Johns Hopkins University, Dept. of Biostatistics Working Papers

6-15-2005

Polydesigns and Causal Inference

Fan Li
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, fli@jhsph.edu

Constantine E. Frangakis
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, cfrangak@jhsph.edu

Suggested Citation
Li, Fan and Frangakis, Constantine E., "Polydesigns and Causal Inference" (June 2005). *Johns Hopkins University, Dept. of Biostatistics Working Papers*. Working Paper 80.
http://biostats.bepress.com/jhubiostat/paper80

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.
Copyright © 2011 by the authors

Polydesigns and Causal Inference

Fan Li and Constantine E. Frangakis

Department of Biostatistics, The Johns Hopkins University

615 N. Wolfe St., Baltimore, MD 21205, USA.

emails: fli@jhsph.edu cfrangak@jhsph.edu

June 15, 2005

SUMMARY. In an increasingly common class of studies, the goal is to evaluate causal effects of treatments that are only partially controlled by the investigator. In such studies there are two conflicting features: (1) a model on the full cohort design and data can identify the causal effects of interest, but can be sensitive to extreme regions of that design's data, where model specification can have more impact; and (2) models on a reduced design (i.e., a subset of the full data), e.g., conditional likelihood on matched subsets of data, can avoid such sensitivity, but do not generally identify the causal effects. We propose a framework to assess how inference is sensitive to designs by exploring combinations of both the full and reduced designs. We show that using such a "polydesign" framework generates a rich class of methods that can identify causal effects and that can also be more robust to model specification than methods using only the full design. We discuss implementation of polydesign methods, and provide an illustration in the evaluation of a Needle Exchange Program.

KEY WORDS. anchor function; causal effects; needle exchange program; partially controlled studies; polydesign; principal stratification.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

1. Introduction

A frequent goal is to evaluate causal effects of treatments in studies that can only be partially controlled by the investigator. Such partially controlled studies have two conflicting features: (1) a model on the full cohort design of data can identify the causal effect of interest, but can be sensitive to extreme regions of that design's data where model specification can have more impact; (2) models induced to data from a reduced design, i.e., from a rule for selecting a subset of the full data (e.g., a conditional likelihood on matched subsets of data), which can avoid such sensitivity, do not generally identify the causal effects.

To make analogies to a motivating example of a partially controlled study, we consider an evaluation of the Baltimore's Needle Exchange Program (NEP), as originating from the combination of the ALIVE and NEP studies (Vlahov et al., 1997; Strathdee et al., 1999). In the ALIVE study, a cohort of injection drug users (IDUs) has been enrolled and is being followed, with regular six month (semester) visits in which the subjects are offered clinic services, including blood tests for HIV. In parallel to the clinic, the NEP study operates sites in Baltimore where IDUs can exchange a used needle for a sterile one, with the hope of reducing HIV transmission. The goal in the NEP evaluation is to estimate the degree to which exchanging versus not exchanging needles reduces (if at all) HIV incidence among IDUs.

The NEP is partially controlled in the sense that (1) it does not directly control either who exchanges needles or who provides outcome – HIV blood tests at the clinic, but (2) it directly controls the placement of the NEP sites offering needles, hence the distance of the NEP sites from the IDUs. If distance of the NEP sites to the subjects influences who exchanges at the NEP and who provides HIV blood tests, these relations can be useful to evaluate the effect that distance has on HIV incidence and that is attributable to exchanging needles. However, a general result for such studies (Frangakis and Rubin, 1999) implies that it is not appropriate to use standard evaluation of the NEP, such as by comparing exchangers with nonexchangers on observed

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

HIV outcomes (e.g., Keende et al., 1993; van Ameijden et al., 1994; Drucker et al., 1998), or by using distance of NEP to IDUs as a standard instrumental variable on the ”complete data” (e.g., McClellan et al., 1994). In contrast, the framework of “principal stratification” (Frangakis and Rubin, 2002) has been shown to allow both definition of more appropriate estimands, and also to allow separation of the assumptions made for the mechanisms of the effects (structural assumptions) from the assumptions made on the design of what data is being collected (e.g., Barnard et al., 2003; Gilbert, Bosch, and Hudgens, 2003; Frangakis et al., 2004).

Part (1) of the conflict stated in the first paragraph arises from the necessary relative complexity of the models involved in principal stratification in partially controlled studies such as the NEP. This complexity makes it important to investigate whether results are robust to the design used. Particular relevance to alternative designs arises also when the outcome of interest is rare. For example, in the NEP study, over an average follow-up of 9 semesters for 1170 subjects, 52 subjects underwent HIV seroconversion. With such low incidence, results can be sensitive to the design and the model used, for example, in the region of controls' covariates that does not overlap with the covariates of the cases. One can argue that if such low incidence is to exacerbate the consequences of a model misspecification, then a possible way to address this can be to change the model to a semiparametric one, as Robins et al. (2001) adopted in other settings. Semiparametric models may or may not have problems with complex structures, depending on the choice of such model. However, in problems with such challenging structure as with principal stratification, current semiparametric formulations cannot guarantee estimability, and miss the advantages of efficiency of parametric models which can be justified even from a non-parametric perspective (e.g., Frangakis and Rubin, 2001). More importantly, the choice of a good design is central for the quality of the results no matter the methods of analysis used. For this reason, we choose to explore sensitivity of the results to different designs, an approach often preferred in epidemiology in other settings (e.g., Breslow and Day, 1980).

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

Specifically, we consider designs for selecting a subset of the full data, and which we call here "reduced designs".

Part (2) of the above stated conflict arises because the information lost in "reduced designs" can be necessary for estimating the causal effects well in partially controlled studies. Suppose, for example, for a "reduced design" in the NEP, we choose all the HIV cases, and, for each case, we keep one control who closely matches the case on covariates, as in nested case-control designs (Langholz and Thomas, 1990). Then the conditional likelihood that the model on the full design induces on the data of the reduced design, unlike with a standard conditional logistic regression likelihood, does not necessarily identify the causal effect, mainly because of a latent structure inherent in the framework of principal stratification. A related inadequacy of the standard conditional logistic regression also arises in the framework of sequential ignorability, as indicated by Robins, Greenland and Hu (1999, Sec. 7). Also, if the reduced design is based on matching on predefined covariate classes so that all probabilities of selection to the reduced design are well away from 0 and 1, then inverse probability of selection methods may recover sufficient lost information for estimation. However, more general settings can include rules of selection to the reduced design conditionally on the full design that resemble more a deterministic process than a probabilistic one, and then inverse probabilty methods are not applicable. The question that arises then is how to explore results from different designs in a way that preserves enough information to well estimate the causal effects.

