

Efficient Orthogonal Designs:

Testing the Comparative Effectiveness
of Alternative Ways of Implementing
Patient-Centered Medical
Home Components



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov



Prevention & Chronic Care Program
IMPROVING PRIMARY CARE



Efficient Orthogonal Designs: Testing the Comparative Effectiveness of Alternative Ways of Implementing Patient-Centered Medical Home Components

This brief focuses on using efficient orthogonal designs to evaluate patient-centered medical home (PCMH) models. It is part of a series commissioned by the Agency for Healthcare Research and Quality (AHRQ) and developed by Mathematica Policy Research under contract, with input from other nationally recognized thought leaders in research methods and PCMH models. The series is designed to expand the toolbox of methods used to evaluate and refine PCMH models. The PCMH is a primary care approach that aims to improve quality, cost, and patient and provider experience. PCMH models emphasize patient-centered, comprehensive, coordinated, accessible care, and a systematic focus on quality and safety.

I. Efficient Orthogonal Designs

A recent review by AHRQ (Peikes, Dale, Lundquist et al., 2012) shows that more evaluations are needed to assess and refine PCMH models. The review argues that future evaluations of medical homes should not only address the effectiveness of the entire medical home model, but also examine how to structure the components of a medical home to create the most effective—and cost-effective—“package” of medical home features. Efficient orthogonal designs test the effects of different ways of deploying each component of an intervention as well as how the effects of individual components interact with one another. Such designs can be used to explore the best ways to operationalize the five key features of medical home models outlined by AHRQ: patient-centered, comprehensive, coordinated, accessible care, and a systematic focus on quality and safety (see www.pcmh.ahrq.gov).

Efficient orthogonal designs have been used extensively in manufacturing and other fields but are rarely found in published health care evaluations or effectiveness studies. They provide an opportunity to test interventions in real-world settings, rather than under idealized randomized, controlled trial (RCT) conditions, and to study how intervention components are implemented. These designs are well suited for evaluating and refining the medical home model for three reasons:

- ▲ They incorporate planned variation in implementation of key components of the model, whereas such variation would otherwise occur haphazardly.
- ▲ They combine the rigor of experimental design with the ability to produce rapid results on the effectiveness of many different medical home features in a single experiment.
- ▲ They enable evaluators to directly assess whether more resource-intensive components yield sufficient improvement in outcomes to warrant the investment.

Orthogonal design is not an estimation strategy meant to tease out the effects of ad hoc variations in implementation. Rather, orthogonal design is a study design tool; to use it, evaluators must specify *in advance* the variations in implementation of key model components that they wish to test. The power of this methodology comes from this thoughtful planning.

The remainder of this first section introduces efficient orthogonal designs and describes key design and implementation considerations. Because this approach is fairly technical, some readers may wish to skip to Section II, which illustrates how orthogonal designs can be applied in studies of medical homes. The authors use an example from a study of care management delivery by Special Needs Plans to individuals who are covered by both Medicare and Medicaid. Sections III and IV discuss advantages and limitations of the method and Section V presents conclusions about it.

Introduction to Efficient Orthogonal Designs

Efficient orthogonal designs are experimental designs used to simultaneously compare alternative ways of implementing multiple intervention components. In this brief, we illustrate the power of such designs by describing how one could be used to test five components of a PCMH model, with two alternatives for implementing each component (referred to here as options *a* and *b*). Examples of PCMH components include protocols followed (such as telephoning patients who have not received a preventive service versus emailing them) or level of intensity (such as meeting with a patient to explain key self-care issues once versus doing so twice). The orthogonal design method is implemented by randomly assigning each of the practices participating in the study to a specified combination of the five components. Combinations of the different components would include sequences such as *aabaa*, *baaaa*, *ababa*, and so on.

