Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Stat Probab Lett. Author manuscript; available in PMC 2016 October 01.

Published in final edited form as:

Stat Probab Lett. 2015 October 1; 105: 130–135. doi:10.1016/j.spl.2015.06.011.

Exact Confidence Intervals in the Presence of Interference

Joseph Rigdon and

Quantitative Sciences Unit, Stanford University, Palo Alto, California 94304, U.S.A

Michael G. Hudgens

Department of Biostatistics, CB 7420, University of North Carolina, Chapel Hill, North Carolina 27516, U.S.A

Joseph Rigdon: jrigdon@stanford.edu; Michael G. Hudgens: mhudgens@bios.unc.edu

Abstract

For two-stage randomized experiments assuming partial interference, exact confidence intervals are proposed for treatment effects on a binary outcome. Empirical studies demonstrate the new intervals have narrower width than previously proposed exact intervals based on the Hoeffding inequality.

Keywords

Causal Inference; Exact Confidence Interval; Interference; Randomization Inference

1. Introduction

In a randomized experiment, it is commonly assumed that an individual only has two potential outcomes: an outcome on control, and an outcome on treatment. That an individual has only two potential outcomes assumes no interference (Cox, 1958) between individuals, i.e., an individual's potential outcomes are unaffected by the treatment assignment of any other individual in the study. There are many settings where this assumption of no interference is clearly violated (Hong and Raudenbush, 2006; Sobel, 2006; Rosenbaum, 2007).

Partial interference holds when individuals can be partitioned into groups such that there is no interference between individuals in different groups. In settings where partial interference holds, two-stage randomized experiments have been suggested as a study design for drawing inference about treatment (i.e., causal) effects. Two-stage randomized experiments proceed by *(i)* randomizing groups to treatment strategies and *(ii)* randomizing individuals within groups to different treatments based on the treatment strategy assigned to their group in stage *(i)*. Two-stage randomized experiments are found in many fields of study, e.g., infectious diseases (Baird et al., 2012), medicine (Borm et al., 2005), economics (Duflo and

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Saez, 2003), and political science (Ichino and Schündeln, 2012; Sinclair et al., 2012). Building upon ideas in Halloran et al. (1991), Hudgens and Halloran (2008) defined and derived unbiased estimators for the direct, indirect, total, and overall effects of treatment in a two-stage randomized experiment assuming partial interference. Liu and Hudgens (2014) showed that Wald-type confidence intervals based on these estimators perform well when the number of groups is large; however, often the number of groups may not be large enough. For example, Moulton et al. (2001) describe a group-randomized vaccine trial involving approximately 9,000 individuals but only 38 groups. Tchetgen Tchetgen and VanderWeele (2012), henceforth TV, proposed exact confidence intervals using the Hoeffding inequality for these four effects in a two-stage randomized experiment with partial interference. Unfortunately, as will be shown below, the TV intervals can be very wide and conservative.

In this paper, we propose different exact confidence intervals based on inverting exact hypothesis tests that tend to be less conservative than TV. The remainder of the paper is organized as follows. In §2, treatment effects in the presence of interference are defined and existing inferential results are reviewed. In §3, the assumption of stratified interference is presented and bounds are derived for the causal effects under this assumption. In §4 the proposed new exact confidence intervals are described by inverting certain permutation tests. In §5 a simulation study is conducted comparing the TV, asymptotic, and new exact confidence intervals. §6 concludes with a discussion. An R package is available implementing the proposed confidence intervals.

2. Preliminaries

2.1. Estimands

Consider a finite population of N individuals partitioned into k groups with n_i individuals in group i for $i = 1, \ldots, k$. Assume partial interference, i.e., there is no interference between individuals in different groups. Consider a two-stage randomized experiment wherein h of k groups are assigned to strategy a_1 and k–h are assigned to a_0 in the first stage, where strategy a_s specifies that m_i^s of n_i individuals will receive treatment. For example, strategy a_0 might entail assigning (approximately) 1/3 of individuals within a group to treatment whereas strategy a_1 might entail assigning (approximately) 2/3 of individuals within a group to treatment (see TV for further discussion about different types of treatment allocation strategies). Let $S_i = 1$ if group i is randomized to a_1 and 0 otherwise so that $\Pr[S_i = 1] = h/k$. In the second stage, individuals will be randomized to treatment conditional on group assignment in the first stage. Let $Z_{ij} = 1$ if individual j in group i is assigned treatment and 0 otherwise. Let $Z_i = (Z_{i1}, \ldots, Z_{in_i})$ be the random vector of treatment assignments for group i taking on values $z_i \in \mathscr{R}(n_i, m_i^s)$, the set of all vectors of length n_i composed of m_i^s elements equal to 1 and $n_i - m_i^s$ elements equal to 0. Additionally, let $Z_{i(j)}$ denote the random vector of treatment assignments in group i excluding individual j taking on values $z_{i(j)} \in \mathscr{R}(n_i - 1, m_i^s - z_{ij})$.