We propose a class of methods that are based on the combination of the full design together with reduced designs. The key of such "polydesign" methods is that they provide a continuum between the full design and reduced designs. We show that the members of this continuum can identify the causal effects and can achieve better robusteness to model specification than the full design. Section 2 defines polydesigns, discusses inferential perspectives and main general properties. Section 3 discusses implementation of polydesigns; Section 4 demonstrates the

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

methods in an experiment; Section 5 applies the methods to the NEP study, and Section 6 provides remarks.

2. Polydesigns

Consider a cohort design I_{FULL} for selecting units from the population and assigning them to treatments, and let D_{FULL} be the data arising from this design. We call I_{FULL} and D_{FULL} the "full" design and "full" data. Let also $\text{pr}_{\text{FULL}}(D_{\text{FULL}} \mid \theta)$ denote the likelihood of the full data based on the full design, where θ represents the parameters describing the population of study, and where the causal estimand of interest is some function of θ. We consider cases where θ is identifiable from the likelihood of the full data, but where we wish to investigate sensitivity of estimation to different designs.

To assess such sensitivity, consider first a "reduced" design I_{REDU}, that is, a design that describes how to select a subset of the units from the full design I_{FULL}, e.g., through a certain matching rule as in case-control studies (Breslow and Day, 1980). Let D_{REDU} denote the "reduced" data, including indicators of which units have been selected, that arise from this design. For the reduced data, we consider a likelihood

$$\text{pr}_{\text{REDU}}(D_{\text{REDU}} \mid \theta) \tag{2.1}$$

that is induced from the full likelihood when also taking into account the reduced design. We allow that the reduced likelihood can be the marginal likelihood of the reduced data, but also allow that it can be a conditional or a partial likelihood (Cox and Oakes, 1984) of the reduced data, which is the reason for also distinguishing between the distributions pr_{FULL} and pr_{REDU}.

The reduced design can avoid sensitivity to specification of the full model on those features of the full data that are not modeled in the likelihood (2.1). Nevertheless, estimation of the causal effects of interest in partially controlled studies generally requires more information

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

from the full data than does estimation of causal effects in studies that make more assumptions. As a result, a reduced likelihood (2.1) is not generally sufficient to identify the causal effects.

The goal here is to provide an approach that assesses sensitivity to different designs for inference on the parameters in θ for more general partially controlled studies. The idea is that, although a reduced design may not identify the full parameter θ, it identifies some functions of the parameter from (2.1), and thus can provide estimates of these functions that are robust to data omitted from the reduced likelihood (2.1). This point suggests that it is useful to consider the following definitions.

DEFINITION

(a) An *anchor* function $A(\theta)$ is a function that is identifiable from the likelihood of the reduced design (2.1).

(b) A *polydesign* with respect to a full design I_{FULL} of a population is a collection of the full design together with one or more reduced designs I_{REDU}.

An example of an anchor function is the conditional odds ratio (the exponential of the regression coefficient of exposure) that is identifiable from the conditional likelihood of a logistic regression model (Section 4).

In a polydesign, the reduced designs are, by definition, nested within the full design, but do not need to satisfy a nesting structure among each other, and the data they produce can overlap. We focus discussion here on basic polydesigns with one reduced design, although the discussion can be extended to the more general case.

The purpose of a polydesign is to synthesize inference for anchor functions of the parameter, identifiable from the reduced likelihood, with inference for the remaining parts of the parameter from the full likelihood. The former likelihood helps reduce dependence of inference

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

to data omitted from the reduced design, whereas the latter likelihood, if needed, completes identification of the causal effect. Such synthetic inference can be expressed with a Bayesian or a maximum likelihood perspective. In the following, we assume that standard regularity conditions such as for consistency of parametric maximum likelihood estimation hold.

Bayesian perspective.

From the full likelihood and a prior distribution, we obtain the full posterior distribution, $\text{pr}_{\text{FULL}}(\theta|D_{\text{FULL}})$. For an anchor function $A(\theta)$ of the parameters, consider the decomposition of the full posterior distribution, into the marginal distribution of $A(\theta)$ and the distribution of the remaining functions, denoted by $\theta - A(\theta)$, required to specify θ,

$$\text{pr}_{\text{FULL}}(\theta|D_{\text{FULL}}) = \text{pr}_{\text{FULL}}(\theta - A(\theta) \mid A(\theta), D_{\text{FULL}})\, \text{pr}_{\text{FULL}}(A(\theta)|D_{\text{FULL}}). \tag{2.2}$$

Also, from the reduced likelihood and the prior distribution, we can obtain the reduced posterior distribution, $\text{pr}_{\text{REDU}}(\theta|D_{\text{REDU}})$, as if we had only observed reduced data D_{REDU}. For an anchor function $A(\theta)$, the resulting reduced posterior distribution $\text{pr}_{\text{REDU}}(A(\theta)|D_{\text{REDU}})$ is not sensitive to specification of the prior in the sense that $A(\theta)$ is identifiable from the reduced likelihood. It is then helpful to consider the "polydesign" distribution defined as

$$\text{pr}_{\text{POLY}}(\theta) = \quad \text{pr}_{\text{FULL}}(\theta - A(\theta) \mid A(\theta), D_{\text{FULL}})\text{pr}_{\text{REDU}}(A(\theta)|D_{\text{REDU}}). \tag{2.3}$$

The distribution (2.3) is a recalibration of (2.2) to the posterior distribution of the anchor function that arises by the reduced design alone. In the special case where the reduced likelihood $\text{pr}_{\text{REDU}}(D_{\text{REDU}} \mid \theta)$ is a marginal distribution of the full likelihood, the factor $\text{pr}_{\text{REDU}}(A(\theta)|D_{\text{REDU}})$ of (2.3) is proportional to an integrated likelihood.

