HHS Public Access
Author manuscript
Adm Policy Ment Health. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:
Adm Policy Ment Health. 2015 September ; 42(5): 524–532. doi:10.1007/s10488-014-0596-7.

Optimal Design and Purposeful Sampling: Complementary Methodologies for Implementation Research

Naihua Duan,
Division of Biostatistics, Department of Psychiatry, Columbia University; 722 West 168th Street, R206; New York, NY 10032, USA

Dulal K. Bhaumik,
Division of Epidemiology and Biostatistics, University of Illinois at Chicago; Chicago, IL, USA

Lawrence A. Palinkas, and
School of Social Work, University of Southern California, Los Angeles, CA, USA

Kimberly Hoagwood
Department of Child and Adolescent Psychiatry, New York University, New York, NY, USA

Naihua Duan: Naihua.Duan@Columbia.Edu

Abstract

Optimal design has been an under-utilized methodology. However, it has significant real-world applications, particularly in mixed methods implementation research. We review the concept and demonstrate how it can be used to assess the sensitivity of design decisions and balance competing needs. For observational studies, this methodology enables selection of the most informative study units. For experimental studies, it entails selecting and assigning study units to intervention conditions in the most informative manner. We blend optimal design methods with purposeful sampling to show how these two concepts balance competing needs when there are multiple study aims, a common situation in implementation research.

Keywords

A-optimality; Mixed methods; Multiple aims; Optimal allocation; Purposeful sampling; Total survey error; Weighted A-optimality

Introduction

The theory of optimal design is a valuable statistical methodology for designing a wide variety of field studies to optimize their statistical efficiency (see, e.g., Bellhouse 1984; Berger and Wong 2005; Bhaumik and Whittinghill 1991; Bhaumik 1993, 1995; Goos and Jones 2011; Heiberger, Bhaumik, and Holland 1993; Kiefer 1959; Liski et al. 2002; Raudenbush 1997; Raudenbush and Liu 2000; Raudenbush et al. 2011; Shah and Sinha 1989; Spybrook et al. 2011). An optimal design maximizes the information yield for the study, taking into account the budget constraint limiting the resources available for the study. For observational studies, the concept of optimal design can be applied to select study units (organizations, providers, or patients) in the most informative manner. For experimental studies, the same concept can also be applied to assign study units to intervention conditions in the most informative manner.

The application of optimal design is usually based on a pre-specified statistical model (such as a regression model) that is believed to represent the phenomenon being studied. The aims for the study are usually represented as parameters in the pre-specified model, such as the slope parameter in the regression model. The information yield is usually measured using a pre-specified *target criterion* that measures the level of statistical information available for a candidate design, such as the reciprocal for the variance for the estimated slope coefficient for the regression model. This article provides an overview of the potential application of this methodology to facilitate the design of implementation research studies.

There is a close similarity between optimal design and purposeful sampling, a design methodology widely used in qualitative research (Palinkas et al. 2013). There is an important opportunity in mixed methods implementation research to integrate these methodologies for broader applications.

The essence of purposeful sampling is to select information-rich cases for the most effective use of limited resources (Palinkas et al. 2013, Patton 2002). One type of example is the selection of extreme or deviant (outlier) cases for the purpose of learning from unusual manifestation of phenomenon of interest, and the selection of cases with maximum variation for the purpose of documenting unique or diverse variations that have emerged in adapting to different conditions. This approach can be used to identify important common patterns that cut across variations. This type of extreme selection will increase the between-unit variance. Another type of example is the selection of homogeneous cases for the purpose of reducing variation, simplifying analysis, and facilitating group interviewing. This homogeneity is nested within the units chosen in the first type of selection and hence it will reduce the within-unit variability. So between-unit variability is increased by diversification but within-unit variability is reduced by selection of homogeneous cases. Therefore, the reduction of within-unit variability compensates the between-unit variability. As a result, the total variability may be reduced even further after capturing the extreme units. We will discuss below scenarios in which the theory of optimal design leads to designs similar to those resulting from purposeful sampling.

We review the principles of optimal design in the next section, and provide illustrative examples of studies that aim to assess a single parameter. We then review the theory of optimal design for studies with multiple aims, and therefore multiple parameters of interest, to illustrate the important role for optimal design to balance competing study aims. The final section discusses specific issues for mixed methods implementation research studies.

Optimal Design for Single-Aimed Studies

In this section, we review the principles of optimal design, illustrated with several hypothetical implementation studies that aim to assess a single parameter.

