5-20-1997

Pooling Community Data for Community Interventions When the Number of Pairs is Small

Paula Diehr
University of Washington, pdiehr@u.washington.edu

Ted Lystig
Chalmers University of Technology, Sweden, lystig@math.chalmers.se

Holly Andrilla
University of Washington, hollya@u.washington.edu

Ziding Feng
University of Washington & Fred Hutchinson Cancer Research Center, zfeng@fhcrc.org

Introduction

A recent trend in health promotion is the development of interventions that target an entire community, such as community-based health promotion. Controlled evaluations of such interventions typically compare a group of intervention communities with a group of control communities in terms of reducing the prevalence of an unhealthy behavior, such as smoking. The number of communities has usually been small, for budgetary reasons: 6 for the Minnesota Heart Health Program (3 treatment and 3 control)[i], just 2 for North Karelia Project [ii] and the Pawtucket Heart Health Program [iii], and 3 and 5 for two studies at Stanford [iv,v]. Community intervention studies are also described by Shea and Basch[vi].

An analysis of variance model for a community intervention with 3 communities per treatment group is shown in Table 1, where Y, the dependent variable, is an individual level measure of change. Because the communities were randomized to treatment group, the proper F test is F_c, which has 1 and 4 degrees of freedom and requires a critical value of 7.71. If the less correct person-level test had been performed, F_p would have 1 and 598 degrees of freedom, and require a critical value of 3.84. If it were known that σ^2_b was zero, F_p would have much higher power. A test of significance for whether σ^2_b is zero if F_v. Note that the procedure effectively uses F_c when it is least favorable, and uses F_p when there is no variation and so F_c would tend to perform well. This perverse sort of cheating will keep the method from performing well.

[Table 1 about here]

There is some statistical literature describing the use of preliminary tests of significance for pooling mean squares in the analysis of variance.[vii,viii] A major issue is what the α level of significance should be for the initial test of significance. Bozivich et al.[ix] suggest an α level of .25 or .50 and the investigator must make some assumption about the variance being zero. Mead et al.[x] recommend that in our situation, the never-pool procedure is best because the potential gain is small while there could be a loss. Wolde et al. [xi] and Donner [xii] also recommend against pooling. However, the analytic and simulation studies that were studied did not consider the configuration that we have, with large degrees of freedom for the residual. For this reason, and because of its importance for community interventions, we decided to study this matter further.

We studied the performance of pooling using simulation. Table 2 defines the terms used in the simulation. Table 3 shows the simulation model. Table 4 describes the simulation process.

[Table 2 and Table 3 and Table 4 about here]

Best Case Analysis:

The situation most favorable to the pooling procedure is when $\sigma_c = 0$; that is, there is no community variation, and the person-level test (F_p) would be appropriate. A low level of α, such as .05, is favorable because rejection will seldom occur, and the person-level test (F_p) will usually be performed. We chose N = 500, although it is not obvious whether this is optimal.

[Table 5 about here]

Table 5 shows simulation results for the best case simulation. The first set of simulations has 2 communities per treatment group. The first line shows power for a person-level test, the second for a community-level test, and the third from the pooling procedure.

The treatment effect varies from left to right. When the treatment effect is 0 (the null hypothesis is true) the size is .048 for the person level test, .049 for the community level test, and .046 for the pooled test. Thus, in each case, the size is about 0.05. This should be the case, under the null hypothesis. Note that under the alternative hypothesis, when the effect size is .14, the person-level test has power .882, the community-level test has power .414, and the pooled test has power .837. This is encouraging. The difference between the pooled power and the community level power is .837 - .414 = .423. We call this the "Maximum difference" or "maxdif". If maxdif is not big, there is no point in using the pooling process. Note that in the bottom of the table, for 5 communities per group, maxdif (effect size = .08) is only .09, and it is considerably less than .09 for other parameter combinations. This suggests we limit ourselves to 2, 3, and 4 observations per group. Figure 1 makes this point by plotting maxdif against the number of communities in this best case situation. If we want to gain at least 10 percentage points in power, it makes no sense to look at 5 or more pairs. This is true for 50, 100, or 500 people per community. The lowest line in the graph shows results for N = 50,000, which is comparable to a community that can be studied from some external data source that covers all people, such as

surgery rates.