The polydesign distribution can form a basics of estimating θ in a number of ways. To demonstrate, we discuss here operating characteristics of the median of the distribution, although similar results hold for more general summaries of this distribution. For any given

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

polydesign I_{POLY}, Appendix 1 outlines a proof of the property for a parameter θ that can be well approximated with fine enough discretizations.

PROPERTY 1

(a) If the assumed full likelihood is correct then, for any anchor function, the median of the polydesign distribution (2.3) is consistent for the true parameter θ_0.

(b) For any misspecification of the full likelihood, and with respect to any loss function for a true parameter, there exists an anchor function so that the median of the polydesign distribution (2.3) performs uniformly at least as well as the median of the posterior distribution (2.2).

A proof of the first part is based on showing that the right multipliers in the right hand side of both distributions (2.3) and (2.2) converge to the point mass at the anchor function's value at θ_0. The proof of the second part is straightforward because the class of distributions (2.3) generated by a polydesign provides a continuum that contains the distribution (2.2) in the trivial case when $A()$ is free of θ. In fact, part (b) would hold even if we relaxed identifiability of $A(\theta)$ from the reduced likelihood, which is, nevertheless, desired to allow for use of diffuse prior distributions, and, hence, for more dependence on the data. Moreover, the optimum can be achievable within the continuum, i.e., with a nonconstant anchor function. Finding an analytic function giving the optimum depends on how a misspecification is expressed, and is not here our goal. A more practical implication of the property is that use of an anchor function $A(\theta)$ in the polydesign will focus its estimation to be from the reduced design and to not depend on the structure of the remaining data.

Maximum likelihood perspective.

Analogously to the Bayesian perspective, we consider a synthetic estimation of $\theta = (\theta -$

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

$A(\theta), A(\theta))$ by basing estimation for the anchor function $A(\theta)$ on the reduced likelihood, and estimation for the remaining part of the parameter on the full likelihood. Specifically, define:

$$\theta_{\text{POLY}} := ((\theta - A)_{\text{FULL}},\ A_{\text{REDU}}), \quad \text{where} \tag{2.4}$$

$$\text{pr}\{D_{\text{REDU}} \mid [\theta - A(\theta), A(\theta)]\} \quad \text{is maximum at} \quad [(\theta - A)_{\text{REDU}},\ A_{\text{REDU}}], \text{ and}$$

$$\text{pr}\{D_{\text{FULL}} \mid [\theta - A(\theta), A_{\text{REDU}}]\} \quad \text{is maximum at} \quad [(\theta - A)_{\text{FULL}},\ A_{\text{REDU}}]$$

The estimator θ_{POLY}, therefore, is the maximizer of the full likelihood after having profiled (Murphy and van der Vaart, 2000) that likelihood on the value of the anchor function that maximizes the reduced likelihood. Standard errors for θ_{POLY} can be obtained by the delta method, by bootstrap, or by simulation from the model. In the special case where the reduced likelihood $\text{pr}_{\text{REDU}}(D_{\text{REDU}} \mid \theta)$ is a marginal distribution of the full likelihood, the estimator $A(\theta)_{\text{REDU}}$ is a "marginal" (or "restricted", Patterson and Thompson, 1971) likelihood estimator. The reason why A_{REDU} but not $(\theta - A)_{\text{REDU}}$ is used in θ_{POLY} is because the latter is not generally unique and not necessarily a good estimate.

In analogy to the Bayesian perspective, for any given polydesign I_{POLY}, the following property holds for a sufficiently fine discretized approximation of the parameter.

PROPERTY 2

(a) If the assumed full likelihood is correct then, for any anchor function, the polydesign estimator θ_{POLY} defined in (2.4) is consistent for the true parameter θ_0.

(b) For any misspecification of the full likelihood, and with respect to any loss function for a true parameter, there exists an anchor function so that the polydesign estimator θ_{POLY} performs uniformly at least as well as the maximum likelihood estimator of the full likelihood.

Part (a) can be shown using a variation of Wald's proof for consistency of maximum likelihood. The proof of part (b) is again straightforward by analogous arguments to those of

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

Property 1(b).

Thus, polydesigns offer estimation, from either the Bayesian or the maximum likelihood perspective, that shares qualitatively two main properties. It provides an alternative inference to the full design that (a) is valid if the assumed model is correct; and (b) focuses estimation of a selected component of the estimand, namely the anchor function, to be based on a reduced design of the full data. Therefore, polydesigns also offer a way to assess how sensitive the results are to using different designs to estimate different components of the model. If results differ substantially between the full design and the polydesign, this would suggest the model is mispecified. On the other hand, similarity between the results of the full design with those of a polydesign with an anchor function that depends on the estimand, would be more reassuring than simply relying of results of the full design. This does not mean that any polydesign gives better inference than any full design, and justification in terms of principles such as matching is needed to make a case for choosing a polydesign.

3. Implementation

Obtaining the value of θ_{POLY} defined in (2.4) is a maximization problem, although possibly a challenging one in terms of computational stability if the anchor function is complicated.

In contrast, the polydesign distribution (2.3) can be relatively easily obtained through the following simulation.

Step 1: Approximate $\text{pr}_{\text{FULL}}(A(\theta)|D_{\text{FULL}})$. To do so, one can simulate a large number, N, of random draws from the normal approximation based on the MLE of θ of the full likelihood, and then simulate $N_{\text{FULL}}(< N)$ random draws from $\text{pr}_{\text{FULL}}(\theta|D_{\text{FULL}})$ using sampling importance resampling (SIR, Rubin, 1987). For each draw, calculate $A(\theta)$ and estimate $\text{pr}_{\text{FULL}}(A(\theta)|D_{\text{FULL}})$, e.g., with a kernel approximation.

Step 2: Approximate $\text{pr}_{\text{REDU}}(A(\theta)|D_{\text{REDU}})$. Because maximization of the reduced likelihood

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

to obtain a normal approximation may be unstable with multiple modes of θ, one can start from the random draws of $\text{pr}_{\text{FULL}}(\theta|D_{\text{FULL}})$ obtained in step 1, and then simulate $N_{\text{REDU}}(< N_{\text{FULL}})$ random draws from $\text{pr}_{\text{REDU}}(\theta|D_{\text{REDU}})$ using SIR. For each draw, calculate $A(\theta)$ and then estimate $\text{pr}_{\text{REDU}}(A(\theta)|D_{\text{REDU}})$ as in step 1.