Comparison between Two Implementation Strategies

Consider an implementation study that compares two implementation strategies applied to a group of agencies that provides mental health services to families of children with emotional or behavioral disorders. We assume that agencies are recruited and assigned randomly to the two implementations strategies. We assume further that all clinics are similar with respect to factors such as organizational structure, culture and climate, availability of funding and resources, and characteristics of clients. Therefore, we assume a simple two-sample model for the outcomes,

$$Y_{ij}=\mu_i+\varepsilon_{ij}, \quad \varepsilon_{ij}\sim N(0,\sigma^2); \quad i=0,1; j=1,\ldots,n_i \quad (1)$$

The index i denotes the two implementation strategies: i=0 denotes the standard implementation strategy, i=1 denotes the novel implementation strategy being evaluated. For strategy i, the index j denotes the agencies assigned to this strategy; n_i denotes the number of agencies assigned to strategy i. We do not make the usual assumption of equal sample size between the two strategies, for reasons to be discussed later. The outcome for the j-th agency assigned to strategy i is denoted as Y_{ij}; the expected outcome for an agency assigned to strategy i is denoted as μ_i. Furthermore, the random disturbance for the j-th agency assigned to strategy i is denoted as ε_{ij}; the disturbance terms are assumed to be statistically independent and follow a normal distribution with common variance σ^2.

The primary aim for the study is to estimate the *implementation effect*, i.e., the expected difference between the two implementation strategies

$$\delta=\mu_1-\mu_0. \quad (2)$$

We assume for now that the cost to recruit an agency into the study and obtain the outcome measure is constant across agencies and does not depend on the specific strategy the agency is assigned to. We denote the constant cost per agency as C, the total budget available as B. The total number of agencies the study can afford, n_{total}, is therefore

$$n_{total}=\lfloor B/C \rfloor, \quad (3)$$

where $\lfloor B/C \rfloor$ denotes the largest integer smaller than or equal to the fraction B/C, a.k.a. the floor function for B/C. For example, if B = \$40,000, and C = \$3,000, then B/C = 133.333…, and $\lfloor B/C \rfloor$ = 133.

Under these assumptions, it is straightforward to obtain an unbiased estimate for the primary parameter of interest, the *implementation effect* parameter δ in Equation (2), by taking the sampling mean m_1 for the outcomes obtained from agencies assigned to strategy 1, and the sample mean m_0 for the outcomes obtained from agencies assigned to strategy 0, then taking the difference between the two sample means:

$$d = m_1 - m_0. \quad (4)$$

With the availability of the unbiased estimate d in Equation (4), it is reasonable to specify the *target criterion* (TC) to be maximized for the optimal design as the reciprocal of the variance for the unbiased estimate d:

$$TC = 1/Var(d) = 1/[\sigma^2 \times ((1/n_1) + (1/n_0))], \quad (5)$$

where n_1 and n_0 denote the number of agencies assigned to strategy 1 and 0, respectively. More specifically, the optimal design task for this study is to determine n_1 and n_0 that maximize the *target criterion* TC in Equation (5) under the budget constraint

$$n_1 + n_0 = n_{total} = \lfloor B/C \rfloor. \quad (6)$$

It can be shown that the optimal sample allocation is given as follows:

$$n_1 = n_0 = \lfloor B/C \rfloor / 2 \qquad \text{if } \lfloor B/C \rfloor \text{ is an even number;} \quad (7a)$$

$$\{n_1, n_0\} = \{(\lfloor B/C \rfloor - 1)/2, (\lfloor B/C \rfloor + 1)/2\} \quad \text{if } \lfloor B/C \rfloor \text{ is an odd number.} \quad (7b)$$

In Equation (7b), either n_1 or n_0 can take either of the two values shown on the right hand side of the equation. For example, if $\lfloor B/C \rfloor = 133$, then we can either take n_1=61 and n_0=62, or take n1=62 and n_0=61. To avoid unnecessary technical complexity, we simplify Equations (7a-7b) as follows:

$$n_1 \cong (B/C)/2, \text{and} \quad (7c)$$

$$n_0 \cong (B/C)/2, \quad (7d)$$

with the approximate equalities in Equations (7c-7d) indicating the possibility of rounding when (B/C) / 2 is not an integer.

The optimal design shown in Equations (7a-7b/7c-7d) validates the usual design of equal allocation to the two implementation strategies, with the minor deviation in (7b) when the total number of agencies allowed under the budget constraint is an odd number.

However, the assumptions that lead to the optimal design (7a-7b/7c-7d) might be inappropriate for some implementation studies. Deviations from these assumptions might lead to alternative optimal designs. We discuss several such variations in the subsections below.

Unequal cost—The cost per agency might differ between the two implementation strategies, with a higher cost C_1 required for the novel strategy than the cost C_0 for the standard strategy. Under this assumption, a smaller share of agencies should be allocated to the novel strategy due to its higher cost (and lower "bang per buck").

For this scenario, the budget constraint is given as follows:

$$C_1 \times n_1 + C_0 \times n_0 \leq B. \quad (8)$$

The optimal design task is to determine n_1 and n_0 that maximize the *target criterion* TC in Equation (5) under the budget constraint (8).

It can be shown that the optimal sample allocation is given as follows:

$$n_1 \cong B/[\mathrm{sqrt}(C_1) \times \theta], \text{ and} \quad (9a)$$

$$n_0 \cong B/[\mathrm{sqrt}(C_0) \times \theta], \quad (9b)$$

where $\theta = \mathrm{sqrt}(C_1) + \mathrm{sqrt}(C_0)$. Equations (9a-9b) are approximate due to rounding when the right hand sides of these equations are not exact integers.