Select particular orthogonal design. The number of components that can be tested depends on the number of experimental units, or the number of primary care practices in the case of the PCMH. Including all possible combinations, called a “full factorial design,” would require $2n$ practices. In our example, 32 practices would be needed to represent each of the 32 possible combinations of the five components. However, efficient orthogonal designs do not require representing all possible combinations in order to obtain unbiased estimates of the relative effects of options *a* and *b* for each intervention component. The smaller the number of practices used to estimate a given number of effects, the more “efficient” the design is. For example, one such design enables evaluators to estimate the main effects of five components with only eight practices, each implementing a unique combination of options. For any given number of experimental units that are available for a study, published algorithms show various combinations of components that yield orthogonal designs, and illustrate possible tradeoffs between the number of components that can be tested and the extent to which the main effects are indistinguishable from certain interaction effects. After selecting an orthogonal design, the evaluator randomly assigns (without replacement) one combination to each participating practice, which then administers the set of options to all of its patients.

The key feature of these designs that makes the estimates unbiased is orthogonality. Orthogonality means that the assigned combinations are independent of one another. In practice, this means that half of the practices assigned to option *a* of component 1 are assigned to option *a* of component 2 and half to option *b* for component 2, and so on, for all possible pairs of components. Thus, if better outcomes are observed for practices assigned to option *a* than for those assigned to option *b* of component 1, this result cannot be due to the first group of practices being more likely than the second group to deliver option *a* of some other component—the assignments are designed so that this is never the case.

Minimize potential ambiguity of the findings. The efficiency of these designs comes at a cost, however—designs that test a large number of components in a relatively small number of practices cannot distinguish the main effect for any given component from the interaction effects of certain other components. For example, the main effect of component 1 might be completely indistinguishable from the interaction effect of components 2 and 3. The degree of such confounding increases as the number of intervention components being tested increases relative to the number of practices in the study. Evaluators can address this type of confounding if it is reasonable to assume that certain interaction effects, in particular, higher order interaction effects, are negligible. If the interaction effect of other components is considered to be negligible, we can be reasonably certain that the estimate of the main effect of a specific component really is capturing the effect of that component, not the interactive effect of some other components. The extent of such potential confounding depends on the number of practices in the study, relative to the number of components being tested, as described in more detail in the next section. Once the evaluator has carefully selected the set of components for which he/she wishes to test alternative options, an algorithm can be used to generate the relevant combinations. Each participating practice can then be assigned at random to one of the combinations. To summarize, the basic steps involved are:

1. Ascertain the number of practices available for the intervention.
2. If a full factorial design is not possible, determine the interaction effects that are most likely to exist.
3. Select an algorithm that yields an orthogonal design for your desired number of components and the number of practices available, and note the main and interaction effects that are potentially confounded.
4. Order the interventions so that the potentially confounded effects involve only those components for which theory and experience suggest interactive effects are likely to be negligible. For example, if the interaction of components 4 and 5 is potentially confounded with the main effect for component 1, the evaluator would make sure to array the interventions so that the ones placed in the fourth and fifth positions are ones for which the interaction is most likely to be negligible.
5. Randomly assign each practice to one of the combinations specified by the algorithm.

Design Considerations

The most important step in designing a successful orthogonal design study that yields credible, actionable results is to identify intervention components and options that (1) are expected to improve outcomes and/or reduce operational costs, (2) a key decisionmaker is committed to testing, and (3) are feasible to implement on an ongoing basis. Further, when choosing components to test, it is important to consider (1) the likelihood that the practices will faithfully (and without unsustainable oversight) implement the component options to which they are assigned, (2) the likelihood that effective options will be adopted as tested, and (3) the number of components to test.

Assess sample size. When deciding how many components to test, evaluators must first ensure that a sufficient number of practices are available to obtain estimates of main effects that are not confounded

with each other. If more than the minimum number of practices is available to test the desired components, the evaluator must decide whether to assign multiple practices to each of the possible combinations, test more components than originally planned, or select an orthogonal design that requires somewhat more than the minimum number of practices but ensures less confounding of main effects with two-way or higher-order interactions. If the number of practices is limited, as is often the case, the two options are to (1) test more components with greater confounding of main effects with two-way or higher-order interactions, or (2) test fewer components with less confounding.