Step 3: Approximate the function $r(\theta) = \text{pr}_{\text{REDU}}(A(\theta)|D_{\text{REDU}})/\text{pr}_{\text{FULL}}(A(\theta)|D_{\text{FULL}})$, which is the importance ratio of the polydesign distribution $\text{pr}_{\text{POLY}}(\theta)$ in equation (2.3) to $\text{pr}_{\text{FULL}}(\theta|D_{\text{FULL}})$ in equation (2.2). This is obtained by simply dividing the two functions obtained in steps 1 and 2.

Step 4: Simulate from the polydesign distribution $\text{pr}_{\text{POLY}}(\theta)$. To do so, use simulations of $\text{pr}_{\text{FULL}}(\theta|D_{\text{FULL}})$ from step 1, and use SIR by noting that the ratio of the target (polydesign) to candidate (full) distributions is the function $r(\theta)$ of step 3. When enough draws from the target distribution have been obtained, quantiles, moments, and other summaries can be computed.

The fourth step is an alternative description, from an implementation perspective, that the polydesign distribution is a recalibration of $\text{pr}_{\text{FULL}}(\theta|D_{\text{FULL}})$ to be such that the marginal distribution of the anchor function $A(\theta)$ be equal to the posterior distribution arising from the reduced likelihood $\text{pr}_{\text{REDU}}(A(\theta)|D_{\text{REDU}})$.

4. Demonstration in a simple example

In this section, we demonstrate the polydesign in a simple experimental example as also suggested by reviewers.

Estimand. Consider individuals each of which can be assigned one of two treatments, $d = 0, 1$, in which case the potential outcomes (Rubin, 1974, 1978) will be $Y_i(d)$, which we assume here are binary. For example, the "treatment" can be a medical therapy, or the location of the NEP site, depending on the context, and the outcome can be living status or HIV, respectively. We

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

wish to estimate the marginal causal odds ratio Q:

$$Q = \frac{\text{pr}(Y_i(1) = 1)/\text{pr}(Y_i(1) = 0)}{\text{pr}(Y_i(0) = 1)/\text{pr}(Y_i(0) = 0)} \tag{4.1}$$

Because treatment is not generally randomized, it is important to express Q using distributions conditional on a covariate,

$$\begin{aligned} Q(\theta_0) = & \frac{\sum_x \text{pr}(Y_i(1) = 1 \mid X_i = x, \theta_0)\text{pr}(x \mid \theta_0)}{\sum_x \text{pr}(Y_i(0) = 1 \mid X_i = x, \theta_0)\text{pr}(x \mid \theta_0)} \\ & \times \frac{\sum_x \text{pr}(Y_i(0) = 0 | X_i = x, \theta_0)\text{pr}(x \mid \theta_0)}{\sum_x \text{pr}(Y_i(1) = 0 \mid X_i = x, \theta_0)\text{pr}(x \mid \theta_0)}, \end{aligned} \tag{4.2}$$

where θ_0 denotes the true joint distribution of all variables. For this section, we allow that the actual assignment of treatments D_i is ignorable only after conditioning on the covariate X_i in the sense of Rubin (1978). Then, the potential outcome distributions $\text{pr}(Y_i(d) = 1 \mid X_i = x, \theta_0)$ are equal to $\text{pr}(Y_i^{obs} = 1 | D_i = d, X_i = x, \theta_0)$, which involve only observable data (Rubin, 1978), and so by replacing the former with estimates of the latter in (4.2) allows in principle estimation of Q. We use this example to reflect the situation where, although a standard case-control design can estimate some interesting aspects of the distribution, the estimand of interest, here Q, is only partly specified by those aspects, and needs also some other aspects from the full design, as we see below.

Model, and full- and polydesign methods used for analyses. A discrete three level (0.0, 0.5, and 1.0) variable is used as the covariate X. The model that is used here for analysis assumes that potential outcomes conditionally on X have the logistic regression

$$\text{logit}\,\text{pr}(Y_i(d) = 1 | X_i = x, \theta) = g(x, \theta) + \beta_d\, d \tag{4.3}$$

$$\text{where } g(x, \theta) = \beta_c + \beta_x\, x. \tag{4.4}$$

To estimate the odds ratio (4.2), here we discuss a full design and a polydesign method from the maximum likelihood perspective, because it is feasible and makes easier connection

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

to existing case-control arguments. The full design is a simple random sample of individuals from the population. Using this design, the odds ratio (4.2) is estimated by its MLE $Q(\theta_{\text{FULL}})$ by replacing the MLE of β for the outcome distributions given the covariate, and the observed proportions of the distribution of the covariate X.

For the polydesign method, we consider the same individuals of the full design, and, among those, we form a nested case-control design, by matching every case from the full design, $Y^{obs} = 1$, to a control from the full design with the same covariate. Note that since polydesings can combine different likelihoods of the model corresponding to different designs, polydesigns are not necessarily restricted to 1-1 matching, and one could use also a polydesign with more-than 1 to 1 matching for each case. The 1-1 matching here, though, better represents practical situations where at most one control can be well matched.

By standard results under this design and the assumed model, β_d is estimable from the conditional (reduced) likelihood. Unlike a usual setting, though, the reduced likelihood does not fully identify the estimand Q, because Q also involves other parameters. Instead, β_d can be used as an anchor to obtain a polydesign estimator as in (2.4). Specifically, to do this, we (i) estimate β_d by the maximizer, $\beta_{d,\text{POLY}}$, of the conditional likelihood of the case-control design; (ii) estimate β_c, β_x by the maximizers $\beta_{c,\text{POLY}}, \beta_{x,\text{POLY}}$ of the full likelihood after having replaced $\beta_{d,\text{POLY}}$ for β_d; (iii) estimate the distribution of the covariate X by the observed proportions; and (iv) estimate the odds ratio (4.2) by $\hat{Q}(\theta_{\text{POLY}})$ by replacing the estimates from (i)-(iii) in (4.2).

