a straight line.

The effects in Figure 1 form a nearly straight line. One can conclude that the method is rugged for the factors over the ranges tested. If one or more points on the ends of the line clearly do not lie on the line, one can conclude that the method is not rugged or robust for those factors. Those effects lying off the ends of the line would be investigated for improvement or change of specification after they were considered to be of practical importance by an experienced chemist.

12-RUN DESIGN

A 12-run design can be used when more than seven factors are to be considered. Follow the steps as before and identify up to 11 factors for the experiment. These can again be a combination of internal and external factors. The design matrix is shown in Table III. The *M* values are shown in Table IV.

METHOD TRANSFER

This approach can also be used to demonstrate the transfer of a new method to another laboratory. If both labs can do the same experiment and get essentially the same results, the method is successfully transferred.

REFERENCES

1. "Validation of Compendial Methods," *USP 23,* section ⟨1225⟩ (United States Pharmacopeial Convention, Rockville, MD, 1994), pp. 1982–1984.
2. International Conference on Harmonisation, "Guideline on Validation of Analytical Procedures: Definitions and Terminology; Availability." (*Federal Register,* 1 March 1995).
3. International Conference on Harmonisation, "Extension of the ICH Test: Validation of Analytical Procedures," draft number 7, 28 July 1995.
4. J.F. Box, *R.A. Fisher: The Life of a Scientist* (John Wiley & Sons, New York, 1978).
5. ASTM, *Standard Guide for Conducting Ruggedness Tests, E1169* (American Society for Testing and Materials, Philadelphia, 1989).
6. R.C. Paule, "Ruggedness Tests and Interlaboratory Studies," *ASTM Standardization News,* May 1988.
7. R.C. Paule et al., "Ruggedness Testing, Part I: Ignoring Interactions," *J. Res. Natl. Bur. Stand.* **91-1**, 3–8 (1986).
8. R.C. Paule et al., "Ruggedness Testing, Part II: Recognizing Interactions," *J. Res. Natl. Bur. Stand.* **91-1**, 9–15 (1986).
9. G. Wernimont, "Ruggedness Evaluation of Test Procedures," *ASTM Standardization News* **5** (3), 13–16 (1977).
10. G. Wernimont, "Evaluation of the Ruggedness of an Analytical Process," in *Use of Statistics to Develop and Evaluate Analytical Methods*, W. Spendley, Ed. (AOAC, Arlington, VA, 1985), pp. 78–82.
11. W.J. Youden and E.H. Steiner, "Ruggedness Test for Procedures," in *Statistical Manual of the Association of Official Analytical Chemists* (AOAC, Arlington, VA, 1975), pp. 33–36.
12. G. Kateman and L. Buydens, "The Ruggedness Test," in *Quality Control in Analytical Chemistry* (John Wiley & Sons, New York, 2d ed., 1993), pp. 118–125.
13. E.S. Pearson and H.O. Hartley, *Biometrika Tables for Statisticians* (Biometrika Trust, University College, London, 1976), pp. 29–32, 205–210.
14. G.E. Box, W.G. Hunter, and J.S. Hunter, *Statistics for Experimenters* (John Wiley, New York, 1978), p. 329.