The optimal sample sizes in Equations (9a-9b), n_1 and n_0, are (approximately) inversely proportional to the ratio of the square root of the costs:

$$n_1/n_0 \cong \mathrm{sqrt}(C_0)/\mathrm{sqrt}(C_1). \quad (10)$$

Assume hypothetically that the cost per agency is higher for the novel strategy than the standard strategy by a factor of four,

$$C_1{:}C_0 = 4{:}1. \quad (11)$$

Following Equation (10), the optimal sample size for the novel strategy is lower (approximately) by a factor of two (the square root of four) compared to the optimal sample size for the standard strategy:

$$n_1{:}n_0 \cong 1{:}2. \quad (12)$$

Unequal residual variance—The residual variance σ^2 might be lower for the novel strategy due to its provisions to address agency level idiosyncrasies, leading to an expanded model,

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, {\sigma_i}^2); \quad i=0,1; j=1,\ldots,n_i; \quad (13)$$

where ${\sigma_i}^2 = \mathrm{Var}(\varepsilon_{ij})$ denotes the residual variance for agencies assigned to the i-th strategy.

Under this scenario, the *target criterion* TC is given as

$$\mathrm{TC}=1/\mathrm{Var}(\mathrm{d})=1/[(\sigma_1{}^2/\mathrm{n}_1)+((\sigma_0{}^2/\mathrm{n}_0))]. \quad (14)$$

The optimal design task is to determine n_1 and n_0 that maximize the *target criterion* TC in Equation (14) under the budget constraint (8).

It can be shown that the optimal sample allocation is given as follows:

$$\mathrm{n}_1 \cong \mathrm{B} \times \sigma_1/[\mathrm{sqrt}(\mathrm{C}_1) \times \theta], \mathrm{and} \quad (15a)$$

$$\mathrm{n}_0 \cong \mathrm{B} \times \sigma_0/[\mathrm{sqrt}(\mathrm{C}_0) \times \theta], \quad (15b)$$

where $\theta = \sigma_1 \times \mathrm{sqrt}(C_1) + \sigma_0 \times \mathrm{sqrt}(C_0)$. The optimal sample sizes, n_1 and n_0, are (approximately) proportional to the respective residual standard deviations, while also inversely proportional to the square root of the respective costs:

$$\mathrm{n}_1/\mathrm{n}_0 \cong \sigma_1 \times \mathrm{sqrt}(\mathrm{C}_0)/[\sigma_0 \times \mathrm{sqrt}(\mathrm{C}_1)]. \quad (16)$$

Assume for now that cost is equal across agencies ($C_1 = C_0$), while agencies assigned to the novel strategy have a lower residual standard deviation by a factor of 1.5:

$$\sigma_1{:}\sigma_0{=}1{:}1.5. \quad (17)$$

It follows from Equation (16) that the optimal sample size would be lower, by a factor of 1.5, for the novel strategy:

$$\mathrm{n}_1{:}\mathrm{n}_0 \cong 1{:}1.5. \quad (18)$$

For a second example, assume the cost ratio in Equation (11), and the ratio in Equation (17) for residual standard deviations. It follows from Equation (16) that the optimal sample sizes would differ by a factor of 3:

$$\mathrm{n}_1{:}\mathrm{n}_0 \cong 1{:}3. \quad (19)$$

The optimal sample size for the novel strategy is lower by a factor of two (Equation (12)) due to its higher cost (Equation (11)), and also lower by a factor of 1.5 (Equation (18)) due to its lower residual standard deviation (Equation (17)). Combining the two factors, the optimal sample size for the novel strategy is lower by a factor of 3 (= 2 × 1.5).

Non-random recruitment and assignment—The assumption of random recruitment and assignment of agencies might not be feasible for many implementation studies, resulting in a potential bias in the estimated *implementation effect*, d, in Equation (4). Therefore, the focus on the variance of the estimated *implementation effect*, Var(d), in Equation (5) is not an appropriate *target criterion* to evaluate the performance for candidate designs. Instead, it would be more appropriate to use the reciprocal of the mean squared error,

$$\mathrm{TC}=1/\mathrm{MSE}(\mathrm{d})=1/\mathrm{E}[(\mathrm{d}-\delta)^2]=1/\{\mathrm{Var}(\mathrm{d})+[\mathrm{E}(\mathrm{d})-\delta]^2\}, \quad (20)$$

as the *target criterion* to be maximized, to take into account the possibility of having bias, E(d) − δ, in the estimated *implementation effect*, d, in addition to taking into account the uncertainty due to Var(d).

While the variance criterion, Var(d), is fairly straightforward to analyze for unbiased estimators, as in Equations (5), the characterization of the MSE criterion in Equation (20) is much more complicated due to the presence of the bias component, E(d) − δ. For further discussions on such a comprehensive framework for the design of survey studies, please see Groves and Lyberg (2010).

Dose-Response

As another example, consider an implementation study that aims to assess the dose-response relationship for an implementation program, such as the relationship between agency's performance improvement (the outcome response) and the intensity (the dose) of agency's participation in a training program, such as the number of staff hours devoted to the program. Assume that the intensity of participation is available from administrative records, while performance improvement needs to be assessed through on-site interviews with agency staff and clients. With limited resources, it is not feasible to assess performance improvement for all agencies. The optimal design task, therefore, is to select the most informative agencies to make the most efficient use of the available resources.