Experiment and results. By the virtue of matching on X, the anchor β_{POLY} is estimated consistently by the polydesign estimator, but not by the full design estimator, whether or not the function $g(x, \theta)$ is correctly specified. Moreover, although $Q(\theta)$ depends on the function $g(x, \theta)$ and not just on β_d, under the null hypothesis $H_0 : \beta_d = 0$, the odds ratio $Q(\theta)$ is also 1. By continuity, this implies that, at least around H_0 , $Q(\theta)$ will be estimated will less bias by the polydesign $Q(\theta_{\text{POLY}})$ than by the full design $Q(\theta_{\text{FULL}})$.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

We examined how different the performance can be between the two estimators in conditions that satisfy the model (4.4), and in conditions where the model is violated, as follows and as reflected in Table 1. For the former conditions, we set pr$(Y(0) \mid X = x)(=$ expit$(g(x))) = (1\%, 8\%, 40\%)$, for $x = 0.0, 0.5, 1.0$ respectively, which satisfies approximately the logistic linear model (4.4) with $\beta_x = 2.15$; to reflect conditions violating the model, we interchange the second and third values of the above probability, that is, we set pr$(Y(0) \mid X = x)(=$ expit$(g(x))) = (1\%, 40\%, 8\%)$ for $x = 0.0, 0.5, 1.0$ respectively. For each such condition, we vary the conditional odds ratio $\beta_d = 0.5$, 1.00 (null H_0), and 2.0. The parameters we fix across conditions are: a total sample size of 1000 individuals; the distribution of the covariate, pr$(X_i = x) = (80\%, 10\%, 10\%)$ and the probability of assignment to treatment $d = 1$, pr$(D_i = 1 \mid X_i = x) = (10\%, 50\%, 20\%)$, for $x = (0.0, 0.5, 1.0)$ respectively, to induce rare prevalence and non-linearity of assignment. For the resulting conditions, Table 1 gives, the induced true marginal odds ratio Q; the median of the full design and polydesign estimators for Q, $Q(\theta_{\text{FULL}})$ and $Q(\theta_{\text{POLY}})$ respectively; the standard deviation (empirical, over the simulations) of the estimators for $\log Q$ and their mean squared error; and the coverages of $\log Q$ (equal to those of Q) by the nominal 95% confidence intervals using the empirical standard deviations of the estimators. The results are based on 2500 simulations.

In the conditions of Table 1 where the model (4.4) is correctly specified, both the full design and polydesign methods give, as expected, medians essentially equal to the true odds ratio Q. Moreover, the full design estimator is associated with smaller root MSEs, which reflects the fact that full design method makes more use of the data and more use of the model assumptions. On the contrary, in the conditions of Table 1 where the model (4.4) is misspecified, the full-design method has a poor median under both null and non-null effects, whereas the polydesign method is relatively correctly centered. Moreover, even though the method based on the full design has smaller variance, it has poor overall accuracy as judged by the root MSEs. In contrast, the

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

increased dependence of the polydesign method on the reduced design allows it to preserve relatively good overall accuracy.

In summary, while there are settings where the polydesign does not perform as well as the full design, it can perform better when model misspecification becomes important for the target estimand.

5. Application to Needle Exchange Program Evaluation

5.1 Causal estimand and full design

We now return to the NEP study introduced in Section 1. We first give some further background needed to define the causal effect of interest and the full likelihood in terms of potential outcomes; any omitted details from here on the background and assumptions can be found in Frangakis et al. (2004).

For a person-semester unit i, a needle exchange site can be placed at distances indexed by $d = 1, ..., d_{max}$. Define the principal stratum S_i to be the shortest distance of the needle exchange site, $S_i = 0, ..., d_{max}$, beyond which that person would not exchange needles at that site during that semester. That is, the person's exchange indicator $E_i(d)$ equals 1 if and only if $d \leq S_i$. Denote $Y_i(d), C_i(d)$ to be the potential outcomes of HIV status and of censoring of the HIV status, respectively, if the NEP site is placed at distance d. Also denote D_i to be the actual distance between the nearest NEP site and the person; $E_i^{obs} = E_i(D_i)$ the observed exchanging status; $C_i^{obs} = C_i(D_i)$ the observed censoring status; $Y_i^{obs} = Y_i(D_i)$ the observed HIV status if not censored; and H_i the observed history of that person at that semester which includes a baseline risk score B_i of getting HIV.

The causal effect of interest, Q, is the odds ratio, within stratum $S_i = s, (0 < s < d_{max})$, of getting HIV when exchanging versus when not exchanging needles. This, by definition of the principal stratification, is obtained by comparing between the levels of the controlled factor of

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

distance, that is, Q is the odds of

$$\text{pr}(Y_i(d=1)=1 \mid S_i, H_i) \quad \text{versus} \quad \text{pr}(Y_i(d=d_{max})=1 \mid S_i, H_i),$$

which, for simplicity is assumed here to be common across principal strata S_i and history H_i. Note that this definition involves a variable, the principal stratum S_i, that is generally not observed. For example, if at a semester, a person is at distance $D_i^{obs} = 2$ from the NEP and does not exchange needles ($E_i^{obs} = 0$), then we do not know that person's principal stratum because we do not know exchange behavior, $E_i(d=1)$, if the NEP had been placed at distance 1: a behaviour $E_i(d=1)=0$ would imply that there would be no distance at which the person would exchange, so $S_i = 0$; whereas a behaviour $E_i(d=1)=1$, and knowing that $E_i^{obs}=0$ at distance $D_i^{obs}=2$, would imply that $d=1$ would be the shortest distance beyond which (i.e., ≥ 1) the person would not exchange, so S_i would be 1.