[Figure 1 about here]

Simulation Study for 2, 3, or 4 Pairs.

Size for K=2.

We next consider the size of the pooled test as a function of α, restricting ourselves to N= 2, 3, and 4. Here, σ_c is not zero, and there is no treatment effect.

Figure 2 shows the estimated size of the test as a function of α for N=2. There are separate lines for different values of N, and values are averaged over several values of σ_c. Size should be .05, since the null hypothesis is true. But, it is substantially above .05 unless α = .25 or .50. These levels were also recommended in the literature. However, since these results are averaged over values of σ_c, there may be some values of σ_c or N in which the pooled test would be advantageous.

[Figure 2 about here]

Figure 3 shows the mean size, again under the null hypothesis, as a function of σ_c, averaged over all values of α, for different values of N. Note that for low values of σ_c, the size is lower than .05, but that it increases with σ_c and then, in some cases, decreases again. The picture is puzzling, but makes sense. Consider the line for N=1000 people per community. For low σ_c there is almost no community variation, and so the pooled test is approximately correct. For very high values of σ_c, the preliminary hypothesis would almost always be rejected, and the pooled version would again do the right thing. Size is better for small samples than for large samples.

[Figure 3 about here]

Figure 4 shows size as a function of σ_c and α, averaging over the number of persons per community. Clearly higher values of α give better control of size, but for low values of σ_c the size is too conservative. The size is acceptable for low σ_c, and high α.

[Figure 4 about here]

Power Analysis for K=2 : Maximum Difference in Power (Maxdif).

Figure 5 shows maxdif versus α and N, averaged over σ_c, where maxdif is the maximum difference between the pooled and the community tests in power over all the simulations. Maxdif is highest for moderate values of N (250 or 500) and best for low values of α.

[Figure 5 about here]

Figure 6 shows maxdif versus σ_c for and N, averaged over α. Power is best for low σ_c, and moderate N, except for n=50, which is totally different.

[Figure 6 about here]

The above figures have all been for K = 2. Similar figures could be drawn for K=3 and 4, with similar results: there is an extremely complicated relationship between σ_c, α, size, maxdif, K, and n. There are some situations in which size was appropriate, but they were not always situations in which maxdif was also high. Because these relationships are so complicated, we decided to look for specific situations in which the pooling method might be helpful. We reasoned that an investigator might we willing to change the pooling method if size and improvement were reasonable.

Situations in which pooling may be desirable:

Table 6 shows a summary of the simulations for K=2 in which the simulated size was < 6% (close to the nominal 5%). Results are shown as a function of N and α. For example, for N=50 and α=.05, the lowest value of σ_c is .000 and the highest is .060. For that range of σ_c, maxdif varied from 10.3 to 13.5 percentage points. Therefore, if one knew that σ_c was .06 or less, pooling could be expected to provide a modest increase in power (10.3 to 13.5 percentage points).

If N=100 and α=.05, the gain could be as high as 23 percentage points, but σ_c would have to be .02 or less, a more restrictive assumption. In general, the maximum value of σ_c decreases with N, but the potential gain increases with N. As α increases, σ_c increases somewhat and the percent gain decreases somewhat. The maximum power found (for effect size = .18) is about .10 to .13 for the parameter values considered, so an increase of 10 percentage points could be worth it. It is

unlikely that anyone would choose α = .50 because the gain is so small (even negative for N=50). A value of α = .10 seems reasonable, but .25 is perhaps safer if one is not willing to guess that σ_c is .2 or less.

[Table 6 about here]

Table 7 shows similar results for K=3 communities per group. Pooling might be considered valid using α = .05 for N <1000 (with if the assumed values of σ_c could be justified) or α = .10 for N=50 or 100.

[Table 7 about here]

Table 8 shows similar results for K=4 communities per group. One might consider pooling with α = .05 for N < 1000 if σ_c meets the requirements.

[Table 8 about here]

Empirical Values of σ_c

There are situations in which investigators would benefit from pooling if they knew the value of σ_c or at least knew that it was small. We have estimated values of σ_c from a variety of data sets, and rarely find values significantly different from zero -- in fact, the estimate is most often zero. However, these estimates are based on small values of K.

