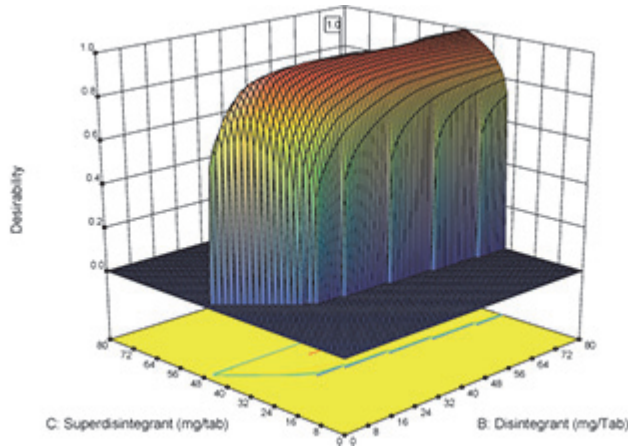


How Design of Experiments Can Improve Formulation Development

Jerry Fireman, Structured Information



Mixtures can be defined as a combination of ingredients where the response is a function of the proportion, rather than the amounts, of the ingredients. Formulation development often boils down to determining the optimum combination of ingredients in a mixture, which can make the difference between success and failure in many diverse fields of research, such as materials, pharmaceuticals,

adhesives and coatings. The traditional approach to experimentation changes only one process factor at a time (OFAT) or one component in a formulation. However, with this approach, it's easy to overlook interactions of factors or components, a likely occurrence in developing formulations.

Statistically-based design of experiments (DOE) provides validated models, including any significant interactions, that make it possible to confidently predict response measures as a function of the inputs. The payoff is the identification of 'sweet spots' where you can achieve all product specifications and processing objectives. Industrial experimenters typically turn to two-level factorials as their first attempt at DOE. These designs consist of all combinations of each factor at its high and low levels. With large numbers of factors, only a fraction of the runs needs to be completed to produce estimates of main effects and simple interactions. However, when the response depends on proportions of ingredients, such as in chemical or food formulations, factorial designs don't work well because they focus on the absolute amounts of the ingredients while it's the proportions that count in mixtures.

ID	Point Type	Blend Type	Gold wt fraction	Copper wt fraction	Melt Point Deg C
1	Vertex	Pure	0.00	1.00	1063
2	"	"	0.00	1.00	1083
3	Axial Check Blend	Quarter	0.25	0.75	955
4	Centroid	Binary	0.50	0.50	926
5	"	"	0.50	0.50	921
6	Axial Check Blend	Quarter	0.75	0.25	952
7	Vertex	Pure	1.00	0.00	1049
8	"	"	1.00	0.00	1036

To begin to explain how DOE can optimize a formulation, let's look at the example of how goldsmiths from ancient times have mixed gold with a small amount of copper to create a lower melting point solder that allowed them to connect intricately designed wire to the backbone of bracelets and necklaces.¹ Even though copper melts at a higher temperature

than gold, when mixed together, these two metals melt at a lower temperature than either one alone. One could never predict this beneficial combination of ingredients without actually mixing them together for experimental purposes.

The experiment described in the above table was performed to determine the temperature at which various mixtures of copper and gold begin to melt. The input values are expressed on a coded scale of zero to one, which statisticians prefer for modeling mixtures. The replication designated in the descriptor columns by ditto marks provides a measure of pure error that quantifies the inevitable variations in blending the materials and measuring the responses. The equation below was fitted from the experimental data by using least squares regression to plot the predicted response of any given composition of a gold-copper mixture. It models the melt point as a function of the two ingredients, gold and copper, symbolized by x_1 and x_2 respectively.

$$\text{Melt point} = 1044 x_1 + 1071 x_2 - 543 x_1 x_2$$

This mixture model, developed by Henri Scheffé (1958), is derived from the second order polynomial for process response surface methods (RSM), also known as a quadratic equation. The mathematical details are spelled out in the accompanying reference.² Two things distinguish Scheffé's polynomial from that used for RSM. First, there is no intercept. Normally this term represents the response when factors are set to zero — set by standard coding to their midpoints for process modeling. However, the constituents of the mixture are coded on a zero-to-one scale so it doesn't make sense to set all components to zero. Although this experiment requires the control of two inputs — gold versus copper, only one X axis is needed on the response surface graph because of the complete inverse correlation of one component with the other. Another difference from RSM is that the formulation equation lacks squared terms. This is because the $x_1 x_2$ term captures the non-linear blending behavior.