Let $y_{ij}(z_i)$ be the binary potential outcome for individual j in group i when group i receives treatment vector z_i. A randomization inference framework is adopted wherein potential

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

outcomes are fixed features of the finite population of N individuals and only treatment assignments S and Z are random (as in Sobel (2006); Rosenbaum (2007); Hudgens and Halloran (2008)). Define the average potential outcome for individual j in group i on treatment $z = 0, 1$ under strategy a_s as

$$\overline{y}_{ij}(z;\alpha_s)=\sum_{\omega\in\mathscr{R}(n_i-1,m_i^s-z)} y_{ij}(z_{ij}=z,z_{i(j)}=\omega)\Pr(Z_{(i)j}=\omega|Z_{ij}=z;S_i=s) \quad (1)$$

where $\Pr(Z_{i(j)}=\omega|Z_{ij}=z;S_i=s)=\begin{pmatrix} n_i-1 \\ m_i^s-z \end{pmatrix}^{-1}$. Henceforth, let $\sum_i=\sum_{i=1}^k$ and $\sum_j=\sum_{j=1}^{n_i}$. For treatment z under strategy a_s define the group average potential outcome as $\overline{y}_i(z;\alpha_s) \equiv n_i^{-1}\sum_j \overline{y}_{ij}(z;\alpha_s)$, and the population average potential outcome as $\bar{y}(z;\ a_s) \equiv k^{-1}\ \Sigma_i\ \bar{y}_i(z;\ a_s)$. Define the average potential outcome for individual j in group i under strategy a_s as

$$\overline{y}_{ij}(\alpha_s) \equiv \sum_{\omega\in\mathscr{R}(n_i,m_i^s)} y_{ij}(z_i=\omega)\Pr(Z_i=\omega;S_i=s), \quad (2)$$

the group average potential outcome as $\overline{y}_i(\alpha_s) \equiv n_i^{-1}\sum_j \overline{y}_{ij}(\alpha_s)$, and the population average potential outcome as $\overline{y}(\alpha_s) \equiv k_i^{-1}\sum_j \overline{y}_i(\alpha_s)$. Define the direct effect of treatment for strategy a_s as $DE(a_s) = \bar{y}(0;\ a_s) - \bar{y}(1;\ a_s)$, the indirect effect of a_0 versus a_1 as $IE(a_0,\ a_1) = \bar{y}(0;\ a_0) - \bar{y}(0;\ a_1)$, the total effect as $TE(a_0,\ a_1) = \bar{y}(0;\ a_0) - \bar{y}(1;\ a_1)$, and the overall effect of a_0 versus a_1 as $OE(a_0,\ a_1) = \bar{y}(a_0) - \bar{y}(a_1)$; see Hudgens and Halloran (2008) and TV for additional discussion regarding these effects.

2.2. Existing Inferential Results

Hudgens and Halloran (2008) derived unbiased estimators for all population average potential outcomes, and thus for the four causal effects. Noting that $\Pr[S_i = s]$ and $\Pr[Z_{ij} = z|\ S_i = s]$ are known by design, the estimator

$$\hat{y}(z;\alpha_s)=k^{-1}\sum_i \frac{1\{S_i=s\}\hat{y}_i(z;\alpha_s)}{\Pr[S_i=s]} \quad (3)$$

where $\overline{y}_i(z;\alpha_s)=n_i^{-1}\sum_j 1\{Z_{ij}=z\}y_{ij}(Z_{ij})/\Pr[Z_{ij}=z|S_i=s]$ is unbiased for $\bar{y}(z;\ a_s)$. Additionally, the estimator

$$\hat{y}(\alpha_s)=k^{-1}\sum_i \frac{1(S_i=s)n_i^{-1}\sum_j y_{ij}(Z_{ij})}{\Pr[S_i=s]} \quad (4)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

is unbiased for $\bar{y}(a_s)$. Unbiased estimators for the effects of interest follow immediately: $\hat{DE}(\alpha_s)=\hat{y}(0;\alpha_s)-\hat{y}(1;\alpha_s), \hat{IE}(\alpha_0,\alpha_1)=\hat{y}(0;\alpha_0)-\hat{y}(0;\alpha_1), \hat{TE}(\alpha_0,\alpha_1)=\hat{y}(0;\alpha_0)-\hat{y}(1;\alpha_1)$, and $\hat{OE}(\alpha_0,\alpha_1)=\hat{y}(\alpha_0)-\hat{y}(\alpha_1)$.