We collected a variety of data sets from community interventions, and used a mixed model analysis of variance to estimate σ^2_e (variance in smoking among people) and σ^2_{ct} (community by time component of variance)[xiii]. (That is, the data included time as a separate factor, and the dependent variable was a single point in time). We then calculated

$$\sigma_c = \sqrt{\frac{2\,\sigma^2_{ct}}{\sigma^2_e}}$$

which is equivalent to the value in our simulation in which σ^2_e was 1.0. The variance we usually compute is half the variance of this simulation (because it's a change score).

The data included community interventions on smoking status (9 times, 3 from kaiser); seatbelt use (7 studies, 3 from Kaiser); health status (4 studies, 3 from Kaiser); health status yes/no (4 studies, 3 from Kaiser); dietary fat (3 studies, all from Kaiser).

The distribution of estimates of σ_c (adjusted) is:

0	21
0-.0199	1
.02-.0299	4
.03-.0399	2
.04-.0499	0
.05-.0599	3
.06-.0699	1
.1175	1
total	33

Or look at average, by intervention, counting Kaiser only once:

Smoking	.00083 (1/8 non-zero)
HYN	.0124 (2)
Seatbelt	.028 (one huge outlier, BRFS)
Dietary Fat	.006 (1 --actually average of 3 Kaiser)+1 zero
EVGFP	.0368 (2)
Fiber	.01575 (3 kaiser, 1 eating pattern)

Table 6 suggests that if $\sigma_c < .06$ and N=50, pooling may be a good idea for any value of α. Based on these empirical values, it seems safe, since only 1 of the 30 estimates was above .06. Plus, these are empirical, not true variances. It looks as though pooling is "safe" for N=50 for any α, and for N=100 for some values of α. σ_c is probably below .2, and almost surely below .07. But there is, unfortunately, one huge outlier. Additional data mayu make the user more confident that the values are usually small.

Relationship to breaking the matches paper

Power is extremely low for K=2. This seems to disagree with a previous paper [xiv] which said that an unmatched t-test has power .39 when the effect size is 3.0. The first paper did not allow "n" to have any effect, so it's most similar to runs for n =

1000. σ_c was effectively 1.0 in that paper, which is very large. An effect size of 3 is $3\sigma_c$. For 2 pairs, N=1000, σ_c =.010, an effect size of .03 would be between .111 and .252. That's a little lower than .39.

Discussion

It is always preferable to study more communities, but desperate times call for desperate measures. If you must use 2, 3, or 4 pairs communities per group, 3, and 4 pairs, some of these situations may be helpful. Perhaps you can convince yourself that community variation should be small because sugjects were effectively randomized. Otherwise, look at our values of σ and see if they fit your situation. Given the relatively low gain, and the problems inherent in using this method, it is probably not advisable to use pooling in community level trials.

Table 1
Anova Table

Source	**SS**	**df**	**df'**	**MS**	**EMS**
Due Tx	Ss_a	1	1	MS_a	$\sigma^2_e + n\sigma^2_b + bn\sigma^2_a$?
Within Tx Among Communities	Ss_b	2(K-1)	4	MS_b	$\sigma^2_e + n\sigma^2_b$
Residual (Person) Ss_r	2K(n-1)	594	s^2_e	σ^2_e	

F-tests:		**DF (k=3,n=100)**	**Value**	.05 **Crit**
Community-level:	$\mathbf{F_c} = F_{(1,2(K-1))} = MS_a/MS_b$	1,4		7.71
Person-level (pooled)	$\mathbf{F_p} = F_{2(K-1),2K(n-1)} = MS_a/[(SS_b+SS_r)/(a(Kn-1))]$	1,598		3.84
Test for significant community variation	$\mathbf{F_v} = F_{2(K-1),wK(n-1)} = MS_b / s^2_e$	4,594		

Example for K=3, N=100. The pooling process tests whether $\mathbf{F_v}$ is significant at the α level (α not necessariily = .05). If significant, test for treatment effect using F_c. If not F_v not significant, test for the treatment effect with F_p.

Table 2.

