




Bristol Myers Squibb Webinar
for Charu Chaudhry and Friends






Pat Whitcomb
Stat-Ease, Inc.
pat@statease.com


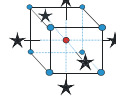



Shari Kraber
Stat-Ease, Inc.
shari@statease.com

Bristol Myers Squibb 1

Introduction to Design of Experiments
via three AAO Biomek® FX case studies




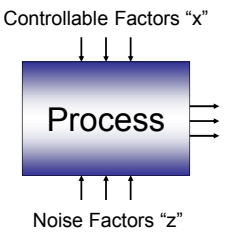




- Factorial design
- Mouse Cell Assay
- Response Surface design
- Enzyme assay optimization
- Mixture design
- Low Cost Media (Candida Bombicola)

Bristol Myers Squibb 2

Design of Experiments






DOE (Design of Experiments) is:
"A systematic series of tests, in which purposeful changes are made to input factors,

Responses "y"
so that you may identify causes for significant changes in the output responses."

Bristol Myers Squibb 3

DOE versus OFAT



Traditional Approach to Experimentation


- Study one factor at a time (OFAT), holding all other variables constant
- Simple process, but doesn't account for interactions
- It is inefficient (serial processing)


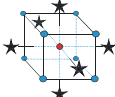

Design of Experiment Approach

- Study multiple factors changing at once (parallel processing)
- Accounts for interactions between variables
- Maximize information with minimum runs

Bristol Myers Squibb 4

Agenda Transition




- **Factorial design**
 - Mouse Cell Assay
- Response Surface design
 - Enzyme assay optimization
- Mixture design
 - Low Cost Media (Candida Bombicola)

Bristol Myers Squibb 5

Motivation for Factorial Design



Would like to:

- Estimate factor effects well;
 - This implies estimating effects from averages.
- Obtain the most information in the fewest number of runs.
- Estimate each factor effect independent of the existence of other factor effects.
- Keep it simple.

Bristol Myers Squibb 6

Two-Level Full Factorial Design

Run all high/low combinations of 2 (or more) factors
Use statistics to identify the critical factors

2^2 Full Factorial

$$\text{Effect}(\Delta y) = \frac{\sum y_+}{n_+} - \frac{\sum y_-}{n_-}$$

What could be simpler?

Bristol Myers Squibb 7

Design Construction

2^3 Full Factorial

Std	A	B	C	AB	AC	BC	ABC	
1	-	-	-	+	+	+	-	y_1
2	+	-	-	-	-	+	+	y_2
3	-	+	-	-	+	-	+	y_3
4	+	+	-	+	-	-	-	y_4
5	-	-	+	+	-	-	+	y_5
6	+	-	+	-	+	-	-	y_6
7	-	+	+	-	-	+	-	y_7
8	+	+	+	+	+	+	+	y_8

Bristol Myers Squibb 8

Relative Efficiency

Factorial versus OFAT

Relative efficiency = $6/4 = 1.5$

Relative efficiency = $16/8 = 2.0$

Bristol Myers Squibb 9

General Factorials

2×3 Factorial

$2 \times 3 \times 3$ Factorial

Bristol Myers Squibb 10

Agenda Transition

- Factorial design
 - Mouse Cell Assay
- Response Surface design
 - Enzyme assay optimization
- Mixture design
 - Low Cost Media (Candida Bombicola)

Bristol Myers Squibb 11

Application of DOE to Mouse Cell Assay*

This case study highlights how a development team applies DOE to study a mouse-cell fluorescent assay performed in a 96-well plate format. They are concerned about the effects of several key factors.

* Detailed in "How Experimental Design Optimizes Assay Automation" by Thomas Erbach & Lisa Fan, Beckman Coulter, Inc., Shari Kraber, Stat-Ease, Inc., Advance, June 28, 2004, Vol. 16, No. 13, pp 18-21.

Bristol Myers Squibb 12

Mouse Cell Assay General background

This assay system is used to detect cell activity.

- Cells are pipetted into a 96-well plate where stimulant is added to induce the mouse cells to express a biomarker.
- The plate is normalized with media and incubated for two hours.
- A fluorescently tagged substrate is added to bind to the biomarker.
- The plate is read on a fluorescent plate reader at a specific wavelength.