Estimate statistical power. The evaluator should perform power calculations to determine how many practices are needed to detect substantively meaningful effect sizes. Any test of the medical home requires a clustered design in which practices are assigned to implement a given set of components for all of their patients, so power depends predominantly on the number of practices (not patients) involved in the study. To produce credible evidence about what works best in designing medical homes, studies should include enough practices to provide a high likelihood that true impacts of meaningful magnitude will be detected in the study sample as statistically significant.¹

A key feature of orthogonal designs is that, due to the random assignment of practices to combinations of components, the relative effect of option *a* versus option *b* for any intervention component can be estimated by simply comparing the mean outcome of practices assigned to *a* to the mean for those assigned to *b*. Thus, the standard methods of computing statistical power for clustered designs in RCTs can be used to compute power for orthogonal designs.

Analysis

Evaluators calculate the effect of a given option for implementing an intervention component by comparing the mean outcome over all patients for practices that provide option *a* to the mean for patients of those that provide option *b*. As in any other design, one can use regression analysis with patient-level data to achieve greater precision in estimates of intervention component effects and to control for any pre-intervention differences in patient and practice characteristics across practices. The regressions should account for clustering at the practice level.

Use intent-to-treat approach. The analyses should take an “intent-to-treat” approach, in which component effects are computed by comparing outcomes of those assigned to the two options, regardless of whether or how thoroughly the options were actually delivered. Thus, observations from a practice that was assigned to option *a* but instead implemented option *b*, did not implement that component at all, or implemented something entirely different, should be treated in the analysis as implementing option *a*, as assigned.

Evaluate implementation and fidelity to the intervention. As with any effectiveness study, to fully understand the quantitative results it is important to document how the components were

¹See Peikes, Dale, Lundquist et al. (2011) for a discussion of sample sizes and how measuring certain outcomes among higher-risk patients can improve power to detect effects in medical home evaluations.

implemented and to evaluate fidelity to the planned intervention. Discussions with practice staff involved in the study can help explain why certain components were effective and others were not, and can identify facilitators and barriers to implementation of any component options that were not implemented as planned. This is particularly important, because a finding that option a and option b were equally effective for some component may lead the investigator to conclude that the less expensive of the two options is just as effective as the more expensive one. However, if options a and b were not actually implemented fully, such an inference may be incorrect.

II. Uses of Efficient Orthogonal Designs

Orthogonal designs are being used in a number of applications and in different ways. Three such variants are described here.

Mathematica’s Study of Care Coordination Intervention Options for Special Needs Plans. Two of the authors of this brief are currently conducting a study to help identify care management strategies that work best for members served by three Special Needs Plans— Medicare managed care plans serving disabled or frail elderly patients who are dually eligible for Medicare and Medicaid. The main outcomes of interest are hospital admissions, readmissions, and emergency room visits. One part of the study involves 25 care managers who are implementing either the current approach or an enhanced approach for each of 11 components of care coordination. We selected the components and options to test in collaboration with the care managers and leaders in the participating plans, based on their past experience with care coordination for this population and desire to learn more about the effects of particular care coordination features. We also drew on our own knowledge of various care coordination models being used in Medicare demonstrations and other settings.

For each of the 11 components, we are studying options such as how often the component is provided or which procedures or protocols are used for implementing it. The components we are examining include conducting routine contacts with patients, medication management, depression screening, patient education and coaching, integration of mental health and medical care, and management of care transitions. For example, for the care transitions component we are testing the effectiveness of the plan’s current practice of conducting one followup visit with patients after discharge from an inpatient setting, versus an enhanced option that adds a second followup visit within a week of discharge. Another component introduces the “brownbag method” of reviewing the patients’ medications— having them bring in all of their prescription bottles in a bag—if they are prescribed at least four medications, versus the current ad hoc approach to reviewing medications.

To illustrate how some of the intervention components were specified, the table below shows 4 of the 11 intervention components tested in one Special Needs Plan that serves individuals with severe and persistent mental illness. Each care manager is randomly assigned to a combination of the 11 components, with either option *a* or option *b* specified for each component. For example, one care manager was assigned to implement *aabaaabbbab*, another was assigned to implement *bbaaababbab*, and so on.

