Center on HEALTH CARE EFFECTIVENESS

ISSUE BRIEF

APRIL 2012

Orthogonal Design: A Powerful Method for Comparative Effectiveness Research with Multiple Interventions

by Jelena Zurovac and Randy Brown

There is a growing interest in new or modified research methods that can balance the needs for quick results and statistical rigor in studies of intervention effectiveness. Orthogonal design, which has been extensively used in manufacturing and other fields but not in published health care evaluations or comparative effectiveness research (CER), affords an opportunity to design interventions in real-world settings and to study intervention components that can be implemented in various ways. This brief introduces orthogonal design, describes key design and implementation considerations, and illustrates how it can be applied in CER studies by describing its use in an ongoing study of care coordination delivery by Special Needs Plans (SNPs) to dually eligible individuals.

The Need for New Methods

The aim of comparative effectiveness research (CER) is to assess the relative ability of two or more interventions to influence a desired outcome, and to evaluate how their effectiveness differs across patients and circumstances. The ongoing focus on CER, and the desire to obtain findings more quickly than typically possible with randomized controlled trials, have created an interest in new or modified methods. Of particular interest are new methods amenable to multi-component interventions that can be implemented in various ways. Learning how best to implement intervention components in a real-world setting is important because existing evidence often provides no guidance on whether more intensive or more expensive ways to implement interventions yield significantly better outcomes.

Experience shows that, even after traditional evaluation methods (such as randomized trials) find a given intervention effective, the programs often do not produce comparably favorable results when the intervention is repeated in other settings (due in part to differences in *how* the intervention was implemented). In the last decade, several models of transitional care have been shown to be effective in reducing readmission rates (Naylor et al. 1999; Coleman et al. 2006; Jack et al. 2009). However, less is known about how best to implement the various components. For example, evidence shows that post-discharge follow-up helps reduce readmission rates, but there is little information about how quickly this follow-up visit needs to occur, how many times, and which protocols to use. Mahoney (2010) notes that the success of multi-component interventions depends critically on the features of the intervention, the methods used to engage patients and providers, and the target population for the intervention components.

WHY IS ORTHOGONAL DESIGN WELL-SUITED FOR CER?

- It can identify unplanned variations occurring in real world practice and test them to determine which is best.
- It combines the rigor of experimental design with the ability to produce rapid results by testing multiple components at once.
- It can be used to assess whether more resource-intensive ways of implementing interventions yield sufficiently better patient outcomes to warrant the investment.

MATHEMATICA Policy Research

The Center on Health Care Effectiveness (CHCE) was founded in 2010 to inform health care policy and improve patient outcomes. Building on more than four decades of Mathematica's substantive expertise, methodological skill, and corporate infrastructure, CHCE is at the forefront of comparative effectiveness policy analysis and research.

ORTHOGONAL DESIGN OVER THE YEARS

Sir Ronald Fisher developed orthogonal design, described in his seminal book *Design of Experiments* (1935), based on agricultural experiments in England. Over the years, this design has been used widely in the chemical and automotive industries, natural science research, business and marketing, and operations research. Recently, health plans have used orthogonal design for internal quality improvement initiatives, although these studies are rarely in the public domain (Moore 1994; Jones and Moore 1995).

Orthogonal Design

Orthogonal design is an experimental design used to test the comparative effectiveness of multiple intervention components—referred to here as "interventions"—each of which takes on two or more variants. In a two-level design, the researcher specifies two variants (referred to here as options *a* and *b*) for each intervention—such as which of two protocols are followed or which of two levels of intensity are used. An algorithm is then used to generate a specific set of combinations of a's and b's that constitute an orthogonal design for the number of interventions to be tested. The researcher then randomly assigns (without replacement) one of these predetermined combinations of variants to each experimental unit (for example, nurses or classrooms), who will then administer this combination of variants to all of their subjects (such as patients or students). The matrix on this page shows a design in which 12 experimental units are each assigned to implement a specific variant for each of the 11 interventions to be tested. The set of combinations ensure "orthogonality," which refers to a property of the assignments that ensures that the effect of any one intervention is unconfounded with the effects of any other single intervention. These combinations ensure that half of the experimental units are assigned to option a, and half to option b, for each intervention. Half of the units assigned to option a for intervention 1 are assigned to option a for intervention 2 and half to option b, and so on for all possible intervention pairings. However, in efficient orthogonal designs, any single intervention is confounded with two-way or higher order interactions between other interventions. The degree of such confounding depends on the number of experimental units used relative to the number of interventions being tested.

The effect of an individual intervention is calculated by comparing the mean outcome over all subjects for experimental units that provide one variant (*a*) to the mean for subjects of those who provide the other (*b*). As in any other design, one can use regression analysis with subject-level data to compute intervention effects, controlling for any differences between units with respect to the characteristics of their subjects. The regression can also control for potentially important characteristics of the experimental units, such as the baseline value of the outcome variable for the experimental units, if the number of units is substantially larger than the number of interventions being tested. Box (2005) provides a comprehensive discussion of methods.