The objective of this study is to find the settings of three factors (cell number, stimulant and substrate concentration) that maximize signal (fluorescents).

Bristol Myers Squibb 13

Mouse Cell Assay Replicated Factorial Design

Factors studied and their levels:

Factor	-1 level	+1 level
Cell Number	5000	10000
Stimulant	5 μ L	10 μ L
Substrate concentration	0.15 μ M	0.30 μ M

The 2^3 full factorial is replicated 5 times for a total of 40 runs.

Bristol Myers Squibb 14

DOE via AAO Process Overview

The figure below provides a detailed summary of the steps involved when using Design-Expert and AAO FX Software.

Bristol Myers Squibb 15

DOE & AAO Flowchart Step by Step Process

Bristol Myers Squibb 16

Step 1: Design your Experiment Export Design to AAO file...

Bristol Myers Squibb 17

Step 2: Biomek FX Method Creation Import your Design into AAO software

Bristol Myers Squibb 18

Step 3: Perform the Assay

Destination: Plate_1_1
Dispense 100 µL of Water to Plate_1_1 using the Sea technique.

Examine Plate Maps

Plate Information: Plate correct link enabled

Plate Map: A 96-well plate layout with wells colored in purple and blue.

Bristol Myers Squibb 25

Step 4 & 5: Data Acquisition & Collection

Collect the data and analyze the results

Design Information
Pipetting Information
Method Information
Data Collection

Data Collector
Data Deconvolution

Microsoft Excel - student mouse.xls

BECKMAN COULTER

Bristol Myers Squibb 26

Step 5: Data Collection

Collect the data and analyze the results

Microsoft Excel - student mouse.xls

A	B	C	D	E	F	G
A:Cell number	B:Stimulant	C:Substrate	Plate	Well	Protocol	Result
1	5000	5	5 Plate1_1_F8	Total		1655
2	5000	5	5 Plate1_1_C6	Total		3603
3	5000	5	5 Plate1_1_G2	Total		1503
4	5000	5	5 Plate1_1_F10	Total		1465
5	5000	5	5 Plate1_1_F2	Total		1771
6	5000	5	5 Plate1_1_G1	Total		2176
7	10000	5	5 Plate1_1_F6	Total		2298
8	10000	5	5 Plate1_1_E10	Total		2442
9	10000	5	5 Plate1_1_C1	Total		1516
10	10000	5	5 Plate1_1_G12	Total		1738
11	5000	10	5 Plate1_1_C7	Total		2997
12	5000	10	5 Plate1_1_E6	Total		2475
13	5000	10	5 Plate1_1_E5	Total		3234
14	5000	10	5 Plate1_1_E2	Total		2376
15	5000	10	5 Plate1_1_E3	Total		3222
16	10000	10	5 Plate1_1_E3	Total		2075
17	10000	10	5 Plate1_1_E5	Total		3333
18	10000	10	5 Plate1_1_E6	Total		3333

Bristol Myers Squibb 27

Mouse Cell Assay

Half-Normal Plot of Effects Reveal a Vital Few

Design-Expert Software
Fluxus Data

Shapiro-Wilk test
W=0.958
p-value=0.279

A:CellNumber
B:Stimulant
C:Substrate

Positive Effects
Negative Effects

Half-Normal Plot

Half-Normal % Probability

(Standardized Effect)

Bristol Myers Squibb 28

Mouse Cell Assay

Analysis of Variance (ANOVA) Shows Significance

ANOVA for selected factorial model

Source	Sum of Squares	df	Mean Square	F Value	p-value
Model	4.792E+004	3	1.597E+004	9.26	0.0001
B:Stimulant	3.074E+004	1	3.074E+004	17.82	0.0002
C:Substrate	5.896E+003	1	5.896E+003	3.40	0.0736
BC	1.132E+004	1	1.132E+004	6.57	0.0147
Residual	6.209E+003	26	2.388E+002		
Lack of Fit	2.280E+003	4	5.699E+002	0.30	0.8725
Pure Error	5.980E+003	22	2.718E+002		
Cor Total	1.100E+007	29			

Std. Dev. 475.27 R-Squared 0.4356
Mean 2479.39 Adj R-Square 0.3986
C.V. % 61.76 PRed R-Square 0.3023
PRESS 7.664E+004 Alex Precision 6.795