Details of Simulation

K = # of Communities.

2,3,4 in great detail
5 through 15 in less detail

N = # of people per community
50,100,250,500,1000

σ^2_e = variation among persons within community. (1.0)

sigc = σ_c, true variation among communities.

Here, the dependent variable is a change score, not the individual values. Its variance is twice as big as the variance of the individual change scores (in a cross-sectional design). So, this variance is twice as big as the variance we usually discuss. (Thanks, ZF)

0,.005,.01, .025, .05, .10, .15, .20, .30

(PFAT is about .06, Fiber is about .03)

Observed community variation is a combination of σ_c and σ^2_e/N, so it decreases if there are more people per community.

Alpha = probability of type I error when we test whether $\sigma_c = 0$. (That is, when we decide whether it's ok to pool).

0,.05,.10,.25,.50

.25 is usually recommended in the literature

In the simulation, we generate a data set, test to decide whether to "pool" (using alpha), then calculate either a person-level or a community-level analysis.

Size = percent of time we reject when the null hypothesis (treatment effect is zero) is true. Estimated from simulation.

Maxdif = maximum difference in power between the usual (community level) test and the "pooled" test. Estimated from simulation.

Best Case: $\alpha = .05$ and $\sigma_c = 0$. That is, there really IS no community-level variation, and we give ourselves the maximum chance to detect it (pool 95% of the time)

Table 3
Simulation Model.

Model:

Let Y_{ikN} be the change score for person N, community K, treatment I

I = 2 tx
K = # of communities/tx
N = # of subjects per community

$Y_{ikN} \sim N(\mu_{ik}, \sigma^2_e)$

Where:

$\mu_{1k} \sim N\ (0, \sigma_c)$

$\mu_{2k} \sim N\ (\Delta, \sigma_c)$

sample means $Y_{ik.} \sim N(\mu_{ik}, \sigma^2_e/N)$ where N = # people/community

So, true change in control communities is zero, on average and true change in treatment communities is Δ, on average. If $\sigma_c = 0$, then all tx communities have mean 0, all control communities have mean Δ. [This is the simulation that was done in the breaking the matches paper --- true?].

Table 4 Simulation Process

Choose a value of K,N, σ_c, (σ^2_e=1), alpha (for pooling test)

There are K communities per group.

First, generate 2K true means, $\mu_{ik} \sim N(0, \sigma^2_C)$

Then, generate 2K sample means, $Y_{ik.}$ from $N(\mu_{ik}, \sigma^2_e/N)$

Calculate the variance of the $Y_{1k.}$ and the $Y_{2k.}$ - this is **MSA**

Then, generate a MSE value. The true value is 1, based on the model. I don't want to generate a value for each of 500 people per community, so I am generating a **MSE** separately.

s^2_e ~ chi-square / df, df = [IK(N-1)]

Chi-square ~ ? N[IK(N-1), 2IK(N-1)]

So, generate V fom N(IK(N-1),2IK(N-1)),

let MSE = s^2_e = V/[IK(N-1)]

Letting Δ vary, calculate

Numerator = $Y_{1k.}-(Y_{2k.}+ \Delta)$

Person-level denominator = $s^2/(IKN)$

Community-level denominator = MSA/(2K)

Person-level F-test is squared numerator over PLD

Community-level F test is squared numerator over CLD

Pooled method: (Bad name: call it peek and test?)

See if there is significant community variation

F = CLD/PLD "large" (p < ALPHA)

If F is large, use community-level analysis

If F is small, use person-level analysis

Change value of Δ and repeat

Do this 100,000 times.

Count number of times the hypothesis is rejected for each value of Δ and for each method = "size" or "power".

Table 5 best case, N=500 ,alpha=.05

[is N=500 really the best case?]