Here are some general guidelines for setting up a formulation experiment and analyzing the results, starting with the Scheffé equations for predicting the response from two components.

$$\text{First order (linear): } \hat{y} = \beta_1 x_1 + \beta_2 x_2$$

$$\text{Second order (quadratic): } \hat{y} = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

The hat (^), properly known as a circumflex, over the letter y symbolizes that we want to predict this

response value. The β (beta) symbols represent coefficients to be fitted via regression.

DOE software has the potential to eliminate the need for statistical expertise on the part of the users by walking the user through the complete process. For example, the software prompts the user to enter the factors and responses and select the type of experiment while providing information that will help the user pick the best type. The software will then generate a randomized list of experimental runs. As each run is completed in the order given, the results are entered into the software. The software then generates tabular and graphical data that helps define the region where quality product is produced.

As an example of how these methods are used in the real world, VerGo Pharma Research Laboratories Pvt. Ltd was recently hired by a generic pharmaceutical manufacturer to develop a bioequivalent with different polymorphic forms for an anti-depressant drug that had been patented in crystalline form only. Bioequivalence requires that a drug be pharmaceutically

equivalent and that it be delivered at the same rate and same level of bioavailability so that its efficacy and safety can be expected to be the same as the original product. Using conventional one-factor-at-a-time testing methods, it would have taken several years to determine the right combination of inactive ingredients to achieve the required in-vitro dissolution and in-vivo plasma drug profile. VerGo compressed this development process to only four months by using Design-Expert software for DOE to reduce the number of tests required to determine the effects of inactive ingredients on bioavailability in both fed and fasting conditions.

Results of experiment evaluating tablet dissolution									
Std	Run	Factor 1 A.Diluent mg/tab	Factor 2 B.Disintegrant mg/Tab	Factor 3 C.Superdisintegrant mg/tab	Response 1 Disso-1.6- 20min %	Response 2 Disso-1.6- 45min %	Response 3 Disso-1.6- 90min %	Response 4 Disso-5.5- 20 min %	Response 5 Disso-5.5- 45 min %
9	1	0	0	28.8	71	79	84	50	57
8	2	97.7	72.0	8.0	60	73	79	54	59
12	3	55.8	35.4	44.8	80	90	94	75	77
7	4	174.6	40.0	6.8	67	80	84	61	67
14	5	194.0	32.0	47.6	83	93	96	66	73
11	6	121.3	0	34.4	67	79	85	46	52
18	7	0	51.6	80.0	91	94	96	78	82
17	8	0	51.6	80.0	91	94	96	74	80
5	9	0	80.0	0	45	57	67	51	57
19	10	194.0	80.0	80.0	94	101	103	77	84
2	11	194.0	0	0	33	44	51	17	22
15	12	125.1	0	80.0	84	89	92	73	77
1	13	194.0	0	0	54	63	69	32	39
10	14	0	0	28.8	61	72	79	42	50
6	15	194.0	80.0	0	68	83	87	52	60
20	16	194.0	80.0	80.0	90	97	99	74	81
4	17	0	80.0	0	57	68	77	55	62
16	18	116.4	48.3	80.0	89	97	100	85	89
13	19	77.6	80.0	47.6	77	83	86	68	73
3	20	77.6	32.0	0	50	63	71	40	48

The software selected values for a total of 20 runs with the diluent ranging from 0 to 194 milligrams

per tablets and the two disintegrating agents ranging from 0 to 80 mg per tablet. The experiment included 5 replicates which were used to measure the reproducibility of the results.

After running the experiments, Subrata Kundu, Principal Scientist, Formulation Development for VerGo, entered the results into the DOE software along with the ideal values for the dissolution rate at each pH value/time point pair. These dissolution rate values were selected to match the values achieved by the original drug based on the assumption that if the proposed generic performs the same as the original drug in the lab it is likely to also perform the same in clinical testing. The software generated a prediction of the concentration of each variable required to meet all of the target dissolution values. VerGo's scientists then prepared a new batch of tablets with the recommended concentration values. These tablets matched the desired dissolution profile within +/- 5%, which is within the acceptable margin of error.

VerGo then prepared a larger batch of tablets with this formation for use in clinical testing with volunteer patients. The patients took the drugs in both fed and fasting conditions, and the concentration of the drug in their blood was measured at set intervals. The results showed that VerGo was the first company that was able to match the blood concentration levels of the active ingredient to the original pharmaceutical over the full time profile within an acceptable margin of error. The application provides a good example of how DOE can compress the development process for formulation development by identifying potential effects caused by interactions between ingredients.

- For more information about Design-Expert software for DOE, visit www.statease.com [1]

References

1. Mark Anderson and Pat Whitcomb, "A Primer on Mixture Design: What's In It for Formulators? [2]," Stat-Ease, 2009
2. John Cornell, "Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data [3]," Wiley, 2002.

Jerry Fireman is President of Structured Information.