The critical advantage of orthogonal design relative to typical controlled trials is that it allows the researcher to test the effectiveness of many interventions simultaneously in a single experiment (and possibly identify some of their interactions) with far fewer experimental units than it would take to exhaust all possible intervention combinations. This feature makes it particularly valuable for testing the best way to implement complex interventions with many facets. For example, it can be applied in testing ways to implement the numerous components or activities involved in care management of complex patients. Variations across units in how interventions are implemented will occur regardless of whether the variants are tested; explicit testing as part of the experimental design allows the program operator to learn which variant is best for each intervention.

This design yields useful results even if the difference between the two variants of any intervention is not statistically significant. Often, one of the two tested variants is more difficult, burdensome, or expensive to implement; thus, a finding of no significant difference means that the more burdensome variant does not produce better results than the less burdensome variant.

Design and Implementation

The most important step in designing a successful orthogonal design study that yields credible, actionable results, is to identify intervention components and alternative options that are feasible to implement and can improve patient outcomes and/or operational costs. The interventions should also be ones that the program implementer wishes to test.

EXAMPLE OF AN ORTHOGONAL DESIGN MATRIX

Experimental Unit	Intervention										
	1	2	3	4	5	6	7	8	9	10	11
1	a	a	b	a	a	a	b	b	b	a	b
2	b	a	a	b	a	a	a	b	b	b	a
3	a	b	a	a	b	a	a	a	b	b	b
4	b	a	b	a	a	b	a	a	a	b	b
5	b	b	a	b	a	a	b	a	a	a	b
6	b	b	b	a	b	a	a	b	a	a	a
7	a	b	b	b	a	b	a	a	b	a	a
8	a	a	b	b	b	a	b	a	a	b	a
9	a	a	a	b	b	b	a	b	a	a	b
10	b	a	a	a	b	b	b	a	b	a	a
11	a	b	a	a	a	b	b	b	a	b	a
12	b	b	b	b	b	b	b	b	b	b	b

Further, when designing interventions or choosing components, it is important to consider:

- The likelihood that the experimental units will faithfully implement the component variants to which they are assigned without excessive oversight beyond what would be expected to occur in an ongoing program.
- The likelihood that variants shown to be effective will actually be adopted as tested.
- The number of interventions to test.

Orthogonal design studies face implementation considerations similar to those of other studies, but resistance from participants might be stronger. People who implement the interventions often feel initially that testing many interventions is "too complicated." Engaging these experimental agents in selecting the intervention variants to be tested and using tools such as implementation guides and individualized assignment sheets can ease concerns about fidelity and help the researcher to better understand whether (and how) the interventions might be implemented on an ongoing basis. Requiring the agents to record the provision of interventions can yield important data for measuring fidelity to interventions and interpreting findings. In addition, researchers should assess facilitators and barriers to implementation to understand whether the interventions were actually implemented and the factors that are necessary for effective implementation. Researchers should also keep in mind that the health care environment differs from other disciplines in which orthogonal design has historically been used (such as manufacturing), which generally have more controlled environments. The analyses should always take the "intent-to-treat" approach, in which effects of interventions are computed by comparing outcomes of those assigned to two variants of a given intervention, regardless of whether or how thoroughly the variants were actually delivered.

When deciding how many interventions to test, researchers must first ensure that a sufficient number of experimental units are available to obtain unconfounded main effects. (For example, the matrix shown on p. 2 requires at least 12 experimental units to estimate main effects of 11 interventions.) If more than the minimum number of experimental units is available to test the desired interventions, the researcher must decide how to use the additional experimental units. The two options are to (1) test more interventions or (2) improve "resolution," that is, to reduce the extent of confounding of main effects with potentially important second-order interactions of other interventions. In other words, the study can test more interventions with greater confounding of main effects with second-order interactions, or test fewer interventions with less confounding. If the number of experimental units is not fixed, the researcher must perform power calculations to determine how many experimental units are needed for the desired precision before finalizing the number of interventions. Because this is a clustered design in which groups of research subjects all receive the same set of interventions, power depends almost entirely on the number of experimental units.

Limitations and Caveats

Two key assumptions embedded in orthogonal design are that the experimental units (agents) are homogenous, and that all third or higher order intervention interactions can be considered to be negligible. Homogeneity simply means that the agents would have roughly similar outcomes if implementing the same interventions, which can be assessed by comparing average outcomes prior to the study for each agent. Outliers that have a high likelihood of distorting the comparison of means can then be excluded from the study. The degree to which an orthogonal design study is subject to confounding depends primarily on the number of interventions and experimental units and somewhat on the selected design. Although the researcher can choose how much confounding to tolerate and which interaction effects to estimate, the limitation of designs that use few experimental units relative to the number of interventions tested (such as the matrix on p. 2) is that the effect of any one intervention is confounded with many two-way interactions. Doubling the number of experimental units in this design to 24 would ensure that each of the 11 intervention effects being tested in this "foldover" design is confounded only with three-way and higher order interactions, which are usually expected to be negligible. The larger number of experimental units also increases the statistical power of the tests for all interventions.