TV proposed exact confidence intervals based on the Hoeffding inequality for the effects of interest in a two-stage randomized experiment where partial interference is assumed. In particular, for any $\gamma\in\{0, 1\}$, $\hat{DE}(\alpha_s)\pm\varepsilon^*_D(\gamma,\alpha_s,q_s,k)$ is a $1-\gamma$ exact confidence interval for $DE(a_s)$ where $\varepsilon^*_D(\gamma,\alpha_s,q_s,k)$ is given in equation (17) of TV for $s = 0, 1$. Additionally, $\hat{IE}(\alpha_0,\alpha_1)\pm\varepsilon^*(\gamma,\alpha_0,q_0,\alpha_1,q_1,k), \hat{TE}(\alpha_0,\alpha_1)\pm\varepsilon^*(\gamma,\alpha_0,q_0,\alpha_1,q_1,k)$, and $\hat{OE}(\alpha_0,\alpha_1)\pm\varepsilon^*(\gamma,\alpha_0,q_0,\alpha_1,q_1,k)$ are all $1-\gamma$ exact confidence intervals for their target parameters where $\varepsilon^*(\gamma, a_0, q_0, a_1, q_1, k)$ is given in Theorem 3 of TV.

Liu and Hudgens (2014) examined conditions under which Wald-type intervals $\hat{DE}(\alpha_s)\pm z_{(1-\gamma/2)}\{\hat{var}(\hat{DE}(\alpha_s))\}^{1/2}$ and Chebyshev-type intervals $\hat{DE}(\alpha_s)\pm\{\hat{var}(\hat{DE}(\alpha_s))/\gamma\}^{1/2}$ are valid, large sample confidence intervals for $DE(a_s)$, where $z_{(1-\gamma/2)}$ is the $1-\gamma/2$ quantile for the standard normal distribution and $\hat{var}(\hat{DE}(\alpha_s))$ is an estimator of the variance of $\hat{DE}(\alpha_s)$ for $s = 0, 1$. They also considered Wald and Chebyshev-type confidence intervals for the indirect, total, and overall effects.

3. Bounds Under Stratified Interference

Exact randomization based inference about the four effects is challenging without further assumptions as the experiment reveals only N of the $\sum_i\sum_j\{\binom{n_i}{m_i^0}+\binom{n_i}{m_i^1}\}$ total potential outcomes. One such additional assumption is stratified interference (Hudgens and Halloran, 2008), which assumes that individual j in group i has the same potential outcome when assigned control or treatment as long as a fixed number of other individuals in group i are assigned treatment, i.e.,

$$y_{ij}(z_i)=y_{ij}(z_i') \text{ for all } z_i, z_i'\in\mathscr{R}(n_i,m_i^s) \text{ such that } z_{ij}=z_{ij}'. \quad (5)$$

Under (5), individual j in group i only has four potential outcomes, which we denote by $y_{ij}(z; a_s)$ for $z, s = 0, 1$, so that the experiment reveals the observed outcome $Y_{ij} = \Sigma_{z,s=0,1}\, 1\{Z_{ij} = z; S_i = s\}y_{ij}(z; a_s)$ for each individual and thus N of the $4N$ total potential outcomes. Furthermore, (5) implies that $\bar{y}_{ij}(z; a_s) = y_{ij}(z; a_s)$, and that $\bar{y}_{ij}(\alpha_s)=w_i^s y_{ij}(1;\alpha_s)+(1-w_i^s)y_{ij}(0;\alpha_s)\equiv y_{ij}(\alpha_s)$ where $w_i^s=\Pr[Z_{ij}=1|S_i=s]=m_i^s/n_i$.