Table 1. Intervention Components Testing Ways of Operationalizing Care Management in Special Needs Plan Study

No.	Intervention Component	Option a: Current Approach	Option b: New Approach
1	Frequency of routine contact between care manager and member	Contact frequency based on member risk	More frequent contact (also based on member risk)
2	Involvement of a medical nurse in management of complex medical cases	A medical nurse is involved as needed	Medical nurse is always involved in the management of complex medical cases
3	Follow up during hospital admission and post-discharge (for psychiatric admissions)	Care manager contacts the member during the admission; a licensed mental health professional conducts an in-person followup at discharge and monitoring as needed	Current practice plus additional licensed mental health practitioner followup within a week of discharge and monitoring for 2 months post-discharge
4	Brownbag medication review	No brownbag medication review	A care manager will perform a brownbag medication review for members with more than 4 prescriptions

Possible uses to determine optimal features of medical homes. Efficient orthogonal designs could be used to answer similar questions about how best to operationalize medical home models, such as the following:

1. Should practices be required to have a social worker on the care team to improve access to medical and social support services, or can nurses or less-specialized staff help patients achieve the same level of access?
2. Which alternative care coordination strategies or protocols should care managers follow?
3. Which care transitions models or protocols should a medical home adopt and how?

Several models of transitional care have been shown to be effective in reducing readmissions (Naylor, Brooten, Campbell, et al., 2004; Coleman, Parry, Chalmers, et al., 2006; Jack, Chetty, Anthony, et al., 2009). For example, evidence shows that post-discharge followup is essential to a safe transition from an institution to home; however, little evidence exists about how quickly this followup visit needs to occur, how many visits are needed, and which topics to cover with patients. Further, the optimal approach to post-discharge followup may also differ by the type of primary care practice and the other providers in the medical neighborhood, making studies testing different combinations of approaches in different types of practices tremendously useful.

Using orthogonal design in a multi-phase testing strategy for behavioral health interventions.

A variant of efficient orthogonal designs that is increasingly being used to study and refine complex interventions in behavioral health interventions, such as smoking cessation, is the Multi-Phase Optimization Strategy (MOST) (Collins, Baker, Mermelstein, et al., 2007, 2011). MOST consists of the following three evaluation phases:

1. *Screening phase.* An efficient orthogonal design is used to identify and refine promising components of service packages.

2. *Refining phase.* Another efficient orthogonal study is performed to fine-tune these components (for example, by finding optimal intervention dosages).
3. *Confirmatory phase.* The package of the components from phases 1 and 2 found to be the most effective and efficient is tested in a traditional, large-scale RCT.

Using efficient orthogonal designs in the screening and refining stages allows evaluators and practitioners to test and modify intervention components and implementation strategies to create the most effective intervention package. The confirmatory RCT tests this final product to obtain a rigorous estimate of its effectiveness and learn more about it before the large-scale implementation. The MOST approach could be used to test medical home components. However, this three-phase process increases the time required to fully test a set of intervention components, and some users of efficient orthogonal design find it redundant.

III. Advantages of Efficient Orthogonal Designs

Efficient orthogonal designs are a useful tool for evaluating and refining medical home models, because these models contain many different components and there are many ways of implementing each one. These designs provide an opportunity to test—rigorously and simultaneously—how best to provide the different components. Below we discuss several advantages of these designs over traditional evaluation designs.

Facilitates rapid learning. Efficient orthogonal designs allow evaluators to test the effectiveness of several components of an intervention in a single experiment. This leads to faster learning than would be possible with a traditional RCT, where such learning would be sequential.

Saves resources. The design allows evaluators to test several components with fewer practices than needed to test the same number of components with a traditional RCT.