POWER FOR 2 PAIRS 10000 ITERATIONS
ALPHA = .05 N = 500

SC	METH	EFFECT SIZE .0	.02	.04	.06	.08	.10	.12	.14	.16	.18
.000	PERS	.048	.074	.151	.273	.436	.619	.771	.882	.948	.981
.000	COMM	.049	.059	.087	.128	.189	.257	.335	.414	.497	.572
.000	POOL	.046	.070	.143	.258	.412	.585	.730	.837	.900	.931

POWER FOR 3 PAIRS 10000 ITERATIONS
ALPHA = .05 N = 500

SC	METH	EFFECT SIZE .0	.02	.04	.06	.08	.10	.12	.14	.16	.18
.000	PERS	.051	.085	.197	.378	.588	.778	.905	.969	.993	.998
.000	COMM	.054	.072	.140	.250	.386	.542	.688	.813	.899	.947
.000	POOL	.048	.080	.186	.359	.558	.742	.865	.931	.961	.975

POWER FOR 4 PAIRS 10000 ITERATIONS
ALPHA = .05 N = 500

SC	METH	EFFECT SIZE .0	.02	.04	.06	.08	.10	.12	.14	.16	.18
.000	PERS	.054	.097	.243	.479	.711	.876	.963	.991	.999	1.000
.000	COMM	.050	.085	.185	.352	.563	.739	.873	.952	.985	.997
.000	POOL	.052	.091	.230	.453	.677	.841	.936	.976	.993	.998

POWER FOR 5 PAIRS 10000 ITERATIONS
ALPHA = .05 N = 500

SSc	METH	EFFECT SIZE .0	.02	.04	.06	.08	.10	.12	.14	.16	.18
.000	PERS	.052	.112	.299	.561	.808	.940	.987	.998	1.000	1.000
.000	COMM	.053	.105	.245	.464	.694	.867	.955	.989	.998	1.000
.000	POOL	.050	.107	.283	.537	.780	.917	.975	.994	.999	1.000

POWER FOR 6 PAIRS 10000 ITERATIONS
ALPHA = .05 N = 500

Sc	METH	EFFECT SIZE .0	.02	.04	.06	.08	.10	.12	.14	.16	.18
.000	PERS	.046	.122	.341	.650	.877	.973	.996	1.000	1.000	1.000
.000	COMM	.046	.105	.293	.560	.799	.936	.986	.998	1.000	1.000
.000	POOL	.045	.117	.327	.627	.856	.961	.992	.999	1.000	1.000

Table 6

Summary of Simulation Results Size < 6
(2000 replications)

NPAIR
2.00

	N				
	50	100	250	500	1000
ALPHA					
.05					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.020	.010	.010	.010
MAXDIF					
Minimum	10.30	22.05	40.00	41.20	41.75
Maximum	13.50	23.40	40.10	42.25	43.30
.10					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.020	.020	.010	.010
MAXDIF					
Minimum	10.15	19.60	36.05	37.15	37.75
Maximum	12.60	20.80	38.70	38.75	38.05
.25					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.050	.040	.030	.020
MAXDIF					
Minimum	4.40	13.35	20.45	21.45	21.60
Maximum	8.10	15.10	27.25	26.25	24.40
.50					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.060	.060	.060	.050
MAXDIF					
Minimum	-.95	2.25	3.65	1.95	2.35
Maximum	1.80	3.95	7.95	7.70	6.70

Table 7

Summary of Simulation Results Size < 6 and maxdif > 10

NPAIR
3.00

	N				
	50	100	250	500	1000
ALPHA					
.05					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.050	.040	.020	.020	.010
MAXDIF					
Minimum	8.70	17.30	19.35	19.80	17.75
Maximum	11.65	20.05	21.25	22.40	20.10
.10					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.050	.010	.000	.010
MAXDIF					
Minimum	6.75	13.55	15.35	17.30	15.15
Maximum	10.05	16.30	17.95	17.30	15.30
.25					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.060	.050	.030	.020
MAXDIF					
Minimum	3.05	6.10	7.50	6.90	7.50
Maximum	4.85	9.05	8.25	8.00	8.15
.50					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.060	.060	.060	.060
MAXDIF					
Minimum	-2.05	-1.05	.15	.20	.15
Maximum	-1.00	-.35	.75	.70	.70

