An Overview of Multiphase Optimization Strategy (MOST) for Behavioral Intervention Development

Liang Li, PhD

Biostatistics @ MD Anderson

Genetic Epidemiology Work Group Meeting, December 5, 2013

The Multiphase Optimization Strategy (MOST)

- An increasingly popular study design strategy in behavioral sciences research
- Proposed by Linda Collins @ Penn State
- NIH is currently funding proposals specifically for the development and application of MOST methodology
- June 18, 2013 Workshop by Dr. Collins at MDACC
- Not a statistical method, Not a study design, but a strategy to develop behavioral interventions (it involves some statistics, mainly the factorial design methodology)

Multicomponent Behavioral Interventions for Prevention and Treatment

- **Examples:**
  - ✓ Treatment for depression
  - ✓ School-based drug abuse prevention
  - ✓ Prevention/treatment of obesity
  - ✓ Smoking cessation treatment
    1. Precessation nicotine patch (Y/N)
    2. Precessation nitotine gum (Y/N)
    3. Precessation in-person consulting (Y/N)
    4. Cessation in-person consulting (minimal/intensive)
    5. Cessation phone consulting (minimal/intensive)
    6. Maintenance medication duration (short/long)

Research Question

- **Goal: Develop a behavioral intervention to maximize the probability of successful quitting**
  - ✓ **Choose from a set of candidate components**
  - ✓ **Set the component levels**
  - ✓ **Consider constraints (cost, compliance, etc.)**

Possible Approaches

- Individual experiments:
  - ✓ Run K independent randomized trials for each of the K candidate components
- Single factor experiments: K+1 group randomized trial

	A	B	C
0	N	N	N
1	Y	N	N
2	N	Y	N
3	N	N	Y

- Problem:
  - ✓ It is unclear how the various components work together

Possible Approaches

- Treatment package approach:
  - ✓ Set each component to the highest level
  - ✓ Run a two group randomized trial with controls
- Problems:
  - ✓ If it is better than the control, does not indicate which component makes a difference
  - ✓ If not better, does not indicate which component did affect an improvement
  - ✓ Does not take cost or other constraints into account

The MOST Approach

- Motivated by engineering applications
  - ✓ Start with a clear goal, including constraints
  - ✓ Using the resources available, design an efficient experiment to gather the needed information (e.g., individual effects of components) (factorial design)
  - ✓ Based on the information, choose components and levels to achieve the goal (optimization)
  - ✓ Then compare new intervention to the old one (RCT)
- Optimization (in engineering): the process of finding the best possible solution to a problem … subject to given constraints

The Engineering Perspective

- Manufacture the truck leaf string:
  - ✓ furnace temperature (low, high)
  - ✓ heating time (short, long)
  - ✓ transfer time on conveyer belt (short, long)
  - ✓ hold down time in high pressure press (short, long)
  - ✓ quench oil temperature (low, high)