Under (5), the observed data form bounded sets for all effects contained in the interval $[-1, 1]$. The bounded sets have widths less than two where here and in the sequel the width of a set is defined to be the difference between its maximum and minimum values. Consider $DE(\alpha_0)=k^{-1}\sum_i n_i^{-1}\sum_j\{y_{ij}(0;\alpha_0)-y_{ij}(1;\alpha_0)\}$ for illustration. For the $\Sigma_i\,\Sigma_j(1 - S_i)(1 - Z_{ij})$ individuals with $S_i = Z_{ij} = 0$, $y_{ij}(0; a_0)$ is revealed; however, for the $N - \Sigma_i\,\Sigma_j(1 - S_i)(1 -$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Z_{ij}) individuals with $S_i = 1$ or $Z_{ij} = 1$, $y_{ij}(0; a_0)$ is missing and only known to be 0 or 1. Let $\vec{y}(z; a_s)$ be the N-dimensional vector of potential outcomes for treatment z under strategy a_s. Under (5), a lower bound for $DE(a_0)$ is found by filling in all missing potential outcomes in $y(\vec{0}; a_0)$ as 0 and all missing potential outcomes in $y(\vec{1}; a_0)$ as 1. An upper bound for $DE(a_0)$ is found by filling in all missing potential outcomes in $y(\vec{0}; a_0)$ as 1 and all missing potential outcomes in $y(\vec{1}; a_0)$ as 0. Simple algebra shows that width of the bounded set for $DE(a_0)$ is equal to $2 - (k - h)/k$. The width of this bounded set approaches 1 as $(k - h)/k \to 1$, i.e., as more groups are randomized to a_0.

Similar logic leads to bounds for the other effects. The width of the bounded set for $DE(a_1)$ is equal to $2 - h/k$ which approaches 1 as $h/k \to 1$. The width of the bounded set for $IE(a_0, a_1)$ is equal to $2-k^{-1}\sum_i n_i^{-1}\sum_j (1-Z_{ij})$ which approaches 1 as the proportion of individuals assigned $Z_{ij} = 0$ approaches 1. The width of the bounded set for $TE(a_0, a_1)$ is equal to $2-k^{-1}\sum_i n_i^{-1}\{(1-S_i)\sum_j (1-Z_{ij})+S_i\sum_j Z_{ij}\}$ which approaches 1 as the proportion of individuals with $S_i = Z_{ij} = 0$ or $S_i = Z_{ij} = 1$ approaches 1. Lower and upper bounds for $OE(a_0, a_1)$ can be derived similarly but the corresponding width does not have a simple closed form.

4. EIT Confidence Intervals

In addition to leading to unbiased estimators and bounds, the observed data can be used to form $1 - \gamma$ confidence sets for the four effects. The confidence sets are formed by inverting hypothesis tests about the potential outcomes that define the effect of interest. This section is divided into two parts: §4.1 outlines how the confidence sets are formed and §4.2 presents a computationally feasible algorithm for constructing an interval that contains the exact confidence set. Henceforth this interval is referred to as the exact inverted test (EIT).

4.1. An Exact Confidence Set

The methods to follow can be generalized to any effect, so consider $DE(a_0)$. Inference about $DE(a_0)$ concerns the vectors $y(\vec{0}; a_0)$ and $y(\vec{1}; a_0)$, which are partially revealed by the experiment. A hypothesis about these vectors is considered sharp if it completely fills in the potential outcomes not revealed by the experiment. A sharp null $H_0 : y(\vec{0}; a_0) = y^0(\vec{0}; a_0), \vec{y}(1; a_0) = y^0(1; a_0)$ maps to a value of $DE(a_0)$, which we denote $DE^0(a_0)$. Only sharp null hypotheses that are compatible with the observed data need to be tested as other sharp nulls can be rejected with zero probability of making a type I error. Thus for each sharp null to be tested, the implied null value $DE^0(a_0)$ will be a member of the bounded set derived in §3. There are $B_1 = 2^{\sum_i (1-S_i) n_i} 4^{\sum_i S_i n_i}$ sharp null hypotheses to test, as individuals with $S_i = 0$ have only one missing potential outcome with two possible values $\{0, 1\}$, and individuals with $S_i = 1$ have two missing potential outcomes with four possible values $\{0, 1\} \times \{0, 1\}$.