Table 8

Summary of Simulation Results Size < 6 and maxdif > 10

NPAIR
4.00

	N				
	50	100	250	500	1000
ALPHA					
.05					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.050	.040	.020	.020	.000
MAXDIF					
Minimum	8.35	12.35	11.95	11.35	13.80
Maximum	11.65	13.25	13.80	13.85	13.80
.10					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.040	.010	.010	.010
MAXDIF					
Minimum	5.25	9.10	9.60	10.05	10.20
Maximum	7.85	10.60	10.20	11.25	10.55
.25					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.060	.060	.060	.060
MAXDIF					
Minimum	1.25	3.15	2.65	1.70	.60
Maximum	4.05	4.20	4.95	4.70	4.20
.50					
SIGCT					
Minimum	.000	.000	.000	.000	.000
Maximum	.060	.060	.060	.060	.060
MAXDIF					
Minimum	-2.05	-1.95	-.20	.00	.00
Maximum	-.75	-.50	.15	.30	.20

References

i. Blackburn H, Lueker RV, Kline FG, Bracht N, Carlow R, Jacobs D, Stauffer L, Taylor HL. The Minnesota Heart Health Program: A research and demonstration project in cardiovascular disease prevention. In Matarazzo JD, Weiss SM, Miller NE and Weiss SW (eds.). A Handbook of Health Enhancement and disease Prevention. Wiley, New York, 1984, pp. 1171-1178.

ii. McAlister A, Puska P, Salonen J, Tuomilehto J, Koskela K. Theory and action for health promotion: illustrations from the North Karelia project. American Journal of Public Health 72:43-50. 1982.

iii. Lefebvre R, Lasater TM, Carlton RA, Peterson G. Theory and delivery of health programming in the community: the Pawtucket Heart Health program. Preventive Medicine 16:80-95 1987.

iv. Farquhar JW, Fortmann SP, MacCoby N, Haskell WL, Williams PT, Flora JA, Taylor CB, Brown BW Jr., Solomon DS, Hulley SB. The Stanford five-city project: Design and methods. American Journal of Epidemiology, 122(2):323-334. 1985.

v. Fortman SP, Williams PT, Hulley SB, Haskell Wl, Farquhar JW. Effect of health education on dietary behavior: the Stanford three community study. American Journal of Clinical Nutrition 34:2030-2038. 1981.

vi. Shea S, Basch CE. A review of five major community-based cardiovascular disease prevention programs. Part I: Rationale, design, and theoretical framework. American Journal of Health Promotion 4(3):203-213. 1990.

vii. Bancroft T.A. On biases in estimation due to the use of preliminary tests of significance. Annals of Mathematical statistics 15:190-204. 1944.

viii. Paull A. E. On a preliminary test for pooling mean squares in the analysis of variance. Annals of Mathematical Statistics 21:539-556.

1950.

ix. Bozivich H, Bancroft T.A., Hartley H.O. Power of analysis of variance test procedures for certain incompletely specified models. Annals of Mathematical statistics 27:1017-1043. 1956.

x. Mead R, Bancroft T.A., Han Chien-Pai. Power of analysis of variance tests procedures for incompletely specified fixed models. Annals of statistics 3:797-808. 1975

xi. Wolde-Tsadik G, Afifi A.A. A comparison of the "sometimes pool", "sometimes switch", and "never pool" procedures in the two-way ANOVA random effects model. Technometrics 22:367-373. 1980.

xii. Donner A, Klar N. Statistical considerations in the design and analysis of community intervention trials. Journal of clinical epidemiology 49:435-439. 1996.

xiii. Andrilla CH et al. **Values of** σ_c. Unpublished.

xiv. Diehr P, Martin DC, Koepsell TD, Cheadle A. Breaking the matches in a paired t-test for community inverventions when the number of pairs is small. Statistics in Medicine 14:1491-1504. 1995.