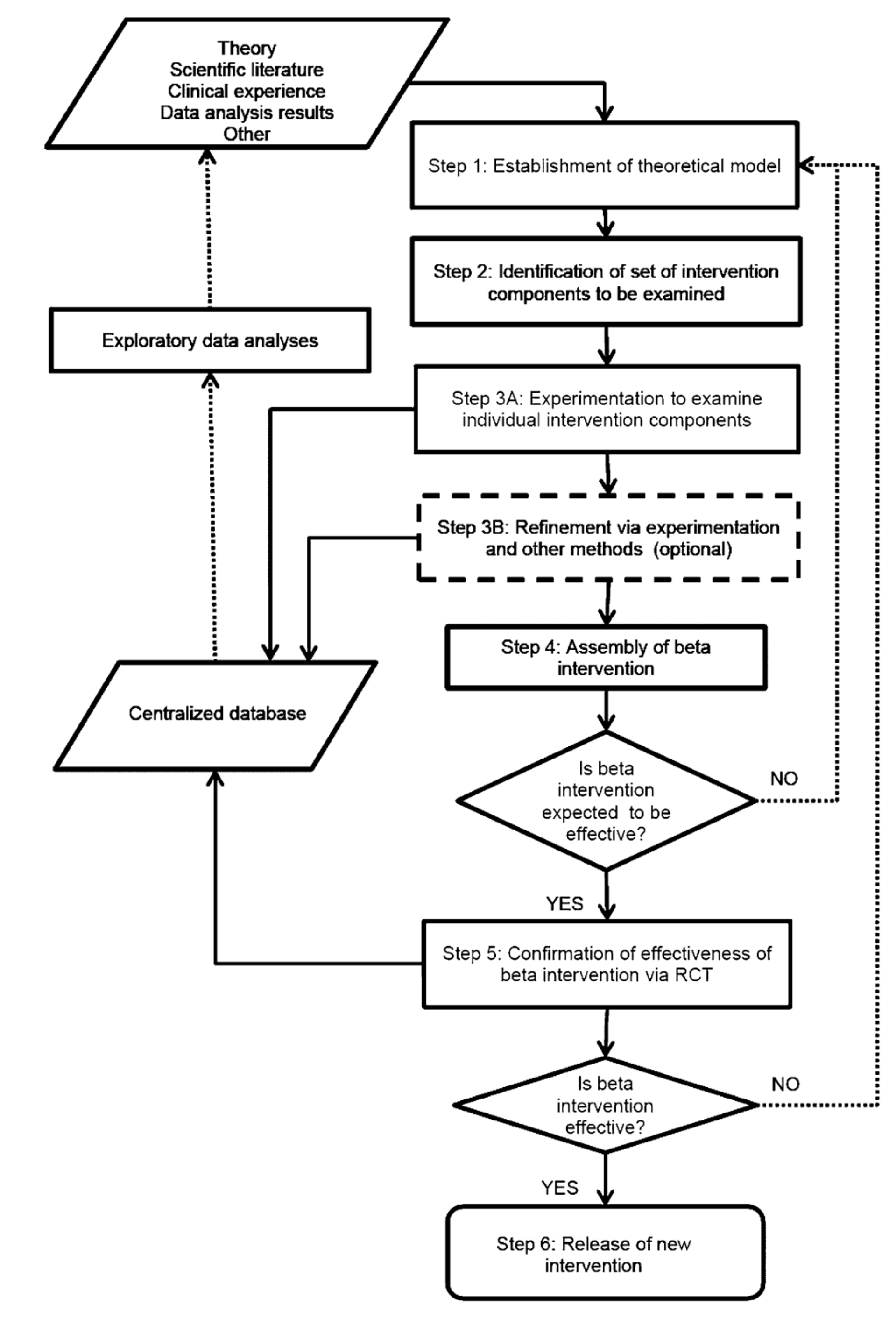
The MOST Approach

- **Features:**
  - ✓ Indicates which components are active and which are redundant
  - ✓ Ensures an incremental improvement, and therefore is the fastest way to the best intervention in the long run
  - ✓ Readily incorporates costs/constraints of any kind
- **Note:**
  - ✓ There is no “MOST design”; it is a research strategy
  - ✓ At its core is a factorial design
  - ✓ “Optimal” does not mean “Best” in engineering
  - ✓ Not for causal effect of individual components (not a RCT)
  - ✓ Not confirmatory, but exploratory

The Two Principles of MOST

- Resource management principle (engineering way of thinking):
  - ✓ Huge (e.g., 64-arm RCT for 6 components) would be definitive, but not feasible
  - ✓ Given the resources and constraints, what is the most efficient way to achieve the goal
- Continuous optimization principle (engineering):
  - ✓ I have finished developing this product and it is ready to market
  - ✓ Now I am going to start developing the new, improved product
  - ✓ Optimization is a “cyclic process”