Bristol Myers Squibb 29

Mouse Cell Assay

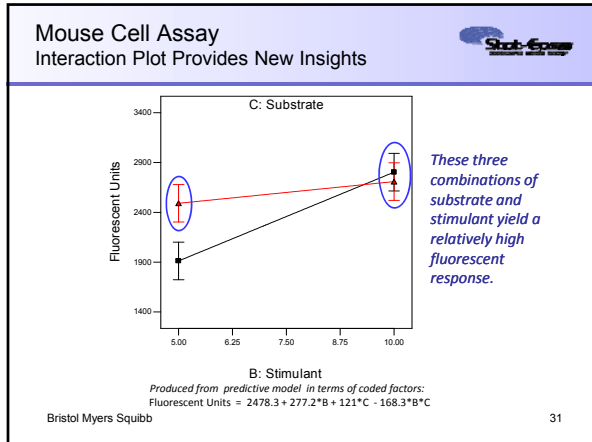
Insignificant Lack of Fit – Good!

Lack of Fit: Compares the variation of the data points about the model to the variation of the replicates about their mean.

In this case: The variation about the model surface is a bit smaller than expected given the variation in the replicates.

C: Substrate
B: Stimulant

Bristol Myers Squibb 30



Factorial Design Selection Handbook for Experimenters (page 1-8; 1 of 3)

Two Level: Selection of full and fractional factorial designs where each factor is run at 2 levels. These designs are color-coded in Stat-Ease software to help you identify their nature at a glance.

- **White:** Full factorials (no aliases). All possible combinations of factor levels are run. Provides information on all effects.
- **Green:** Resolution V designs or better (main effects (ME's) aliased with four factor interactions (4FI) or higher and two-factor interactions (2FI's) aliased with three-factor interactions (3FI) or higher). Good for estimating ME's and 2FI's. Careful: If you block, some 2FI's may be lost!
- **Yellow:** Resolution IV designs (ME's clear of 2FI's, but these are aliased with each other [2FI - 2FI]). Useful for screening designs where you want to determine main effects and the existence of interactions.
- **Red:** Resolution III designs (ME's aliased with 2FI's). Good for ruggedness testing where you hope your system will not be sensitive to the factors. This boils down to a go/no-go acceptance test. Caution: Do not use these designs to screen for significant effects.

Bristol Myers Squibb 32

Factorial Design Selection Handbook for Experimenters (page 1-8; 2 of 3)

Minimum Run Resolution V ("Min Run Res V"): Balanced (equireplicated) two-level designs containing the minimum runs to estimate all ME's and 2FI's. Check the power of these designs to make sure they can estimate the size effect you need. Caution: If any responses go missing, then the design degrades to Resolution IV.

Minimum Run Resolution IV ("Min Run Res IV"): Estimates main effects only (the 2FI's remain aliased with each other). Check the power. Caution: even one missing run or response degrades the aliasing to Resolution III. To avoid this sensitivity, accept the Stat-Ease software design default adding two extra runs (Min Run +2).

Irregular Fraction*: These special Resolution V designs may be a good alternative to the standard full or Res V two-level factorial designs.

**(Not powers of two, for example 12 runs for 4 factors.)*

Bristol Myers Squibb 33

Factorial Design Selection Handbook for Experimenters (page 1-8; 3 of 3)

General Factorial: Good for categorical factors with any number of levels: Run all possible combinations. If too many, use Optimal design.

Optimal: Choose any number of levels for each categorical factor. The number of runs chosen will depend on the model you specify (2FI by default). D-optimal factorial designs are recommended.

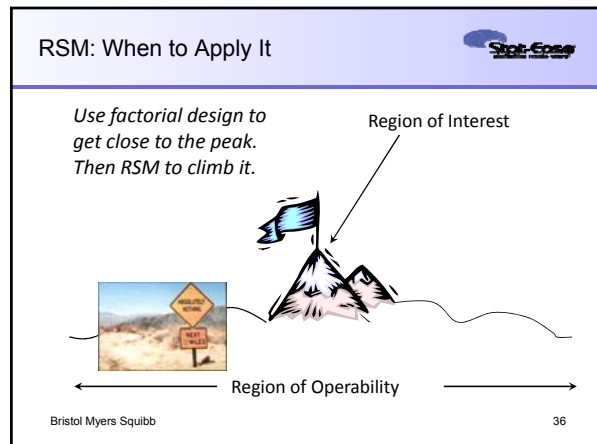
Plackett-Burman: OK for ruggedness testing but not much else due to extremely complex Resolution III alias structures. Use only with great care!