We assume that the dose-response relationship is anticipated to be linear, with homoscedastic variance,

$$\mathrm{Y_j}=\alpha+\beta\times\mathrm{X_j}+\varepsilon_\mathrm{j}, \varepsilon_\mathrm{j}\sim\mathrm{N}(0,\sigma^2); \quad \mathrm{j}=1,\ldots,\mathrm{n}. \quad (21)$$

The performance improvement for the j-th agency is denoted as $\mathrm{Y_j}$; the intensity of participation is denoted as $\mathrm{X_j}$. The intercept (the expected outcome in the absence of participation) is denoted as α. The dose-response parameter (the amount of performance improvement attributable to each unit of participation), denoted as β, is assumed to be the primary aim for the study.

We assume that intensity of participation, X, is statistically independent of residual disturbance, ε, therefore the estimated slope parameter, b, is unbiased for β. Under this assumption, it is reasonable to specify the *target criterion* (TC) to be maximized for the optimal design as the reciprocal of the variance for the estimated slope parameter, b:

$$\mathrm{TC}=1/\mathrm{Var}(\mathrm{b})=(\mathrm{n}-1)\times\mathrm{S}^2(\mathrm{X})/\sigma^2, \quad (22)$$

where $S^2(X)$ denotes the sample variance for the X's among the agencies in the study. The optimal design task is to select the agencies to maximize the *target criterion* in Equation (22) under the budget constraint to be specified below.

For now we assume that the cost C is constant across agencies, so that the budget constraint is given as follows:

$$n = \lfloor B/C \rfloor, \quad (23)$$

similar to the budget constraint in Equation (3) for the first example.

Under this assumption, the sample size n is independent of the composition of the sample selection (which agencies are selected). Therefore the optimal design task (to maximize TC in Equation (22)) is equivalent to selecting the agencies to maximize the sample variance $S^2(X)$. This is achieved by selecting the most extreme agencies, i.e., those with the highest intensities of participation and those with the lowest intensities of participation, to provide the most informative contrast. This is analogous to the strategy of extreme or deviant case, and the strategy of maximum variation, for purposeful sampling.

It is important to note that the strategy of selecting extreme agencies might not be optimal under variations in the underlying assumptions. The subsection below provides an illustration for the contrary when the constant cost assumption is violated.

Unequal cost—Consider a hypothetical example consisting of 100 agencies with the distribution for the intensity of participation (X) shown in Table, with ten agencies at each level of intensity, 0, 1, 2, …, 9:

Assume further that the cost for enrolling each agency is $10,000 for each of the twenty most extreme agencies (the ten agencies with X = 0 and the ten agencies with X = 9); and $2,500 for each of the other 80 agencies with X = 1, 2, …, 8. Assume that the total budget available is B = $40,000.

Under these assumptions, the strategy of extreme or deviant cases will enroll four of the most extreme agencies, namely, two agencies with X = 0 and two agencies with X = 9, utilizing the entire budget of $40,000 ($10,000 for each agency). This design results in the maximum sample variance possible,

$$S^2(X) = 4 \times (4.5)^2 / 3 = 27. \quad (24)$$

Following Equation (22), the *target criterion* for this extreme-agencies design is given as follows:

$$TC = 1/\mathrm{Var}(b) = [3 \times 27]/\sigma^2 = 81/\sigma^2. \quad (25)$$

Alternatively, the study can achieve a more efficient use of the resources by enrolling less extreme agencies that are less costly to enroll. In particular, the optimal design for this study is to enroll none of the most extreme (and more costly) agencies, and instead enroll eight agencies with X = 1 and eight agencies with X = 8, again utilizing the entire budget of $40,000 ($2,500 for each agency), resulting in the following sample variance:

$$S^2(X)=16\times(3.5)^2/15=13.07. \quad (26)$$

The sample variance in Equation (26) is much smaller than the sample variance achieved with the strategy of extreme cases ($S^2(X)$ = 27 given in Equation (24)), because the less extreme agencies provide less statistical information per agency. However, the lower statistical information per agency is more than compensated by the larger sample size available with the less extreme agencies (total n = 16 instead of 4).

Following Equation (22), the *target criterion* for this less-extreme-agencies design is given as follows:

$$TC=1/Var(b)=15\times 13.07/\sigma^2=196/\sigma^2, \quad (27)$$

which is superior to the *target criterion* achieved with the most extreme agencies, 81 / σ^2 given in Equation (25), by a factor of 2.42 (=196/81). Therefore, the selection of the most extreme agencies needs to take into consideration the validity of underlying assumptions, such as the assumption of constant cost across agencies.

Optimal Design for Multiple-Aimed Studies

Many implementation studies have multiple aims and, therefore, multiple parameters of interest. These aims may lead to different designs, making it necessary to seek a balance among the aims. Optimal design provides a useful framework for this balancing task.