The Flow Chart of MOST



Collins, et al 2011
Ann Behav Med

The Flow Chart of MOST

- Step 1 and 2: Theoretical model and Identification of set of candidate intervention components
- Step 3A (main experiment): Use complete factorial design (CFD) or fractional factorial design (FFD)
- Step 3B (optional -- refinement): in the smoking cessation example, one component is maintenance medication weeks: 8 vs. 16. If this component is found to be important in Step 3A, another experiment may be needed to find the optimal number of weeks
- Step 4 (assembly of beta intervention): is it better than the currently standard intervention?
- Step 5: confirmatory two-arm RCT comparing the beta intervention with the currently standard intervention

Overview of Step 3A: Experimentation to Examine Individual Intervention Components

- Objective is to identify the most promising components and level/settings
- NOT to compare each combination to a control or against each other (not a confirmatory RCT)
- Optimization criteria: effectiveness, cost/time and other constraints
- NOT to identify single best combination
  - ✓ Multiple combinations of intervention components may lead to similar results
  - ✓ To identify the single best combination, an enormous RCT is the only way and it is often impractical
  - ✓ Find a good combination, if not the best
  - ✓ MOST may identify the best combination <u>in the long run</u>

Factorial Design

- **Dr. Collins: there is no "MOST" design; it is a research strategy that uses the factorial design (proposed by R. A. Fisher)**
- **Factorial design: subjects randomized to 2^K conditions in order to study the main effect and interactions of K intervention components**

Condition	A	B	C
1	1	1	1
2	1	1	-1
3	1	-1	1
4	1	-1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1

Why Using Factorial Design?

- **It enables examination of individual component effects**
- **It requires SMALLER sample sizes than alternative designs (individual experiment, treatment package, single factor experiments)**
- **But it usually requires more experiment conditions than we may be accustomed to (causing logistic difficulties)**
- **We do not compare the conditions (they are too many of them!); we estimate the main effect and some interactions of scientific interest**
- **The estimates are based on all the subjects (efficiently use subjects and reduce sample size)**

Why are we interested primarily in main effects?

- Using the information for decision making, not necessarily estimating pristine effects
- Effect is likely to be robust if it obtains an average across many other factors --- R. A. Fisher
- If theory and prior research specifies and explains an interaction, it must always be dealt with
  - ✓ But we know little about interactions
  - ✓ Most theories and models are silent on this topic
- Where theory/prior research do not specify whether or not there is an interaction, we rely on these principles from engineering:
  - ✓ Effect sparsity: there are a lot of effects in a factorial experiment, most are not significant or important
  - ✓ Hierarchical ordering: those that are important are likely to be simpler effects, i.e., main effects first, then two-way interactions

Powering the factorial design

- Power for main effects: sample size requirements for a k-factor experiment about the same as for a t-test
- Power the experiment for the smaller effect size
- Adding a factor generally does not increase sample size requirements, unless that factor is expected to have a smaller effect size
- Power the study for the smallest effect size that you would accept for inclusion in the intervention
- Usually not powered for interactions
  - ✓ Little is known about interactions
  - ✓ Effect sizes are probably much smaller than main effects

Powering the factorial design

- **Three intervention components: A, B, C**

Design	Design	n	# conditions	interactions
Individual experiment	A vs. NULL B vs. NULL C vs. NULL	168	6	None can be estimated
Single factor experiment	A vs. B vs. C vs. NULL	112	4	None can be estimated
Complete factorial	Factorial (A, B, C)	56	8	All can be estimated

Some Misconceptions

- Misconception 1: factorial experimental designs require larger numbers of subjects than available alternative designs
  - ✓ Reality: when used to address suitable research questions, balanced factorial designs often require may FEWER subjects than alternative designs
- Misconception 2: if you want to add a factor to a balanced factorial design, you will need to increase the sample size dramatically
  - ✓ Reality: If the effect size of the added factor is no smaller than the factors already in the experiment, power will be about the same
- Misconception 3: the primary motivation for conducting a factorial design is always to test for interactions
  - ✓ Reality: even if there is no interaction, you can still conduct a factorial design to make economical use of subjects