Taguchi ("Orthogonal Array"): Saturated Resolution III arrays - all columns used for ME's. Via Taguchi's "linear graphs", delete columns to estimate certain interactions. We recommend you not use these designs.

Before running fractional designs, evaluate for aliasing. If any responses go missing or the levels are botched, watch for further aliasing.

Bristol Myers Squibb 34

- ### Agenda Transition
-
- Factorial design
 - Mouse Cell Assay
 - **Response Surface design**
 - Enzyme assay optimization
 - Mixture design
 - Low Cost Media (Candida Bombicola)
- Bristol Myers Squibb 35



RSM vs OFAT

Bristol Myers Squibb 37

Agenda Transition

- Factorial design
 - Mouse Cell Assay
- Response Surface design
 - Enzyme assay optimization
- Mixture design
 - Low Cost Media (Candida Bombicola)

Bristol Myers Squibb 38

Enzymatic Assay Background

The best existing condition had the following buffer composition:
 25 mM HEPES, pH 7.5, 10 mM MgCl₂, 50 mM KCl,
 0.01% CHAPS, 0.2% bovine serum albumin, and
 200 μM Tris(2-carboxyethyl) phosphine (TCEP).

A full factorial design was created to explore four factors (each at four levels) and resulted in the assembly 4⁴ = 256 buffers.

Factor	Units	Level 1	Level 2	Level 3	Level 4
NaCl	mM	0	15	75	150
Glycerol	%	0	1	5	10
Sucrose	mM	0	25	125	250
CHAPS	%	0	0.01	0.05	0.1

Bristol Myers Squibb 39

Enzymatic Assay Build the Design (page 1 of 2)

- Select "User-Defined" on the "Response Surface" tab.
- Enter the factors name, units, change the Type to "Discrete" and their levels:

Factor	Name	Units	Type	Levels	L[1]	L[2]	L[3]	L[4]
A (Numeric)	NaCl	mM	Discrete	4	0	15	75	150
B (Numeric)	Glycerol	%	Discrete	4	0	1	5	10
C (Numeric)	Sucrose	mM	Discrete	4	0	25	125	250
D (Numeric)	CHAPS	%	Discrete	4	0	0.01	0.05	0.1

Continue >>

Bristol Myers Squibb 40

Enzymatic Assay Build the Design (page 2 of 2)

- The user defined design consists of the full factorial: 4⁴ = 256 combinations.

Continue >>

- Enter the response "S/B" (signal to background):

Response	Name	Units
S/B		

Continue >>

Bristol Myers Squibb 41

Design your Experiment Export Design to AAO file...

Run	Factor 1 A NaCl mM	Factor 2 B Glycerol %	Factor 3 C Sucrose mM	Factor 4 D CHAPS %	Response 1 S/B
163	1	0.00	5.00	125.00	0.05
162	2	15.00	0.00	125.00	0.05
39	3	75.00	1.00	125.00	0.00
88	4	150.00	1.00	25.00	0.01
236	5	150.00	5.00	125.00	0.10
17	6	0.00	0.00	25.00	0.00
41	7	0.00	5.00	125.00	0.00
179	8	75.00	0.00	250.00	0.05
170	9	15.00	5.00	125.00	0.05
250	10	15.00	5.00	250.00	0.10
237	11	0.00	10.00	125.00	0.10
121	12	0.00	5.00	250.00	0.01
242	13	15.00	0.00	250.00	0.10
129	14	0.00	0.00	0.00	0.05
128	15	150.00	10.00	250.00	0.01

Bristol Myers Squibb 42

DOE via AAO Process Overview

The figure below provides a detailed summary of the steps involved when using Design-Expert and AAO FX Software.