Compares effectiveness of intervention options scientifically. Randomly assigning practices to different intervention combinations is a more scientifically based approach for drawing inferences about the best way to implement complex interventions than the usual “single model” approach. In that approach (1) an overall model is loosely defined and tested in a group of treatment practices; (2) the study collects qualitative information about implementation, which it uses to group practices into subgroups; and (3) the study compares the outcomes of the subgroups to identify possible associations between intervention components and outcomes.

Enables practices to test options of interest. Primary care practices may have different approaches and strategies for implementing components of medical home models, some of which may be more resource-intensive than others. Orthogonal designs give practices the flexibility to choose the variations they wish to test in advance and to evaluate their relative effectiveness rigorously.

IV. Limitations

Below we discuss some limitations to using efficient orthogonal designs to assess the effects of different medical home features on outcomes.

Requires relatively homogenous practices. Orthogonal designs assume that the practices have relatively homogenous outcomes prior to the study. Generally, the effects of differences among practices in outcomes can be removed by controlling for the differences with a regression model. However, if there are outlier practices that have very different outcomes than the rest of the group before the intervention, they should be excluded from the study to avoid biasing the results. Homogeneity is particularly important for efficient orthogonal designs because they often include relatively few practices, so the results are particularly susceptible to outliers.

Is subject to potential confounding. Testing many components allows for broader learning about how best to implement multi-component care models such as medical homes. However, as a study tests more components relative to the number of practices, it becomes increasingly difficult to identify the main and interaction effects. Evaluators can decide how much confounding to tolerate and which interaction effects to estimate; however, in the most efficient designs, the main effects may be confounded with many two-way interactions of other components. Furthermore, designs that use few practices assume that all three-way or higher-order intervention component interactions are negligible. If that assumption is problematic, the study can either include more practices and use a less efficient but more discriminating orthogonal design, or reduce the number of components being tested in order to identify higher-order interactions. (For example, a study could use eight practices to identify first-order effects of five components, but no higher-order interactions. However, if the study tests only three components, it could use the eight practices to control for the three-way interaction.)

More complex to implement than traditional studies. Even though the implementation challenges of orthogonal design studies are similar to those of other studies, initial resistance from participants may be stronger. Practices that implement the interventions may worry that testing many different ways of operationalizing a PCMH is “too complicated.” Engaging practices in selecting what to test can help increase their willingness to participate. Further, providing implementation guides and individualized assignment sheets can ease concerns about the difficulty of implementing the assigned combination of options. Individualized assignment sheets outline which options each practice should implement and can prevent confusion. Nonetheless, the success of an orthogonal design study relies on practices implementing the assigned component options. Participating practices must be willing to implement those options, regardless of any pre-conceived notions about the most effective approach.

May fail to detect important differences between options. Although efficient orthogonal designs can use a small number of practices to identify the effects of numerous intervention components, these are still clustered designs, in which practices implement the approaches but patient-level outcomes are analyzed. Thus, power to detect outcomes depends primarily on the number of participating practices, and less on the number of patients involved. If relatively few practices participate, power to detect moderate main effects and interactions will likely be low. This limitation exists regardless of the number of intervention components that are being tested. It is particularly important for orthogonal designs because a conclusion of “no effect” can be interpreted to mean that the less expensive option

is just as effective as the more expensive one. Yet if the lack of statistical significance is due to a low-powered test, that conclusion may be incorrect.

Necessitates expert consultation. The method is fairly technical, so researchers should consult with experts in orthogonal designs when designing and analyzing these studies in order to avoid important mistakes and maximize the power of the approach.

Suffers from multiple-test bias. If many components are being tested, the significance level of t-tests of the individual regression coefficients will be understated, leading to potentially overly optimistic conclusions about whether true effects exist. This problem can be addressed by first conducting an overall F-test of whether the difference between option a and option b is zero for all components. If this test is rejected, the evaluator can use the t-tests to draw inferences about which components have meaningful effects. Researchers can also conduct joint tests of subsets of components, such as access or care coordination in medical homes.

V. Conclusion

Although orthogonal designs have some limitations, evaluators and practices can work together to overcome them and harness the value of this approach. By drawing on their combined knowledge of and experience with medical homes and practice behavior, and having frank discussions about the practices' priorities and the components they are able and willing to test, evaluators and implementers should be able to generate practical and rigorous findings about PCMH models.