The full design here is a cohort design of $n = 1170$ individuals, representative of those IDUs originally consenting to be part of the study. The full data are history H_i, distance D_i, exchanging status E_i^{obs}, censoring status C_i^{obs}, and HIV status Y_i^{obs} if not censored, for all person-semesters i up to the first censoring time. Models for the full data can be specified conditionally on the principal strata as follows:

$$f(y,s,d,h;\theta) = \text{pr}(Y_i(d)=y \mid S_i=s, H_i=h, \theta) \tag{5.1}$$

$$g(c,s,d,h;\theta) = \text{pr}(C_i(d)=c \mid S_i=s, H_i=h, \theta) \tag{5.2}$$

$$k(s,h;\theta) = \text{pr}(S_i=s \mid H_i=h, \theta). \tag{5.3}$$

Because location of the NEP sites is controlled by the study, the distribution of D_i at each time is assumed ignorable (Rubin, 1978) conditionally on the observed history, so a model for that distribution is avoided. Under the remaining assumptions of Frangakis et al. (2004), the full

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

likelihood $\text{pr}_{\text{FULL}}(D_{\text{FULL}} \mid \theta)$ is

$$\prod_i \{\sum_{s_i} f(y_i, s_i, d_i, h_i; \theta)^{(1-c_i)} g(c_i, s_i, d_i, h_i, ; \theta) k(s_i, h_i; \theta)\}, \tag{5.4}$$

where the summation is over all possible s_i values that are consistent with the observed exchange status E_i^{obs} and observed distance D_i for the person-semester unit i.

The first row in Table 2(a) presents results from the posterior distribution of the causal effect Q, obtained from the above full likelihood with a noninformative prior, with two levels of distance (greater or less than 3 miles). Specifically, the median of the effect is estimated as a 91% decrease of HIV incidence attributable to exchanging needles (odds ratio 0.09, 95% posterior interval $(0.00, 0.86)$). The first row of Table 2(b) gives results for the odds ratio of getting HIV under fixed exchange, for comparing higher versus lower principal strata, given fixed history, that is, for comparing between $\text{pr}(Y_i(d+1) = 1 \mid S_i = (d+1), H_i)$ and $\text{pr}(Y_i(d) = 1 \mid S_i = d, H_i)$; and the second row of Table 2(c) gives results for the odds ratio of being at higher (versus lower) principal strata for one s.d. increase in baseline risk score, that is, for comparing between $\text{pr}(S_i \geq d \mid H_i + 1)$ and $\text{pr}(S_i \geq d \mid H_i)$, where $H_i + 1$ is H_i except for one s.d. increase in its baseline risk score.

5.2 Evaluation with polydesign

We now demonstrate the use of a polydesign in the NEP. Specifically, the example of a reduced design I_{REDU} we consider here is a nested case-control design (Langholz and Thomas, 1990), in the sense that it keeps all the uncensored HIV cases, and for each case, selects an uncensored control that closely matches the case on its history H_i. Thus the corresponding reduced data D_{REDU} consist of the observed data of 52 uncensored case-control pairs of IDUs.

We use this design for two reasons. First, because the design matches cases to controls on the history characteristics, it reduces extrapolation, as in other settings (e.g., Rubin and

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

Thomas, 1996), of the models on the range of those characteristics for all the other controls, and is therefore expected to be more robust in detecting the existence of treatment effects. Second, because the reduced likelihood of the above reduced data is not only on the HIV outcome Y^{obs}, but also on the exchange behavior, conditionally on distance, it carries information on both the effect of distance on exchange and the effect of distance on HIV, which is information that puts constraints on the estimand Q. This was supported by evaluating the reduced likelihood at different values of the nuisance parameters, and confirming that it varied as a function of Q.

For an uncensored case-control pair (a, b), where a is the case, i.e., $Y_a^{obs} = 1$, we define a conditional likelihood $\pi_{ab}(\theta)$ by

$$\text{pr}(Y_a^{obs} = 1, E_a^{obs}, Y_b^{obs} = 0, E_b^{obs} \mid D_a, D_b, H_a, H_b, C_a^{obs} = 0, C_b^{obs} = 0, Y_a^{obs} + Y_b^{obs} = 1, \theta), \tag{5.5}$$

where the conditionals on $C_a^{obs}, C_b^{obs}, Y_a^{obs}, Y_b^{obs}$ reflect the no-censoring and outcome constraints of this design. Then, we take $\text{pr}_{\text{REDU}}(D_{\text{REDU}} \mid \theta)$ to be the product of the terms $\pi_{ab}(\theta)$ in (5.5) over all case-control pairs. To express this product in terms of the models (5.1)-(5.3), first denote the probabilities $\text{pr}(Y^{obs} = y, E^{obs} = e | D = d, H = h, C^{obs} = 0, \theta)$ as $\tau(y, e, d, h; \theta)$, then the likelihood (5.5) can be expressed as

$$\frac{\tau(1, E_a^{obs}, D_a, H_a; \theta)\tau(0, E_b^{obs}, D_b, H_b; \theta)}{\tau(1, E_a^{obs}, D_a, H_a; \theta)\tau(0, E_b^{obs}, D_b, H_b; \theta) + \tau(0, E_a^{obs}, D_a, H_a; \theta)\tau(1, E_b^{obs}, D_b, H_b; \theta)}, \tag{5.6}$$

where by the law of total probability over the principal strata S, and by the assumptions of the models (5.1)-(5.3) in Frangakis et al. (2004), each $\tau(y, e, d, h; \theta)$ can be expressed as

$$\tau(y, e, d, h; \theta) \;=\; \frac{\sum_s f(y, s, d, h; \theta) g(0, s, d, h; \theta) k(s, h; \theta)}{\sum_s g(0, s, d, h; \theta) k(s, h; \theta)}, \tag{5.7}$$

where the summation is, again, over all s values that are consistent with observing exchange e at distance of d.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

Relations (5.6) and (5.7) link the reduced likelihood to the full likelihood (5.4). From these expressions, it is evident that the reduced likelihood is generally not sufficient to consistently estimate the causal effect Q. The reason is that the design's constraints on the outcome and no-censoring in the reduced likelihood lose some information that is, in general, necessary for identifying all parameters. Moreover, due to the summation over the principal strata s in (5.7), this problem is not removed when making additive assumptions to models (5.1)-(5.3) as is usually done in conditional logistic regression models.