After filling in the missing potential outcomes under H_0, the null distribution of the test statistic $\hat{DE}(\alpha_0)$ can be found by computing the statistic, denoted by $\hat{DE}_c(\alpha_0)$, for each of the $c = 1, \ldots, C_1$ possible experiments under H_0, where

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$C_1=\sum_{S\in\mathcal{S}}\Pi_{i=1}^{k}\begin{pmatrix} n_i \\ m_i^1 \end{pmatrix}^{S_i}\begin{pmatrix} n_i \\ m_i^0 \end{pmatrix}^{(1-S_i)}$ and $\mathcal{S}$ is the set of all possible values of the vector S such that $|\mathcal{S}|=\begin{pmatrix} k \\ h \end{pmatrix}$. A two-sided p-value to test H_0 is given by $p_0=\sum_{c=1}^{C_1}1\{|\hat{DE}_c(\alpha_0)-DE^0(\alpha_0)|\ \geq\ |\hat{DE}(\alpha_0)-DE^0(\alpha_0)|\}/C_1$. If $p_0 < \gamma$, H_0 is rejected. Note p_0 is a function of the null hypothesis vectors $y^0(0;\ \alpha_0)$ and $y^0(1;\ \alpha_0)$. Let $p(DE^0(\alpha_0))$ denote the set of all p_0 which are functions of compatible vectors $y^0(0;\ \alpha_0)$ and $y^0(1;\ \alpha_0)$ that map to $DE^0(\alpha_0)$. A $1-\gamma$ confidence set for $DE(\alpha_0)$ is $\{DE^0(\alpha_0) : \max\{p(DE^0(\alpha_0))\} \geq \gamma\}$. P-values, and thus confidence sets, can be found in an analogous manner for the other effects.

4.2. A Computationally Feasible Algorithm

Finding the exact confidence set for $DE(\alpha_0)$ described above entails testing B_1 hypotheses, where each hypothesis test involves C_1 randomizations. As N becomes large, the computational time necessary to perform $B_1\times C_1$ operations grows exponentially. For illustration of the problem, consider two examples in which $h = 1$ of $k = 1$ groups are assigned α_0, in which $m_1^0=10$ of $n_1 = 20$ individuals are randomized to treatment such that $B_1 = 2^{20}$ and $C_1=\begin{pmatrix} 20 \\ 10 \end{pmatrix}=184,756$. Suppose there are two cases of observed data: (a) 5 of 10 unexposed experienced an event, and 5 of 10 exposed experienced an event, and (b) 8 of 10 unexposed experienced an event and 2 of 10 exposed experienced an event. Figure 1 displays a plot of $DE^0(\alpha_0)$ versus $p(DE^0(\alpha_0))$ for both examples. The bounded set and 95% exact confidence set for $DE^0(\alpha_0)$ are, respectively, {−0.5, −0.45, …, 0.45, 0.5} and {−0.35, −0.3, …, 0.3, 0.35} in (a) and {−0.2, −0.15, …, 0.75, 0.8} and {0.15, 0.2, …, 0.75, 0.8} in (b).

A computationally feasible algorithm is given below for approximating the confidence sets. The algorithm entails testing a targeted random sample of B_2 of the B_1 total sharp null hypotheses, and computing p-values for each sampled sharp null based on a random sample of C_2 of the C_1 possible randomizations. The set of computed p-values are then used to approximate the confidence set endpoints using local linear interpolation. For intuition underlying the interpolation step, consider the piecewise linear function that connects the maximum p-values for each compatible value of $DE(\alpha_0)$ in Figure 1. Finding the x-coordinates for the intersection points of this function and a horizontal line at γ will conservatively approximate the lower and upper $1-\gamma$ confidence limits for $DE(\alpha_0)$. This suggests the following targeted, local linear interpolation algorithm for estimating the lower bound of a confidence set for $DE(\alpha_0)$. An analogous algorithm can be used to target the upper limit of the confidence set for $DE(\alpha_0)$.

Let $\hat{DE}(\alpha_0)_l$ denote the lower bound for $DE(\alpha_0)$, and $\hat{y}(z;\ \alpha_0)_l$ and $\hat{y}(z;\ \alpha_0)_u$ denote the lower and upper bounds, respectively, for $\bar{y}(z;\ \alpha_0)$.

1. Test the unique sharp null about $y(\vec{0};\ \alpha_0)$ and $y(\vec{1};\ \alpha_0)$ that maps to $\hat{DE}(\alpha_0)_l$. If the corresponding p-value $p_0 \geq \gamma$, let $\hat{DE}(\alpha_0)_l$ be the lower limit of the confidence set

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and do not proceed. Otherwise, let $l = \hat{DE}(\alpha_0)$ and let $p_l = 1 - 1/B_2$. Let $\mathcal{L} = \{\hat{DE}(\alpha_0)_l\}$ and $\mathcal{P} = \{p_0\}$.