VI. References

Coleman EA, Parry C, Chalmers S, et al. The care transitions intervention. *Arch Intern Med* 2006;166(17):1822–8.

Collins LM, Baker TB, Mermelstein RJ, et al. The multiphase optimization strategy for engineering effective tobacco use interventions. *Ann Behav Med* 2011;41(2):208–26.

Collins LM, Murphy SA, Strecher VJ. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent health interventions. *Am J Prev Med* 2007;32:S112–8.

Jack BW, Chetty VK, Anthony D, et al. A reengineered hospital discharge program to decrease rehospitalization: a randomized trial. *Ann Intern Med* 2009;150:178–87.

Naylor MD, Brooten DA, Campbell RL, et al. Transitional care of older adults hospitalized with heart failure: a randomized, controlled trial. *J Am Geriatr Soc* 2004;52:675–84.

Peikes D, Dale S, Lundquist E, et al. Building the Evidence Base for the Medical Home: What Sample and Sample Size Do Studies Need? White Paper (Prepared by Mathematica Policy Research under Contract No. HHS2902009000191 TO2). Rockville, MD: Agency for Healthcare Research and Quality; October 2011. AHRQ Publication No. 11-0100-EF.

Peikes D, Zutshi A, Genevro J, et al. Early Evidence on the Patient-Centered Medical Home, Final Report (Prepared by Mathematica Policy Research, under Contract Nos. HHS290200900019I/HHS29032002T and HHS290200900019I/HHS29032005T). Rockville, MD: Agency for Healthcare Research and Quality; February 2012. AHRQ Publication No. 12-0020-EF.

VII. Resources

Orthogonal design methods and theoretical background

Box GE, Hunter SJ, Hunter WG. *Statistics for experimenters: design, innovation, and discovery*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2005.

Fisher RA. *The design of experiments*. Edinburgh: Oliver and Boyd; 1935.

Ledolter J, Swersey AJ. *Testing 1-2-3: experimental design with applications in marketing and service operations*. Stanford, CA: Stanford University Press; 2007.

Applications of orthogonal design in health research

Collins LM, Baker TB, Mermelstein RJ, et al. The multiphase optimization strategy for engineering effective tobacco use interventions. *Ann Behav Med* 2011;41(2):208–26.

Collins LM, Murphy SA, Strecher VJ. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent ehealth interventions. *Am J Prev Med* 2007;32:S112–8.

Jones FG, Moore CH. Designing and executing experiments in care—a data-driven, scientific approach to quality improvement. In: Schoenbaum SC, ed. *Measuring clinical care: a guide for physician executives*. Providence, RI: American College of Physician Executives; 1995. p. 115–25.

Moore CH. Experimental design in health care. *Qual Manag Health Care* 1994;2(2):1–15.

Power in clustered designs

Peikes D, Dale S, Lundquist E, et al. Building the Evidence Base for the Medical Home: What Sample and Sample Size Do Studies Need? White Paper (Prepared by Mathematica Policy Research under Contract No. HHS290200900019I TO2). Rockville, MD: Agency for Healthcare Research and Quality; October 2011. AHRQ Publication No. 11-0100-EF.

This brief was prepared by Zurovac J, Ph.D. (jzurovac@mathematica-mpr.com), Peikes D, Ph.D., Zutshi A, Ph.D., and Brown R., Ph.D. (Mathematica Policy Research).

Suggested Citation: Zurovac J, Peikes D, Zutshi A, Brown R. Efficient Orthogonal Designs: Testing the Comparative Effectiveness of Alternative Ways of Implementing Patient-Centered Medical Home Models. Rockville, MD: Agency for Healthcare Research and Quality. February 2013. AHRQ Publication No. 13-0024-EF.

This brief and companion briefs in this series are available for download from pcmh.ahrq.gov.



AHRQ Publication No. 13-0024-EF
March 2013