Opportunities and Recommendations

Orthogonal design is potentially a very useful tool for rapid-cycle comparative effectiveness research. It provides an opportunity to test—rigorously and simultaneously—many aspects of delivering multifaceted interventions. This advantage is relevant because nearly all policy research today focuses on whether a broad concept is effective in improving outcomes, without studying the operational details of *how* the interventions are provided. The experimental agents' experience and understanding of current processes of care is invaluable in identifying the intervention components for which there is the most uncertainty about the relative effectiveness of alternative ways to implement the component. Furthermore, variations across health care workers in how interventions are implemented will occur in any case; by formalizing and testing the variations, program operators can learn which ones yield the greatest improvements in patient outcomes. Therefore, to harness the most value from this powerful design, studies should focus on developing interventions in collaboration with those who would be implementing them, such as the care coordinators in our study. The flexibility of orthogonal designs allows researchers to choose whether to screen many interventions at once, or to test a smaller number with greater statistical precision.

APPLICATION: STUDY OF CARE COORDINATION DELIVERY

Special Needs Plans (SNPs) were established in 2003 as part of the Medicare Prescription Drug, Improvement, and Modernization Act, with the goal of improving care for three high-risk target populations of Medicare beneficiaries: (1) dual eligibles (those enrolled in both Medicare and Medicaid), (2) beneficiaries with chronic conditions, and (3) beneficiaries residing in nursing homes. SNPs contract with CMS to provide all covered Medicare services in return for a monthly risk-adjusted capitation payment. Thus, SNPs have the incentive—and the requirement, under their contract with CMS—to engage in care management to help patients reduce their need for expensive services, especially hospitalizations and emergency room visits. Although many protocols, interventions, and screening tools exist, SNPs still have many questions about the most efficacious and cost-effective ways to design and implement interventions to address the unique and varied needs of the high-risk populations they serve.

In collaboration with three SNPs serving dual eligibles, we designed and are conducting an orthogonal design study to help SNPs identify care coordination strategies that work best for their members. The main outcomes we will analyze are hospital admissions, readmissions, and emergency room visits. One part of the study involves 25 care coordinators who are implementing either current practice or an enhanced practice for each of 11 interventions over a 12-month period. The studied intervention variants address how often an intervention is provided, or which procedures or protocols are used in conducting routine contacts with patients, screening for the risk of falls, depression screening, care planning, patient coaching, and management of care transitions. For example, the study is testing a care transitions intervention that compares the current practice of conducting one follow-up visit with patients within three business days of discharge from an inpatient setting versus an enhanced practice that includes a second follow-up visit within a week of the first one. Another care transitions intervention being tested compares the current practice to a more structured follow-up process that uses a checklist and an instrument to assess the patient's understanding of post-discharge instructions. One of the depression screening interventions tests effectiveness of a longer screening instrument versus a shorter one and the other tests variants in the minimum frequency of routine screening contacts. The details of the interventions were developed in collaboration with the experimental agents (the care coordinators) and management of the participating plans.

References

Box, George E. P., Stuart J. Hunter, and William G. Hunter. "Statistics for Experimenters: Design, Innovation, and Discovery. 2nd. Sl." 2005.

Coleman, Eric. A., Carla Parry, Sandra Chalmers, and Sung-Joon Min. "The Care Transitions Intervention: Results of a Randomized Controlled Trial." *Archives of Internal Medicine*, vol. 166, no. 17, 2006, pp. 1822–1828.

Fisher, Ronald A. "The Design of Experiments." Edinburgh: Oliver and Boyd, 1935.

Jack, Brian W, Veerappa K. Chetty, David Anthony, Jeffrey L. Greenwald, Gail M. Sanchez, Anna E. Johnson, et al. "A Reengineered Hospital Discharge Program to Decrease Rehospitalization." *Annals of Internal Medicine*, vol. 150, no. 3, 2009, pp. 178–87.

Jones, Frederic G., and Colletta H. Moore. "Designing and Executing Experiments in Care—A Data-Driven, Scientific Approach to Quality Improvement." *In Measuring Clinical Care: A Guide for Physician Executives*, edited by S. C. Schoenbaum. Providence, RI: American College of Physician Executives, 1995.

Mahoney, Jane E. "Why Multifactorial Fall-Prevention Interventions May Not Work: Comment on "Multifactorial Intervention to Reduce Falls in Older People at High Risk of Recurrent Falls." *Archives of Internal Medicine*, vol. 170, no. 13, 2010, p. 1117.

Moore, Colletta H. "Experimental Design in Health Care." Quality Management in Healthcare, vol. 2, no. 2, 1994, p. 13.

Naylor, Mary D., Dorothy Brooten, Roberta Campbell, Barbara S. Jacobsen, Mathy D. Mezey, Mark V. Pauly, and J. S. Schwartz. "Comprehensive Discharge Planning and Home Follow-Up of Hospitalized Elders: A Randomized Clinical Trial." *JAMA*, vol. 281, no. 7, 1999, pp. 613-620.

For more information, please contact Jelena Zurovac at (609) 275-2383, jzurovac@mathematica-mpr.com, or Randy Brown at (609) 275-2393, rbrown@mathematica-mpr.com.

Mathematica® is a registered trademark of Mathematica Policy Research, Inc.