As an illustrative example, we expand the first example in the previous section of an implementation study comparing two implementation strategies with the primary aim to assess the *(overall) implementation effect*, and incorporate an additional aim to assess the interaction between the strategies and an agency-level moderator, rurality, i.e., whether the agency is located in a rural community. This second aim is motivated by concerns about a possible disparity for the rural agencies where the novel implementation strategy might work less well due to lack of resources and peer support.

Model for Outcomes

We assume that agencies are assigned randomly to the two implementations strategies, stratified by rurality status, with the following two-way analysis of variance (ANOVA) model for the outcomes:

$$Y_{ijk}=\mu_{ij}+\varepsilon_{ijk}, \qquad \varepsilon_{ijk}\sim N(0,\sigma^2); \qquad i=0,1; j=0,1; k=0,1,2,\ldots,n_{ij}. \quad (28)$$

The index i denotes the implementation strategies as before; the index j denotes the rurality status (=0 for a non-rural agency, =1 for a rural agency). For a strategy-rurality combination ij, the index k denotes the agencies with strategy i and rurality status j; n_{ij} denotes the number of agencies with strategy i and rurality status j. The outcome for the k-th agency with strategy-rurality combination ij is denoted as Y_{ijk}; the expected outcome for an agency with strategy-rurality combination ij is denoted as μ_{ij}. Furthermore, the random disturbance for the k-th agency with strategy-rurality combination ij is denoted as ε_{ijk}.

Study Aims and Parameters

The *implementation effect* among rural agencies is given by

$$\delta_1 = \mu_{11} - \mu_{01}, \quad (29)$$

where μ_{11} denotes the expected outcome for a rural agency (j=1) assigned to the novel implementation strategy (i=1), μ_{01} denotes the expected outcome for a rural agency (j=1) assigned to the standard implementation strategy (i=0). Similarly, the *implementation effect* among non-rural agencies is given by

$$\delta_0 = \mu_{10} - \mu_{00}. \quad (30)$$

The moderating effect for rurality on the *implementation effect*, i.e., the interaction between the effect of rurality and the effect of implementation, is therefore given by

$$\gamma = \delta_1 - \delta_0, \quad (31)$$

the difference between the *implementation effect* among rural agencies (δ_1) and the *implementation effect* among non-rural agencies (δ_0).

The second aim for this expanded implementation study is to assess the moderating effect parameter, γ, given in Equation (31).

To address the first aim for this study, namely, to assess the *overall implementation effect*, we need to specify the *overall implementation effect*. Consider the following thought experiment. If we assign all agencies in the study population to the novel implementation strategy, the expected outcome μ_{1+} is given as follows:

$$\mu_{1+} = P_0 \times \mu_{10} + P_1 \times \mu_{11}, \quad (32)$$

where P_0 denotes the proportion of agencies in the study population located in a non-rural community, and P_1 (= 1 − P_0) denotes the proportion of agencies located in a rural community. Similarly, if we assign all agencies in the study population to the standard implementation strategy, the expected outcome μ_{0+} is given as follows:

$$\mu_{0+} = P_0 \times \mu_{00} + P_1 \times \mu_{01}. \quad (33)$$

The *overall implementation effect* $\delta_{overall}$ is the difference obtained in this thought experiment:

$$\delta_{overall} = \mu_{1+} - \mu_{0+} = P_0 \times \delta_0 + P_1 \times \delta_1. \quad (34)$$

Our first study aim is to assess the *overall implementation effect* parameter, $\delta_{overall}$, given in Equation (34).

Parameter Estimation

To estimate the parameters $\delta_{overall}$ in Equation (34) and γ in Equation (31), we first estimate the expected outcome parameters by the corresponding sample means:

$$\text{Estimate}\,\mu_{ij}\,\text{by}\;m_{ij}, \quad i=0,1; j=0,1; \quad (35)$$

where m_{ij} denotes the sample mean for the outcomes from agencies with rurality status j assigned to implementation strategy i. It follows from Equation (35) that the *implementation effect* for rural agencies, δ_1 in Equation (29), is estimated by

$$d_1 = m_{11} - m_{01}. \quad (36)$$

Likewise, the *implementation effect* for non-rural agencies, δ_0 in Equation (30), is estimated by

$$d_0 = m_{10} - m_{00}. \quad (37)$$

The *overall implementation effect* parameter, $\delta_{overall}$ in Equation (34), is estimated by

$$d_{overall} = P_0 \times d_0 + P_1 \times d_1. \quad (38)$$

The moderating effect parameter, γ in Equation (31), is estimated by

$$g = d_1 - d_0. \quad (39)$$

All of the estimates presented in this subsection are unbiased for the respective parameters.

Budget Constraint

We assume equal residual variance σ^2, as shown in Model (28), and constant cost C per agency, resulting in the following budget constraint:

$$n_{00} + n_{10} + n_{01} + n_{11} \le B/C, \quad (40)$$

where n_{ij} denotes the number of agencies with rurality status j assigned to implementation strategy i, and B denotes the total budget available.