Figure 1

Best Case Analysis

(sigma ct = 0, alpha=.05)

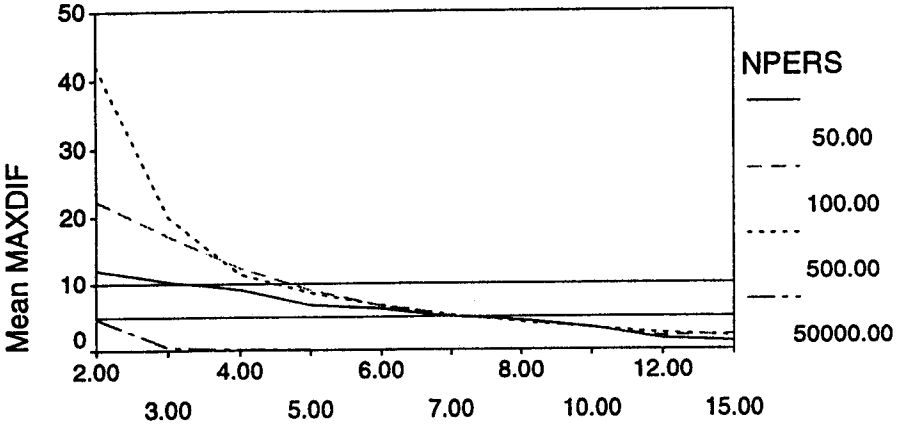
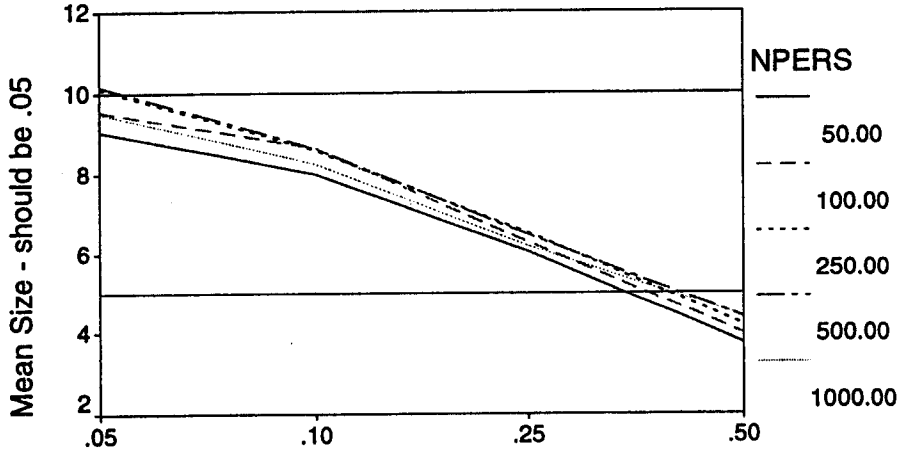


Figure 2

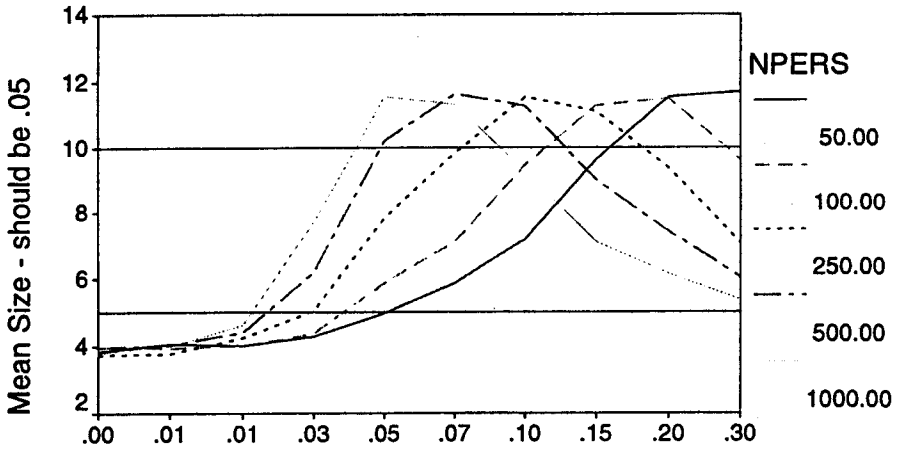
K=2 Size Analysis



(sigma ct = all, alpha=)