For this reason, and in order to make use of the reduced likelihood, we use a polydesign. For an anchor function, consider the entropy-related expectation, $A^*(\theta) = E_{\theta_0}(\log(\pi_{ab}(\theta)))$. The function $A^*(\theta)$ is an anchor function, because $A^*(\theta)$ is identifiable from the reduced likelihood (2.1), since $\pi_{ab}(\theta_1) = \pi_{ab}(\theta_2)$ implies $A^*(\theta_1) = A^*(\theta_2)$ for any θ_1, θ_2. However, we do not know the true θ_0, and hence do not know $A^*(\theta)$ exactly. For this reason, we use its empirical counterpart, the logarithm of the reduced likelihood, $A(\theta)$,

$$A(\theta) := \frac{1}{n_c} \sum_{(ab)} \log(\pi_{ab}(\theta)), \tag{5.8}$$

to approximate $A^*(\theta)$, where n_c is the number of cases. There is no guarantee that $A(\theta)$ defined in (5.8) is exactly identifiable from the reduced likelihood. Nevertheless, $A(\theta)$ converges in probability to $A^*(\theta)$, and the region around $A(\theta_0)$ where $A(\theta)$ is possibly equally likely as $A(\theta_0)$ is expected to be small, so using the empirical function $A(\theta)$ as an actual anchor function is expected to give similar results as using $A^*(\theta)$. Practically, the choice of $A(\theta)$ as an anchor function will calibrate the full design inference to values of θ that give the same entropy to the reduced data as if θ is inferred only using the reduced data.

The second and the third rows of Table 2 present results from using the reduced design and the polydesign, respectively. Specifically, the second rows are obtained from the simulated draws from the reduced likelihood as described in Step 2 of Section 3. The third rows are

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

obtained by simulating from the polydesign distribution (2.3) using the above full and reduced likelihoods and anchor function, and are implemented as described in Section 3.

For the odds ratio Q, the second row of Table 2(a) shows that the point estimate from the reduced design is equal to that from the full design, but with slightly larger uncertainty, which results from using only part of the full data in the reduced design. This similarity suggests that the reduced data here contain most of the information for estimating the parameters. The third row of Table 2(a) shows that both the median and the posterior intervals from the polydesign are similar to that from the full design. The above observations also stand for the two other estimands. Those observation, and the fact that the chosen anchor function was not variation-independent of the estimand provide some evidence that the results are robust to this use of a case-control reduced design.

6. Remarks

Polydesigns combine the full with reduced designs to better focus estimation of selected parts of the parameters to be based on reduced designs of the data. At the same time, polydesigns retain estimability of the target estimand. Therefore, polydesigns offer an alternative inference using a model, and, thus, a way to assess sensitivity of the results to that model through sensitivity to different designs. More study, however, is necessary for the better understanding of the role and good choices of anchor functions.

Although the motivation of polydesigns was through causal inference, these designs, as also noted by a reviewer, can be useful in other settings too. In either case, polydesigns can be especially useful when the estimands of interest have a definition that is in principle over and above a parametric model, as with the potential outcomes in eq. (4.1) and the principal strata in eq. (5.1). In such more general settings, and as indicated by a reviewer, the polydesign has conceptual similarities to Bayesian model averaging, but where the latter combines inferences

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

of different models for one design, whereas the former combines inferences of a model across different designs. Of course, one does not exclude the other, since a larger, "averaging model" could still be used in different ways for different designs. Importantly, the polydesign emphasizes that regardless of the model of analysis, a good design should play a primary role in the quality of the results.

The use of such different designs has been motivated here as retrospective. A more general purpose for using different designs is cost reduction. If we do not use the full design, where we measure all data, but replace the full design with a reduced design, which we would combine with another reduced design, then it is possible to reduce the cost, although, of course, at the expense of precision. The cost of such a polydesign would then be that of the union of the data measured for each of its reduced designs. It is then of interest to develop methods for (a) specifying such prospective polydesigns that can identify the causal effects, and (b) finding members of the polydesigns that have high accuracy per unit cost, while taking into account ethical considerations.

Acknowledgement

We thank the Editor, the Associate Editor, two anonymous reviewers, and Don Rubin for constructive comments, and the National Eye Institute for partial support.

References

Barnard J., Frangakis C.E., Hill J.L., Rubin D.B. (2003). Principal stratification approach to broken randomized experiments: a case study of school choice vouchers in New York City. *Journal of the American Statistical Association* (with discussion) **98** 299-311.

Breslow, N.E. and Day, N.E. (1980). The analysis of case-control data. In *Statistical methods in Cancer Research*, vol 1, World Health Organization.

Cox, D.R., and Oakes, D. (1984). Analysis of survival data. London: Chapman and Hall.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

Drucker, E., Lurie, P., Wodak, A. and Alcabes, P. (1998). Measuring harm reduction: the effects of needle and syringe exchange programs and methadone maintenance on the ecology of HIV. *AIDS* **12**, Suppl A: S217–230.

Frangakis, C.E., Brookmeyer, R.S., Varadhan, R., Safaeian, M., Vlahov, D. and Strathdee, S.A. (2004). Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a needle exchange program. *Journal of the American Statistical Association* **99**, 239-249.

Frangakis, C.E. and Rubin, D.B. (1999). Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika* **86**, 365–379.

Frangakis, C.E., and Rubin, D.B. (2001). Addressing the idiosyncrasy in estimating survival curves using double-sampling in the presence of self-selected right censoring. *Biometrics* (with discussion), **57**, 333–353.

Frangakis, C.E. and Rubin, D.B. (2002). Principal stratification in causal inference. *Biometrics* **58**, 21–29.

Frangakis, C.E., Rubin, D.B., and Zhou, X.H. (2002). Clustered encouragement design with individual noncompliance: Bayesian inference and application to Advance Directive Forms. *Biostatistics* **3**, 147–164.(with discussion).

Gilbert, P. B., Bosch, R. J., and Hudgens, M. G. (2003). Sensitivity analysis for the assessment of causal vaccine effects on viral load in AIDS vaccine trials. *Biometrics* **59**, 531-541.

Keende, J., Stimson, G. V., Jones, S., and Parry-Langdon, N. (1993). Evaluation of syringe-exchange for HIV prevention among injecting drug users in rural and urban areas of Wales. *Addiction* **88**, 1063-1070.

Langholz, B., and Thomas, D. C. (1990). Nested case-control and case-cohort methods of sampling from a cohort: a critical comparison. *American Journal of Epidemiology*, **131**, 169–176.

McClellan, M., McNeil, B.J., Newhouse, J.P. (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *Journal of the*

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

American Medical Association, **272(11)**, 859–866.

Murphy, S.A. and van der Vaart, A.W. (2000). On profile likelihood. *Journal of the American Statistical Association* (with discussion) **95**, 449–485.