2. Fill in the missingness in $y(\vec{0};\ \alpha_0)$ with samples from a Bernoulli distribution with mean $f(\{\hat{y}(0;\alpha_0)_l + \hat{y}(1;\alpha_0)_u + q_{p_l}(\hat{DE}(\alpha_0)_l, l)\}/2)$ and fill in the missingness in $\vec{y}(1;\ \alpha_0)$ with samples from a Bernoulli distribution with mean $f(\{\hat{y}(0;\alpha_0)_l + \hat{y}(1;\alpha_0)_u - q_{pl}(\hat{DE}(\alpha_0)_l, l)\}/2)$ where $q_p(a, b) = (1 - p)a + pb$, and $f(x) = x$ if $0 \le x \le 1$, $f(x) = 0$ if $x < 0$, and $f(x) = 1$ if $x > 1$.

3. If the sampled sharp null maps to a value $DE^0(\alpha_0) \in [\hat{DE}(\alpha_0)_l, l]$, add $DE^0(\alpha_0)$ to the set $\mathcal{L}$, add the corresponding p_0 to $\mathcal{P}$, and if $p_0 \ge \gamma$ then update l to equal $DE^0(\alpha_0)$. Otherwise, do not compute a p-value corresponding to the sampled sharp null and let $p_l = p_l - 1/B_2$.

   Repeat Steps 2 and 3 $B_2/2 - 1$ times.

Let t be the function from $\mathcal{P}$ to $\mathcal{L}$ that maps each p-value p_0 in $\mathcal{P}$ to the null value of $DE^0(\alpha_0)$ in $\mathcal{L}$ which corresponds to the sharp null hypothesis which generated p_0. Let $\mathcal{R} = \{\max\{p \in \mathcal{P} : t(p) = l\} : l \in \mathcal{L}\}$. Let $r_1 = \min\{r \in \mathcal{R} : r \ge \gamma\}$ and let $r_2 = \max\{r \in \mathcal{R} : r < \gamma\}$. Let $l_i = t(r_i)$ for $i = 1, 2$. The lower limit of the confidence set l^* is found by local linear interpolation by finding the x-coordinate for the point at which a line drawn from (l_2, r_2) to (l_1, r_1) intersects a horizontal line at γ, i.e., $l^* = l_2 + (\gamma - r_2)(l_2 - l_1)/(r_2 - r_1)$. The upper limit u^* is found analogously. As $B_2 \to B_1$ and $C_2 \to C_1$, the interval $[l^*, u^*]$ will contain the exact confidence set described in §4.1 with probability approaching 1.

The algorithms for approximating confidence sets for $IE(\alpha_0, \alpha_1)$ and $TE(\alpha_0, \alpha_1)$ are analogous. For $OE(\alpha_0, \alpha_1)$ the algorithm is modified slightly as it involves all four vectors $\vec{y}(z;\ \alpha_s)$, $z, s = 0, 1$. Let $\hat{y}(\alpha_s)_l$ and $y(\vec{\alpha}_s)_u$ be the lower and upper limits, respectively, for $\bar{y}(\alpha_s)$ under (5). If $p_0 < \gamma$ for $OE(\alpha_0, \alpha_1)_l$, set $l = \hat{OE}(\alpha_0, \alpha_1)$ and fill in the missingness in $\vec{y}(0;\ \alpha_0)$ and $y(\vec{1};\ \alpha_0)$ with samples from a Bernoulli distribution with mean $f(\{\hat{y}(\alpha_0)_l + \hat{y}(\alpha_1)_u + q_{p_l}(\hat{OE}(\alpha_0, \alpha_1)_l, l)\}/2)$ where $p_l = 1 - 1/B_2$. A p-value is computed if $OE^0(\alpha_0, \alpha_1) \in [\hat{OE}(\alpha_0, \alpha_1)_l, l]$ and if not $p_l = p_l - 1/B_2$. If $p_0 \ge \gamma$ for $OE(\alpha_0, \alpha_1)_l$, l is set to equal $OE^0(\alpha_0, \alpha_1)$. The upper endpoint can be approximated using an analogous approach.

The R package interferenceCI is available on CRAN (Rigdon, 2015) for computing EIT confidence intervals via this algorithm for the four effects assuming stratified interference when the outcome is binary. The Wald, Chebyshev, and TV intervals are also computed in the package.