Target Criterion and Optimal Design for Overall Implementation Effect

The optimal design task is to determine the most informative sample sizes, n_{00}, n_{10}, n_{01}, and n_{11}, to maximize the *target criterion* (to be specified below), under the budget constraint (40). With the two study aims, we have several options to specify the *target criterion* for the optimal design.

If we focus entirely on the first study aim to estimate the *overall implementation effect* parameter $\delta_{overall}$ in Equation (34), it is reasonable to specify the *target criterion* (TC) to be maximized as the reciprocal for the variance for the estimate $d_{overall}$ in Equation (38):

$$\mathrm{TC}=1/\mathrm{Var}(\mathrm{d}_{\mathrm{overall}})=1/\{\sigma^2\times[\mathrm{P_0}^2\times(\mathrm{n_{10}}^{-1}+\mathrm{n_{00}}^{-1})+\mathrm{P_1}^2\times(\mathrm{n_{11}}^{-1}+\mathrm{n_{01}}^{-1})]\}. \quad (41)$$

The optimal design task for this scenario is to maximize the *target criterion* TC in Equation (41) under budget constraint (40). It can be shown that the optimal sample sizes are given as follows:

$$\mathrm{n_{00}} \cong \mathrm{P_0} \times (\mathrm{B/C})/2, \quad (42a)$$

$$\mathrm{n_{10}} \cong \mathrm{P_0} \times (\mathrm{B/C})/2, \quad (42b)$$

$$\mathrm{n_{01}} \cong \mathrm{P_1} \times (\mathrm{B/C})/2, \text{and} \quad (42c)$$

$$\mathrm{n_{11}} \cong \mathrm{P_1} \times (\mathrm{B/C})/2. \quad (42d)$$

The optimal design shown in Equations (42a-42d) allocates the rural and non-rural subsamples proportional to the corresponding proportions in the study population. Each subsample is then assigned randomly, 50:50, to the two implementation strategies.

Target Criterion and Optimal Design for Moderating Effect

If we focus entirely on the second study aim to estimate the moderating effect parameter γ in Equation (31), it is reasonable to specify the *target criterion* (TC) to be maximized as the reciprocal for the variance for the estimate g in Equation (39):

$$\mathrm{TC}=1/\mathrm{Var}(\mathrm{g})=1/[\sigma^2\times(\mathrm{n_{10}}^{-1}+\mathrm{n_{00}}^{-1}+\mathrm{n_{11}}^{-1}+\mathrm{n_{01}}^{-1})]. \quad (43)$$

The optimal design task for this scenario is to maximize the *target criterion* TC in Equation (43) under budget constraint (40). It can be shown that the optimal sample sizes are given as follows:

$$n_{00} \cong (1/2) \times (B/C)/2, \quad (44a)$$

$$n_{10} \cong (1/2) \times (B/C)/2, \quad (44b)$$

$$n_{01} \cong (1/2) \times (B/C)/2, \text{and} \quad (44c)$$

$$n_{11} \cong (1/2) \times (B/C)/2. \quad (44d)$$

The optimal design shown in Equations (44a-44d) allocates the rural and non-rural subsamples equally. Each subsample is then assigned randomly, 50:50, to the two implementation strategies.

If the study population happens to have the same proportion of rural and non-rural agencies, i.e., $P_0 = P_1 = 50\%$, the optimal designs for the two study aims coincide, therefore the same design is optimal for both study aims.

Alternatively, if the rural and non-rural agencies do not have the same proportions in the study population, i.e., $P_0 \neq P_1$, the two study aims lead to different optimal designs. For example, if only 20% of the agencies in the study population are located in rural communities ($P_1 = 20\%$), the first study aim allocates 20% of the sample to rural agencies, while the second study aim allocates 50% of the sample to rural agencies. The conflict between the two study aims needs to be balanced, taking into account both target criterion (41) and target criterion (43).

Target Criterion and Optimal Design for Both Effects Considered Jointly

In order to balance multiple study aims, a *composite target criterion* needs to be specified for the optimal design to represent the various study aims, such as the following *composite target criterion* (CTC) for the current example:

$$CTC=1/[w_1 \times Var(d_{overall})+w_2 \times Var(g)], \quad (45)$$

where w_1 and w_2 ($=1-w_1$) denote the importance weights specified for the two respective study aims. Such a *composition target criterion* is known in the optimal design literature as a weighted A-optimality criterion (Shirakura and Tong 1993). The optimal design task for this scenario is to specify the most informative sample sizes n_{00}, n_{10}, n_{10}, and n_{11}, to maximize the *composite target criterion* (CTC) in Equation (45) under budget constraint (40).

The weights w_1 and w_2 in Equation (45) are to be specified according to the relative importance of the two study aims to the overall study. In the extreme, if the first study aim is considered overwhelmingly more important than the second study aim, we should specify w_1 as overwhelmingly larger than w_2. In this situation, the optimal design task is dominated by the component $Var(d_{overall})$ in the *composition target criterion*, resulting in a design that is focused on the first study aim. On the other hand, if the second study aim is considered overwhelmingly more important than the first study aim, we should specify w_2 as overwhelmingly larger than w_1, resulting in a design that is focused on the second study aim. If neither study aim is considered to be overwhelmingly more important than the other, the optimal design will accomplish an appropriate balance between the two study aims.