Patterson, H.D., and Thompson, R. (1971). Recovery of inter-block information when blocks are unequal. *Biometrika* **58**, 545-554.

Robins, J.M., Rotnitzky, A., and Bonetti, M. (2001). Comment on "Addressing an idiosyncrasy in estimating survival curves using double-sampling in the presence of self-selected right censoring ", by CE Frangakis and DB Rubin. *Biometrics*, **57**, 343-347.

Robins, J. M., Greenland, S., and Hu, F.-C. (1999). "Estimation of the Causal Effect of a Time-Varying Exposure on the Marginal Mean of a Repeated Binary Outcome " (with discussion), *Journal of the American Statistical Association*, **94**, 687-712.

Rubin, D.B. (1974) Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology* **66**, 688–701.

Rubin, D.B. (1978) Bayesian inference for causal effects. *Annals of Statistics* **6**, 34–58.

Rubin, D.B. (1987) The Calculation of Posterior Distributions by Data Augmentation: Comment: A Noniterative Sampling/Importance Resampling Alternative to the Data Augmentation Algorithm for Creating a Few Imputations When Fractions of Missing Information Are Modest: The SIR Algorithm. *Journal of the American Statistical Association*, **82**, 543-546.

Rubin, D. B. and Thomas, N. (1996). Matching Using Estimated Propensity Scores: Relating Theory to Practice. *Biometrics*, **52**, 249–264.

Strathdee, S.A., Celentano, D.D., et.al. (1999). Needle-exchange attendence and health care utilization promote entry into detoxification. *Journal of Urban Health* **76**, 448–460.

van Ameijden, E.J., van den Hoek, J.A., Hartgers, C. and Coutinho, R.A. (1994). Risk factors for the transition from noninjection to injection drug use and accompanying AIDS risk behavior in a cohort of drug users. *American Journal of Epidemiology* **139**, 1153–1163.

Vlahov, D., Anthony, J.C., Munoz, A., et. al. (1991). The ALIVE study. A longitudinal study of HIV infection among injection drug users: Description of methods. NIDA Research Monograph **107**,

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

HostedbyTheBerkeleyElectronicPress

75–100.

APPENDIX: OUTLINE OF PROOF OF PROPERTY 1**(a)**

Denote $A(\theta)$ by A and $\theta - A(\theta)$ by B. Assume both A and B take finitely many values, with true values a_0 and b_0. Denote $\text{pr}_{\text{REDU},n}(A = a_0|D_{\text{REDU},n})$ as $g_n(a_0)$, $\text{pr}_{\text{FULL},n}(A = a_0, B = b_0|D_{\text{FULL},n})$ as $f_n(a_0, b_0)$, where n is the sample size. Assuming identifibility of (A, B) from the full likelihood and that $A(\theta)$ is an anchor function implies that both g_n and f_n converge to a point mass at the true values, in the sense that

$$g_n(a_0) \to 1 \quad \text{and} \quad f_n(a_0, b_0) \to 1, \tag{A.1}$$

as n increases. Then for the polydesign distribution, we have:

$$\begin{aligned}
\text{pr}_{\text{POLY},n}(A = a_0, B = b_0) &= \text{pr}_{\text{REDU},n}(A = a_0|D_{\text{REDU},n})\text{pr}_{\text{FULL},n}(B = b_0|A = a_0, D_{\text{FULL},n}) \\
&= g_n(a_0)f_n(a_0, b_0)/\sum_{b'} f_n(a_0, b') \\
&= g_n(a_0)f_n(a_0, b_0)/(f_n(a_0, b_0) + \sum_{b' \neq b_0} f_n(a_0, b')).
\end{aligned}$$

In the last line, all terms except $\sum_{b' \neq b_0} f_n(a_0, b')$ converge to 1, by (A.1). Moreover, for $b' \neq b_0$, $f_n(a_0, b')$ converges to 0. So, the expression in the last line converges to 1, and so the median of the polydesign distribution converges in probability to the true value.

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80

Table 1: Comparison of full design and polydesign estimation in the example of Section 4.

varied parameters		marginal odds ratio	median for estimating Q		SE for estimating $\log Q$		$\{\text{MSE}\}^{\frac{1}{2}}$ for estimating $\log Q$		95% CI coverage for estimating Q or $\log Q$	
pr$(Y(0) = 1 \mid X)$	$\exp(\beta_d)$	Q (estimand)	Full	Poly	Full	Poly	Full	Poly	Full	Poly
$(0.01, 0.08, 0.40)$	0.50	0.59	0.60	0.59	0.49	0.52	0.53	0.49	0.98	0.94
(logistic linear)	1.00	1.00	1.00	1.00	0.28	0.37	0.28	0.36	0.95	0.95
	2.00	1.57	1.59	1.57	0.22	0.30	0.22	0.30	0. 95	0.95
$(0.01, 0.40, 0.08)$	0.50	0.59	1.60	0.53	0.34	0.76	1.03	0.78	0.17	0.98
(not logistic linear)	1.00	1.00	2.69	1.00	0.27	0.38	1.01	0.39	0.06	0.95
	2.00	1.57	3.96	1.79	0.23	0.35	0.95	0.37	0.03	0.93

HostedbyTheBerkeleyElectronicPress

Table 2: Summary statistics of the posterior marginal distribution of parameters of interest, based on different designs.

Estimand	Design	Median	2.5% quantile	97.5% quantile
(a) Odds ratio of getting HIV when	Full	0.09	0.00	0.86
exchanging versus when not exchanging	Redu	0.09	0.00	0.85
needles, within a principal stratum.	Poly	0.09	0.00	0.79
(b) Odds ratio of getting HIV under fixed	Full	3.20	0.97	16.9
exchange, for higher versus lower principal	Redu	3.27	0.97	19.9
strata, given fixed baseline risk score.	Poly	3.17	0.97	16.7
(c) Odds ratio of higher versus lower	Full	1.15	1.12	1.18
principal strata for one s.d. increase	Redu	1.15	1.12	1.19
in baseline risk score.	Poly	1.15	1.12	1.18

COBRA
A BEPRESS REPOSITORY
Collection of Biostatistics
Research Archive

http://biostats.bepress.com/jhubiostat/paper80