5. Comparisons Via Simulation

A simulation study was carried out to compare the asymptotic, TV, and EIT confidence intervals. The simulation proceeded as follows for fixed values of α_0, α_1, $DE(\alpha_0)$, $DE(\alpha_1)$, $IE(\alpha_0, \alpha_1)$, k, $n_i = n$ for $i = 1, \ldots, k$ such that $N = kn$:

0 Potential outcomes were generated by first fixing the vectors $\vec{y}(\vec{z};\ \alpha_s)$ for $z, s = 0, 1$ to be length N vectors of all 0s. Group membership was assigned by letting

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

elements $n(i-1)+1, \ldots, ni$ of each vector belong to group $i = 1, \ldots, k$. Then, $N(0.5+DE(a_0)/2)$ elements in $y(\vec{0}; a_0)$ were randomly set to equal 1 and $N(0.5-DE(a_0)/2)$ elements in $y(\vec{1}; a_0)$ were randomly set to equal 1. Then, $N(0.5 + DE(a_0)/2 - IE(a_0, a_1))$ elements in $y(\vec{0}; a_1)$ were randomly set to equal 1. Finally, $N(0.5 + DE(a_0)/2 - IE(a_0, a_1) - DE(a_1))$ elements in $y(\vec{1}; a_1)$ were randomly set to equal 1.

1 Observed data were generated by *(i)* randomly assigning h of k groups to strategy a_1 and *(ii)* randomly assigning $m_i^s = \alpha_s n$ of n individuals per group to treatment for $s = 0, 1$. Observed outcomes followed based on these treatment assignments and the potential outcomes from step 0.

2 For each effect, 95% confidence intervals were computed using the observed data generated in step 1.

3 Steps 1–2 were repeated 1000 times.

In the simulation we let $k = n = 10$ or $k = n = 20$ with $h = k/2$, $m_i^0 = 0.3n$ under a_0, $m_i^1 = 0.6n$ under a_1, $DE(a_0) = 0.95$, $DE(a_1) = 0.3$, and $IE(a_0, a_1) = 0.5$ (such that $TE(a_0, a_1) = 0.8$ and $OE(a_0, a_1) = 0.395$). In the targeted sampling algorithm, $B_2 = C_2 = 100$ such that B_2/B_1 and C_2/C_1 were less than 10^{-20} for all effects. Table 1 displays average widths and coverages for Wald, EIT, Chebyshev, and TV. Wald and Chebyshev fail to achieve nominal coverage for $DE(a_0)$ when $k = n = 10$ and Wald additionally fails to cover for $DE(a_0)$ when $k = n = 20$ and for $IE(a_0, a_1)$ and $TE(a_0, a_1)$ when $k = n = 10$. As guaranteed by their respective constructions, EIT and TV achieve nominal coverage for all setups; however, EIT has narrower width than TV in all setups. In fact, EIT is an order of magnitude narrower than TV in three instances: $DE(a_0)$, $TE(a_0, a_1)$, and $OE(a_0, a_1)$ when $k = n = 20$.

6. Discussion

In this paper new exact confidence intervals have been proposed for causal effects in the presence of partial interference. The new intervals are constructed by inverting permutation based hypothesis tests. These intervals do not rely on any parametric assumptions and require no assumptions about random sampling from a larger population. The confidence intervals are exact in the sense that the probability of containing the true treatment effects is at least the nominal level. As there may be many vectors of potential outcomes that map to one value of the causal estimand, a computationally feasible algorithm was proposed in §4.2 to approximate the exact confidence intervals. Empirical studies demonstrate the new exact intervals have narrower width than previously proposed exact intervals based on the Hoeffding inequality. Nonetheless, the empirical coverage of the proposed intervals still tends to exceed the nominal level, suggesting one possible future avenue of research would be to develop alternative intervals which are less conservative and narrower but maintain nominal coverage.

Acknowledgments

This research was supported by the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the the NIH. The authors thank Mark Weaver for helpful comments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

References

Baird S, Garfein R, McIntosh C, Özler B. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. The Lancet. 2012; 379:1320–1329.

Borm G, Melis R, Teerenstra S, Peer P. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. Statistics in Medicine. 2005; 24:3535–3547. [PubMed: 16007575]

Cox, D. Planning of Experiments. Wiley; New York, NY: 1958.

Duflo E, Saez E. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. The Quarterly Journal of Economics. 2003; 118:815–842.

Halloran M, Haber M, Longini I, Struchiner C. Direct and indirect effects in vaccine efficacy and effectiveness. American Journal of Epidemiology. 1991; 133:323–331. [PubMed: 1899778]

Hong G, Raudenbush S. Evaluating kindergarten retention policy: A case study of causal inference for multi-level observational data. Journal of the American Statistical Association. 2006; 101:901–910.