Some Reasons for Not Using a Factorial Design

- Intervention composed of many components with tiny effects, overall effect is cumulative
  - ✓ May be difficult to power the study for tiny effects
  - ✓ May need to sort the components into bundles and study bundles
- Factorial design requires more experimental conditions:
  - ✓ May reduce it by using fractional factorial design (FFD)

Fractional Factorial Design (FFD)

- Well established statistical theory & software, applied to behavioral science
- Factorial designs in which only a subset of experimental conditions are run
- Carefully choose the subset to answer the scientific questions; omitted conditions are not needed
- FFD requires at most ½ of the cells of a complete factorial design (CFD), often many fewer
- Example: K factors, CFD has 2^K conditions, FFD may have 2^{K-1} or 2^{K-2} conditions

About FFD

- Why run just a subset of conditions?
  - ✓ Economy: K factors, CFD has 2^K conditions; $2^6 = 64$, $2^7 = 128$
  - ✓ Example: FFD may conduct a 2^7 experiment with only 16 conditions
- When you might consider a FFD?
  - ✓ 5 or more factors (FFD exists for 3 or 4 factors, but benefit is small and strong assumptions are needed)
  - ✓ You are primarily interested in main effects and low-order interactions
  - ✓ Remaining effects and high order interactions are negligible

cell	A	B	C	A	B	C	A*B	A*C	B*C	A*B*C
1	Off	Off	Off	-1	-1	-1	1	1	1	-1
2	Off	Off	On	-1	-1	1	1	-1	-1	1
3	Off	On	Off	-1	1	-1	-1	1	-1	1
4	Off	On	On	-1	1	1	-1	-1	1	-1
5	On	Off	Off	1	-1	-1	-1	-1	1	1
6	On	Off	On	1	-1	1	-1	1	-1	-1
7	On	On	Off	1	1	-1	1	-1	-1	-1
8	On	On	On	1	1	1	1	1	1	1

cell	A	B	C	A	B	C	A*B	A*C	B*C	A*B*C
2	Off	Off	On	-1	-1	1	1	-1	-1	1
3	Off	On	Off	-1	1	-1	-1	1	-1	1
5	On	Off	Off	1	-1	-1	-1	-1	1	1
8	On	On	On	1	1	1	1	1	1	1

Statistical Power of FFD

- FFD and CFD have the same statistical power (using FFD does NOT reduce or increase sample size)
- Compared to the corresponding CFD, in a FFD:
  - ✓ Each condition will have more subjects than the CFD
  - ✓ But each effect estimate based on SAME number of subjects

Design	Number of subjects needed for power > 0.9	Number of conditions	Interactions
CFD	512	$2^6 = 64$	All can be estimated
FFD	512	8-32 depending on design	Selected subset can be estimated

Notation for FFD

- Suppose 4 factors, each factor has 2 levels.
- CFD: 2^4 (16 conditions/cells)
- An FFD with 8 conditions is represented as 2^{4-1}
  - ✓ $2^{4-1}=2^3=8$
  - ✓ This notation tells you:
    - ➢ The number of conditions in the original CFD
    - ➢ The number of conditions in the FFD
    - ➢ The fraction by which FFD reduces the original: ½
    - ➢ The number of aliases of each estimable effect in FFD: 2

What is Aliasing?

cell	A	B	C	A	B	C	A*B	A*C	B*C	A*B*C
1	Off	Off	Off	-1	-1	-1	1	1	1	-1
2	Off	Off	On	-1	-1	1	1	-1	-1	1
3	Off	On	Off	-1	1	-1	-1	1	-1	1
4	Off	On	On	-1	1	1	-1	-1	1	-1
5	On	Off	Off	1	-1	-1	-1	-1	1	1
6	On	Off	On	1	-1	1	-1	1	-1	-1
7	On	On	Off	1	1	-1	1	-1	-1	-1
8	On	On	On	1	1	1	1	1	1	1