Bristol Myers Squibb 43

Enzymatic Assay The Analysis

Source	Sequential p-value	Lack of Fit p-value	Adjusted R-Squared	Predicted R-Squared
Linear	< 0.0001		0.5704	0.5589
2FI	< 0.0001		0.7058	0.6906
Quadratic	< 0.0001		0.7899	0.7743
Cubic	< 0.0001		0.8451	0.8181
Quartic	0.0078		0.8640	0.8134

Bristol Myers Squibb 44

Enzymatic Assay The Analysis

Bristol Myers Squibb 45

RSM Design Selection Handbook for Experimenters (page 1-10; 1 of 2)

Central Composite Designs "CCD":

- Standard** (axial levels (α) for "star points") are set for rotatability: Good design properties, little collinearity, rotatable, orthogonal blocks, insensitive to outliers and missing data. Each factor has five levels. Region of operability must be greater than region of interest to accommodate axial runs. For 5 or more factors, change factorial core of CCD to:
 - Standard Resolution V fractional design, or
 - Min-run Res V.
- Face-centered "FCD"** ($\alpha = 1.0$): Each factor conveniently has only three levels. Use when region of interest and region of operability are nearly the same. Good design properties for designs up to 5 factors: little collinearity, cuboidal rather than rotatable, insensitive to outliers and missing data. (Not recommended for six or more factors due to high collinearity in squared terms.)
- Practical alpha** ($\alpha = 4^{\text{th-root of } k}$ – the number of factors): Recommended for six or more factors to reduce collinearity in CCD.
- Small (Draper-Lin):** Minimal design – not recommended; statistical evaluation shows that properties are marginal, and it's very sensitive to outliers and missing data.

Bristol Myers Squibb 46

RSM Design Selection Handbook for Experimenters (page 1-10; 2 of 2)

Box-Behnken "BBD": Each factor has only three levels. Good design properties, little collinearity, rotatable or nearly rotatable, some have orthogonal blocks, insensitive to outliers and missing data. Does not predict well at the corners of the design space. Use when region of interest and region of operability nearly the same.

Optimal: IV, D, A (alphabetic) and distance based designs. Using two exchange algorithms to find the optimal set of points to estimate the designed for model.

- IV - default reduces the average prediction variance.
- D - minimizes the joint confidence interval for the model coefficients.
- A - minimizes the average confidence interval for the model coefficients.
- Distance based - not recommended; chooses points as far away from each other as possible, thus achieving maximum spread.

User-Defined: Generates points based on geometry of design space.

Historical: Allows for import of existing data. Be sure to evaluate this happenstance design before doing the analysis. Do not be surprised to see extraordinarily high variance inflation factors (VIF's) due to multicollinearity. The resulting models may fit past results adequately but remain useless for prediction.


Bristol Myers Squibb 47

Agenda Transition

- Factorial design
- Mouse Cell Assay
- Response Surface design
- Enzyme assay optimization
- Mixture design
- Low Cost Media (Candida Bombicola)

Bristol Myers Squibb 48

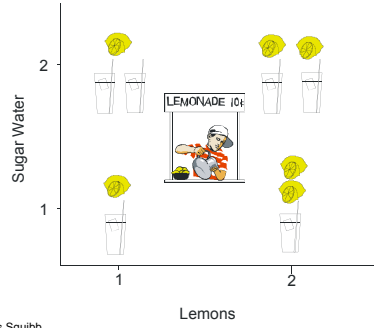
What Makes a Mixture?



1. The factors are ingredients of a mixture.
2. The response is a function of proportions.
Then the ingredients must total to a constant.

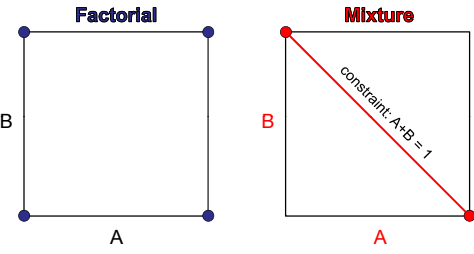
Bristol Myers Squibb 49

Factors are Ingredients Taste Dependent on Proportions



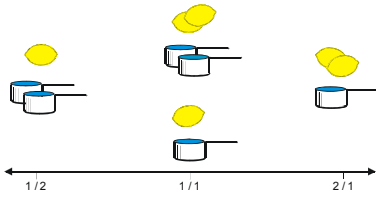
Bristol Myers Squibb 50

Two Component Mixture



Bristol Myers Squibb 51

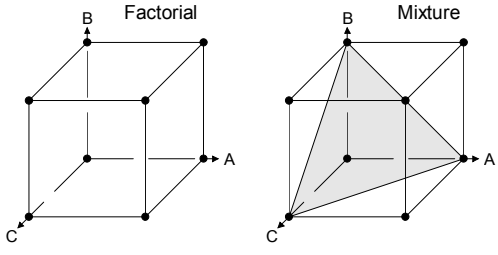
Lemonade Taste is a function of Proportions



To study mixtures you must either use ratios of ingredients or, better yet – a mixture design.