Figure 3

K=2 Size Analysis

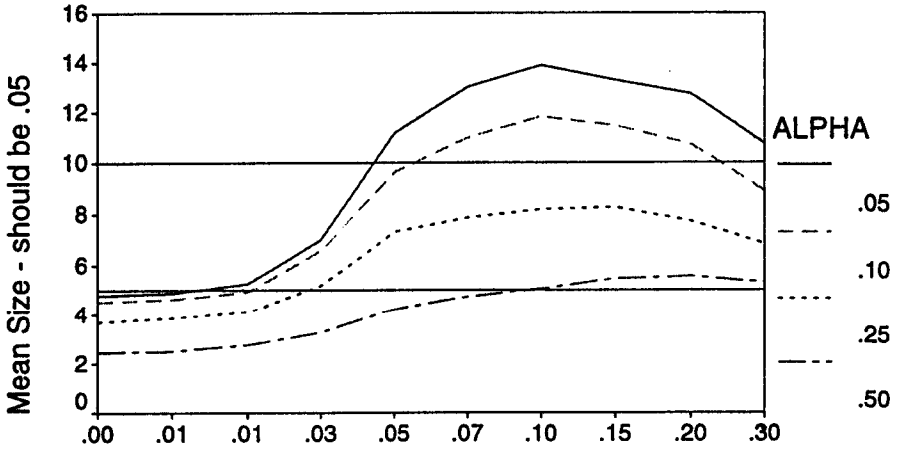


SIGCT

(sigma ct =, alpha= all)

Figure 4

K=2 Size Analysis

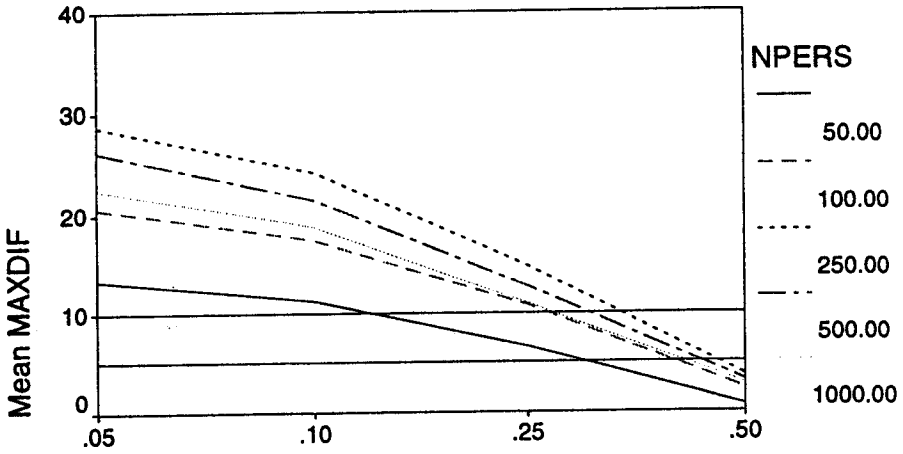


SIGCT

(npers = all)

Figure 5

K=2 Power Analysis

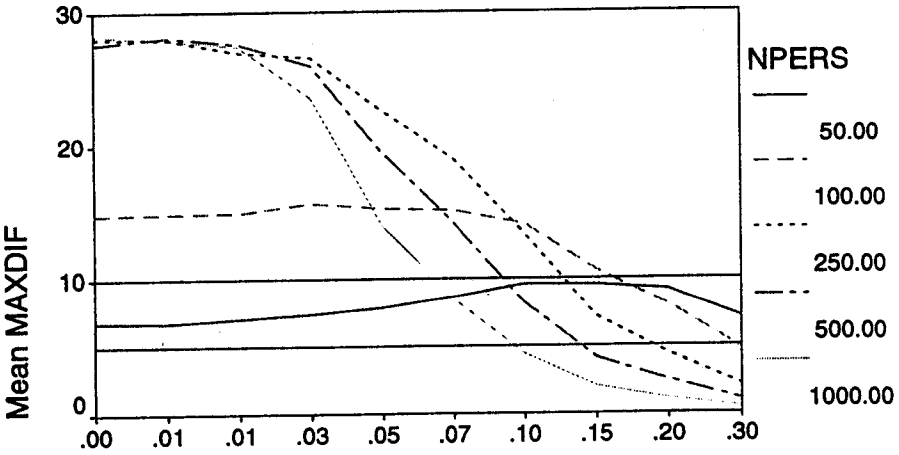


ALPHA

(sigma ct = all, alpha=)

Figure 6

K=2 Power Analysis



SIGCT

(sigma ct =, alpha= all)