- **Estimate the effect of A: compare 2, 3, vs. 5, 8; A is aliased with B*C**
- **Estimate the effect of B: compare 2, 5, vs. 3, 8; B is aliased with A*C**
- **Estimate the effect of C: compare 2, 8, vs. 3, 5; C is aliased with A*B**

(A + B*C) equals A if B*C is negligible

Aliasing of Effects

- Consider a 2^4 factorial design
- 4 factors, 16 conditions/cells
- Effects estimated (TOTAL = 16 effects)
  - ✓ 1 intercept
  - ✓ 4 main effects
  - ✓ 6 two-way interactions
  - ✓ 4 three-way interactions
  - ✓ 1 four-way interactions
- For both CFD and FFD, there are as many estimable effects as the number of conditions/cells
  - ✓ FFD: reduce the number of conditions so that effects of scientific interest are estimable but negligible effects are not estimated (resource management principle)

Aliasing of Effects

- Now consider a 2^{4-1} fractional factorial design
- 4 factors, 8 conditions/cells, 8 estimable effects
- The original 16 effects are combined into 8 estimable effects (aliased)
- In any FFD, it is known which effects are aliased with which
  - ✓ In a ½ FFD, each effect is aliased with 1 other effect (“bundles” of 2)
  - ✓ In a ¼ FFD, each effect is aliased with 3 other effect (“bundles” of 4)
  - ✓ And so on … …
  - ✓ We choose to bundle the effect of scientific interest (main effects & important interactions) with a few other negligible effects
- For 4 factors, there is only 1 FFD
- As the number of factors increases, there are many FFD for each CFD
  - ✓ Choose the one for the specific scientific study

Resolution of an FFD

- FFDs are classified according to their resolution
- For a given CFD, there may be many FFDs with different resolutions
- Resolution is denoted by Roman numbers: II, IV, V, VI
- In general, in a resolution R FFD, F-way interactions are aliased ONLY with R-F way interactions or higher order interactions
  - ✓ Resolution V: 2-ways aliased only with (5-2)=3-ways or higher

Resolution	Main effects are NOT aliased with	2-way interactions are NOT aliased with
III	Main effects	
IV	Main effects, 2-way	Main effects
V	Main effects, 2-way, 3-way	Main effects, 2-way
VI	Main effects, 2-way, 3-way, 4-way	Main effects, 2-way, 3-way

Resolution of an FFD: Example

- CFD would be 2^6=64 conditions
- We choose FFD with 2^{6-1}=32 conditions
- Fraction = half; each effect aliased with another effect
- Resolution VI:
  - ✓ Each main effect aliased with a 5-way interaction
  - ✓ Each 2-way interaction aliased with a 4-way interaction

How to Choose a FFD?

- Classify all effects of CFD into 3 categories
  a) Effects of primary scientific interest: make them estimable
  b) Effects expected to be 0 or negligible
  c) Effects not of scientific interest but may be non-negligible
- Alias (a) and (c) with (b); Do not alias (a) with (c)
- More effects are designated negligible → FFD with fewer conditions
- No effect is negligible → CFD is the only choice; FFD does not exist
- Heuristic guiding principles (engineering):
  - ✓ Hierarchical ordering: priority be given to lower order effects
  - ✓ Effect sparsity (Pareto principle): number of non-negligible effects is a small fraction of the total number of effects (2^K)
- Higher resolution FFD is better than lower resolution ones, because they alias main effects and 2-way interactions with high order interactions

Summary

- MOST (research strategy, not a design) is different from SMART (a design)
- Developed by Penn State Methodology Center; Download references there

  http://methodology.psu.edu/
- It is not a new statistical methodology, but an application of existing statistical methods (factorial design, less well known among behavioral scientists) to behavioral intervention development
- Software: SAS PROC FACTEX
  - ✓ Specify a desired resolution
  - ✓ Specify which effects are in category (a), (b), and (c)
  - ✓ Specify constraints, including costs and maximum number of conditions
  - ✓ PROC FACTEX will generate the design (or infeasible)