It can be shown that the optimal sample sizes for this study are given as follows:

$$n_{00} \cong R_0 \times (B/C)/2, \quad (46a)$$

$$n_{10} \cong R_0 \times (B/C)/2, \quad (46b)$$

$$n_{01} \cong R_1 \times (B/C)/2, \text{and} \quad (46c)$$

$$n_{11} \cong R_1 \times (B/C)/2, \quad (46d)$$

where

$$R_0 = Q_0/(Q_0+Q_1), R_1 = Q_1/(Q_0+Q_1), \text{and} \quad (47)$$

$$Q_0 = \text{sqrt}(w_1 \times {P_0}^2 + w_2), Q_1 = \text{sqrt}(w_1 \times {P_1}^2 + w_2). \quad (48)$$

Furthermore, it can be shown that R_0 (R_1) falls between P_0 (P_1) and ½, the leading terms in Equations (42a-42d) and (44a-44d):

$$\text{If } P_0 \leq \frac{1}{2}, \text{then } P_0 \leq R_0 \leq \frac{1}{2} \leq R_1 \leq P_1; \quad (49a)$$

$$\text{If } \; P0 > \frac{1}{2}, \text{then } P_1 < R_1 < \frac{1}{2} < R_0 < P_0. \quad (49b)$$

The format of Equations (46a-46d) is analogous to the format of Equations (42a-42d) and (44a-44d), with the leading terms P_0 in Equations (42a-42b) and P_1 in Equations (42c-42d), and the leading term ½ in Equations (44a-44d), replaced by the leading terms R_0 in Equations (42a-42b) and R_1 in Equations (42c-42d), to reflect the compromise between the two study aims. More specifically, the first study aim for the overall implementation effect results to the leading terms P_0 in Equations (42a-42b) and P_1 in Equations (42c-42d); the second study aim for the moderating effect results in the leading term ½ in Equations (44a-44d); the *composite target criterion*, CTC in Equation (45), results in the leading terms R_0 in Equations (46a-46b) and R_1 in Equations (46c-46d). According to Equations (49a-49b), the leading terms in Equations (46a-46d) falls between the leading terms in Equations (42a-42d) and the leading terms in Equations (44a-44d), therefore the optimal design based on the *composite target criterion*, CTC in Equation (45), indeed can be interpreted as a compromise between the optimal design for the first study aim (Equations (42a-42d)) and the optimal design for the second study aim (Equations (44a-44d)).

Consider a hypothetical example with 20% of the agencies in the study population located in rural communities (P_1 = 20%). The first study aim allocates 20% of the sample to rural agencies, while the second study aim allocates 50% of the sample to rural agencies. Assume that the total budget B is $40,000, the cost per agency is $400, so the budget constraint (40) allows a total sample size of 100. The optimal design (46a-46d) based on the *composite target criterion* (CTC) in Equation (45), and the budget constraint (40), is shown in Table 2 below:

The first row in Table 2 shows the results based on the first study aim: 80% of the sample is allocated to non-rural agencies (40 assigned to each implementation strategy), 20% of the sample is allocated to rural agencies (10 assigned to each implementation strategy). The last row shows the results based on the second study aim: 50% of the sample is allocated to non-rural agencies, 50% to rural agencies. The intermediate rows show the results based on various compromises between the two study aims. As more importance is placed on the second study aim, the optimal sample allocation shifts from the skewed allocation for the first study aim towards the equal allocation for the second study aim.

The use of optimal design with a *composite target criterion* provides an alternative to the widely used practice of making study design decisions based entirely on the primary study aim. While it might be appropriate to focus on a single primary study aim when this study aim is overwhelmingly more important than the other study aims, many implementation studies do have multiple study aims that need to be considered jointly with no single study aim that is overwhelmingly more important than others. For those studies, it would be worthwhile taking the extra efforts and apply the optimal design approach, using an appropriately specified *composite target criterion*, to allow the competing study aims to be considered jointly in the design decision.

Group Randomized Trial

We assumed implicitly in the previous examples that the study is focused on the agency level, therefore the residual disturbances, ε, are statistically independent. Alternatively, many implementation studies have a multi-level structure, say, with agencies nested within networks. Agencies within the same network share common features such as network policy and culture, and thus are likely to be more similar than agencies from different networks. As a result, the residual disturbances are statistically dependent, with intra-cluster correlation, among agencies in the same network. Raudenbush (1997) and Raudenbush and Liu (2000) examined optimal allocation of resources within and between clusters as a function of variance components and costs at each level. Statistical software developed for these trials has been disseminated widely as a methodological resource through the William T. Grant Foundation (Raudenbush et al. 2011; Spybrook et al. 2011). The software enables one to calculate power analysis for individual and group randomized trials.

Discussion

There is a promising potential for synergy in the integration of optimal design and purposeful sampling in mixed methods implementation research. This conjunction can facilitate the selection of information-rich cases to maximize effective use of limited resources. This is especially important in complex implementation studies where there is a need for balance across multiple study aims.