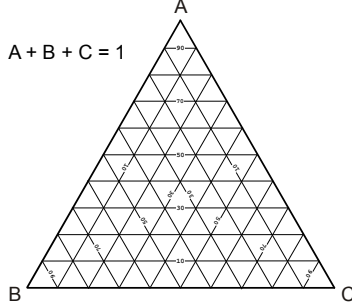
Bristol Myers Squibb 52

Three Component Mixture




Bristol Myers Squibb 53


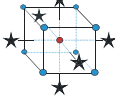
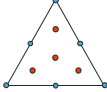
Three Component Mixture



Bristol Myers Squibb 54

Agenda Transition









- Factorial design
 - Mouse Cell Assay
- Response Surface design
 - Enzyme assay optimization
- Mixture design
 - **Low Cost Media (Candida Bombicola)**

Bristol Myers Squibb 55

Low Cost Media Candida Bombicola





Candida bombicola has been grown on a lipophilic substrate in a 96 well plate.

Add 0 to 30 μ L of the following solutions to the experimental wells:


- A. Nitrogen
- B. Phosphorus
- C. Magnesium
- D. Iron
- E. Zinc

Remove the plate from the Biomek FX and incubate for two days in an oxygen rich environment at 60°C.

Bristol Myers Squibb 56

Low Cost Media

Build a Simplex Lattice design (page 1 of 2)



- Choose a "New Design" and click on the "Mixture" tab, choose five components and enter their names:

Response Surface

Factorial

Combined

Mixture

Simplex Lattice

Simplex Centroid

Screening

Optimal

User Defined

Historical Data


Mixture components:	Name	Low	High
A (Mixture)	Nitrogen	0	30
B (Mixture)	Phosphorus	0	30
C (Mixture)	Magnesium	0	30
D (Mixture)	Iron	0	30
E (Mixture)	Zinc	0	30

then click

Bristol Myers Squibb 57

Low Cost Media

Build a Simplex Lattice design (page 2 of 2)



- Choose a quadratic model and accept the defaults:

Simplex points: 15 Order: **Quadratic**

Augment design: 6 Blocks: 1

Number of runs to replicate: 5 0 replicates needed for blocking

Total runs: 26

then click

- Enter the response name and units:


Name	Units
Biomass	kg m ⁻³

then click

Bristol Myers Squibb 58

Design your Experiment


Export Design to AAO file...



Run	Factor 1 A (mM)	Factor 2 B (mM)	Factor 3 C (mM)	Factor 4 D (mM)	Response S/B
1	0.00	0.00	0.00	0.00	4.72
2	15.00	0.00	0.00	0.00	7.29
3	75.00	0.00	0.00	0.00	7.7
4	150.00	0.00	0.00	0.00	3.42
5	0.00	1.00	0.00	0.00	8.39
6	15.00	1.00	0.00	0.00	12.84
7	75.00	1.00	0.00	0.00	3.75
8	150.00	1.00	0.00	0.00	1.5
9	0.00	5.00	0.00	0.00	8.46
10	15.00	5.00	0.00	0.00	5.44
11	75.00	5.00	0.00	0.00	4.49
12	150.00	5.00	0.00	0.00	2.88
13	0.00	10.00	0.00	0.00	9.13
14	15.00	10.00	0.00	0.00	5.46
15	75.00	10.00	0.00	0.00	3.08

Bristol Myers Squibb 59

DOE via AAO Process Overview



The figure below provides a detailed summary of the steps involved when using Design-Expert and AAO FX Software.