Hudgens M, Halloran M. Toward causal inference with interference. Journal of the American Statistical Association. 2008; 103:832–842. [PubMed: 19081744]

Ichino N, Schündeln M. Deterring or displacing electoral irregularities? Spillover effects of observers in a randomized field experiment in ghana. The Journal of Politics. 2012; 74:292–307.

Liu L, Hudgens M. Large sample randomization inference of causal effects in the presence of interference. Journal of the American Statistical Association. 2014; 109:288–301. [PubMed: 24659836]

Moulton L, O'Brien K, Kohberger R, Chang I, Reid R, Weatherholtz R, Hackell J, Siber G, Santosham M. Design of a group-randomized Streptococcus pneumoniae vaccine trial. Controlled Clinical Trials. 2001; 22:438–452. [PubMed: 11514043]

Rigdon, J. interferenceCI: Exact Confidence Intervals in the Presence of Interference. R package version 1.1. 2015. http://CRAN.R-project.org/package=interferenceCI

Rosenbaum P. Interference between unites in randomized experiments. Journal of the Americal Statistical Association. 2007; 102:191–200.

Sinclair B, McConnell M, Green D. Detecting spillover effects: Design and analysis of multilevel experiments. American Journal of Political Science. 2012; 56:1055–1069.

Sobel M. What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference. Journal of the Americal Statistical Association. 2006; 101:1398–1407.

Tchetgen Tchetgen E, VanderWeele T. On causal inference in the presence of interference. Statistical Methods in Medical Research. 2012; 21:55–75. [PubMed: 21068053]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

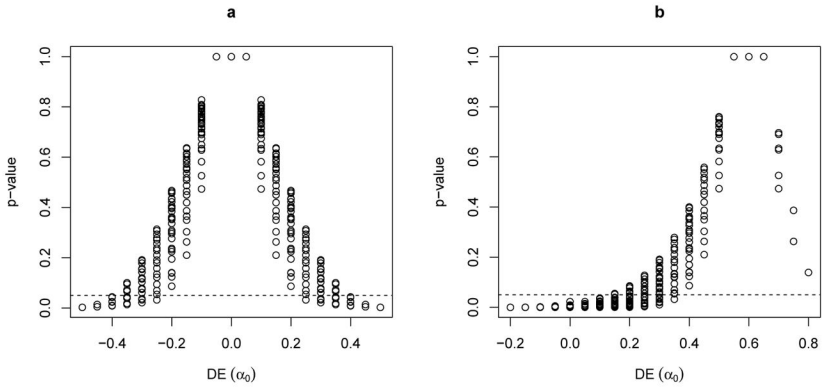

Figure 1.
Plot of $DE(\alpha_0)$ versus $p(DE(\alpha_0))$ for examples (a) and (b) as outlined in §4.2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Empirical width [coverage] of Wald (W), exact inverted test (EIT), Chebyshev (C), and TV 95% CIs for simulation study discussed in §5.

	n	k	$DE(a_0)$	$DE(a_1)$	$IE(a_0, a_1)$	$TE(a_0, a_1)$	$OE(a_0, a_1)$
W	10	10	0.13 [0.84]	0.51 [0.96]	0.39 [0.93]	0.30 [0.93]	0.24 [0.94]
	20	20	0.09 [0.89]	0.26 [0.96]	0.21 [0.95]	0.14 [0.98]	0.11 [0.97]
EIT	10	10	0.28 [0.98]	0.52 [0.98]	0.47 [0.99]	0.31 [0.98]	0.36 [1.00]
	20	20	0.12 [0.98]	0.27 [0.98]	0.24 [0.98]	0.14 [0.98]	0.18 [1.00]
C	10	10	0.22 [0.84]	1.15 [1.00]	0.84 [1.00]	0.54 [1.00]	0.54 [1.00]
	20	20	0.15 [0.99]	0.59 [1.00]	0.49 [1.00]	0.32 [1.00]	0.26 [1.00]
TV	10	10	1.95 [1.00]	2.00 [1.00]	2.00 [1.00]	2.00 [1.00]	2.00 [1.00]
	20	20	1.41 [1.00]	2.00 [1.00]	1.86 [1.00]	1.56 [1.00]	1.96 [1.00]

Stat Probab Lett. Author manuscript; available in PMC 2016 October 01.