The method of optimal design offers a model for the systematic identification of cases and conditions for qualitative or mixed method research that enable efficient use of in depth insight from limited sample sizes at each stage of implementation. Multi-stage purposeful designs also offer better fit with quantitative experimental and quasi-experimental designs frequently used in implementation research.

While optimal design provides a systematic, quantitative approach to select information-rich cases, this methodology requires a variety of assumptions. The more flexible purposeful sampling may serve as a useful complement to optimal design, to allow a qualitative approach to formulate and evaluate these assumptions. At the same time, the optimal design approach provides a useful framework to assess the sensitivity of design decisions to deviations from the usual assumptions.

Acknowledgments

Drs. Duan, Hoagwood, and Palinkas were funded through a Grant from the National Institute of MentalHealth (P30-MH090322:K.Hoagwood, PI). The authors appreciate helpful comments from two anonymous reviewers.

References

Bellhouse DR. A review of optimal designs in survey sampling. The Canadian Journal of Statistics. 1984; 12:53–65.

Berger, MPF.; Wong, WK., editors. Applied Optimal Designs. New York: Wiley; 2005.

Bhaumik DK, Whittinghill DC. Optimality and robustness to the unavailability of blocks in block designs. Journal of the Royal Statistical Society B. 1991; 53(2):399–407.

Bhaumik DK. On optimal block designs in the presence of a linear trend. Sankhya, Series B. 1993 Apr; 55(1):91–102. 1993.

Bhaumik DK. Majorization and D-dptimality under the nearest neighbor correlation model. Journal of the Royal Statistical Society, Series B. 1995; 57(1):139–143.

Duan N, Alegria M, Canino G, McGuire TG, Takeuchi D. Survey conditioning in self-reported mental health service use: Randomized comparison of alternative instrument formats. Health Services Research. 2007; 42(2):890–907. [PubMed: 17362223]

Frankel MR, Shapiro MF, Duan N, Morton SC, Berry SH, Brown JA, et al. National probability samples in studies of low-prevalence diseases. Part II: Designing and implementing the HIV Cost and Services Utilization Study sample. Health Services Research. 1999; 34(5):969–992. [PubMed: 10591268]

Goos, P.; Jones, B. Optimal design of experiments: A case study approach. New York, NY: Wiley; 2011.

Groves RM, Lyberg L. Total survey error: Past, present, and future. Public Opinion Quarterly. 2010; 74(5):849–879. 2010.

Heiberger RM, Bhaumik DK, Holland B. Optimal data augmentation strategies for additive models. Journal of the American Statistical Association. 1993; 88(423):926–938.

Kiefer J. Optimum experimental designs (with discussions). Journal of the Royal Statistical Society, Series B (Methodology). 1959; 21:272–319.

Liski, EP.; Mandal, NK.; Shah, KR.; Sinha, BK. Topics in Optimal Design. New York: Springer; 2002.

Murray DM, Varnell SP, Blitstein JL. Design and Analysis of Group-Randomized Trials: A Review of Recent Methodological Developments. Am J Public Health. 2004 Mar; 94(3):423–432. 2004. [PubMed: 14998806]

Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and Policy in Mental Health. 2013 Nov 6. 2013. Epub ahead of print.

Patton, MQ. Qualitative research and evaluation methods. 3rd. Thousand Oaks, CA: Sage; 2002.

Raudenbush SW. Statistical Analysis and optimal design for cluster-randomized trials. Psychological Methods. 1997; 2:173–185.

Raudenbush, SW.; Spybrook, J.; Congdon, R.; Liu, X.; Martinez, A.; Bloom, H.; Hill, C. Optimal design plus empirical evidence (Version 3.0). 2011. Available from http://www.wtgrantfoundation.org/resources/optimal-design

Raudenbush SW, Liu X. Statistical power and optimal design for mmultisite randomized trials. Psychological Methods. 2000; 5(2):199–213. [PubMed: 10937329]

Shah, KR.; Sinha, BK. Theory of Optimal Designs. New York: Springer; 1989.

Shirakura T, Tong WP. Weighted A-optimality for fractional 2^m factorial designs of resolution V. Journal of Statistical Planning and inference. 1996; 56:243–256.

Spybrook, J.; Bloom, H.; Congdon, R.; Hill, C.; Martinez, A.; Raudenbush, SW. Optimal design plus empirical evidence: Documentation for the “Optimal Design” software Version 3.0. 2011. Available from http://www.wtgrantfoundation.org/resources/optimal-design

Table 1

Intensity of Participation

Intensity (X)	0	1	2	3	4	5	6	7	8	9
Frequency	10	10	10	10	10	10	10	10	10	10

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Optimal Design with $P_1 = 20\%$, $n_{total} = B / C = 100$

w_1	w_2	n_{00}	n_{10}	n_{01}	n_{11}
1.00	0.00	40	40	10	10
0.80	0.20	32	32	18	18
0.60	0.40	29	29	21	21
0.40	0.60	27	27	23	23
0.20	0.80	26	26	24	24
0.00	1.00	25	25	25	25