DESIGN
Design-Expert

AAO FX

ANALYSIS
Design-Expert

Import Design


Generate Plate Map

Configure & Create Method

Run Assay on Biomek FX

Collect Data

Export Data to DXE



Bristol Myers Squibb 60

Low Cost Media Biomass Analysis

Source	Sequential	Lack of Fit	Adjusted	Predicted
	p-value	p-value	R-Squared	R-Squared
Linear	0.9921	0.0002	-0.1688	-0.6313
Quadratic	< 0.0001	0.3426	0.9558	0.8760
Special Cubic	0.8332	0.0480	0.9530	16.4562

Bristol Myers Squibb 61

Low Cost Media Maximum Biomass (Ramps)

Desirability = 0.950

Bristol Myers Squibb 62

Low Cost Media Maximum Biomass (Trace Plot)

Bristol Myers Squibb 63

Mixture Design Selection Handbook for Experimenters (page 1-14; 1 of 2)

Simplex designs: Applicable if all components range from 0 to 100 percent (no constraints) or they have same range (necessary, but not sufficient, to form a simplex geometry for the experimental region).

- Lattice:** Specify degree "m" of polynomial (1 - linear, 2 - quadratic or 3 - cubic). Design is then constructed of m+1 equally spaced values from 0 to 1 (coded levels of individual mixture component). The resulting number of blends depends on both the number of components ("q") and the degree of the polynomial m. Centroid not necessarily part of design.
- Centroid:** Centroid always included in the design comprised of 2^q - 1 distinct mixtures generated from permutations of:
 - Pure components: (1, 0, ..., 0)
 - Binary (two-part) blends: (1/2, 1/2, 0, ..., 0)
 - Tertiary (three-part) blends: (1/3, 1/3, 1/3, 0, ..., 0)
 - and so on to the overall centroid: (1/q, 1/q, ..., 1/q)

Simplex Lattice versus Simplex Centroid

Bristol Myers Squibb 64

Mixture Design Selection Handbook for Experimenters (page 1-14; 2 of 2)

Extreme vertices designs (in order of preference):

- Optimal:** (See RSM design selection for details.) Use when component ranges are not the same, or you have a complex region, possibly with constraints.
- Distance-Based:** (See RSM design selection for details.)
- User-Defined:** (See RSM design selection for details.)

Screening designs: Essential for six or more components. Creates design for linear equation only to find the components with strong linear effects.

- Simplex screening**
- Extreme vertices screening** (for non-simplex)

Bristol Myers Squibb 65

Combined Design Selection Handbook for Experimenters (page 1-15)

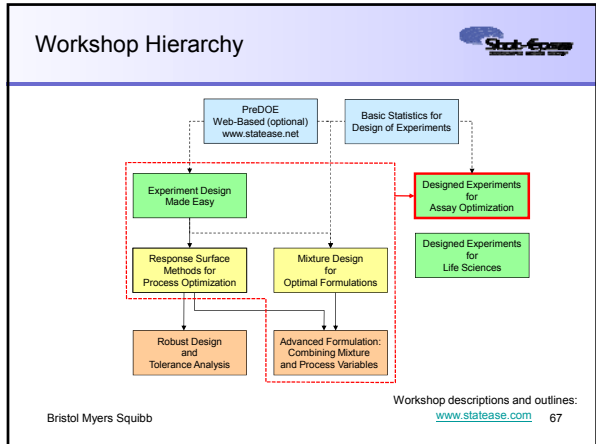
These designs combine either two sets of mixture components, or mixture components with numerical and/or categorical process factors. At least two types of variables, one of them being "Mixture 1", must be specified to achieve the status of a combined design. For example, if you want to mix your filled cupcake and bake it too, using two ovens, identify the number of:

- Mixture 1 components – the cake: 4 for flour, water, sugar and eggs
- Mixture 2 components – the filling: 3 for cream cheese, salt and chocolate
- Numeric factors – the baking process: 2 for time and temperature
- Categorical factors – the oven: 2 types – Easy-Bake or gas.

You must specify models for the mixtures and process factors (quadratic by default). The **User Defined** design option pares down the runs to the bare minimum needed to fit the combined models. Design-Expert software will add by default:

- Lack of fit points (check blends) via distance-based criteria
- Replicates on the basis of leverage.

Bristol Myers Squibb 66



Design-Expert® Software Statistics Made Easy®

Best of luck for your experimenting!

Thanks for listening!

Pat & Shari

Bristol Myers Squibb 68