Using multiple outcomes in intervention studies to increase statistical power: the Adjust NVar approach

D. V. M. Bishop

23/08/2021

## Abstract

The CONSORT guidelines for clinical trials recommend that the researcher should specify a single primary outcome, to guard against the raised risk of false positive findings when multiple measures are considered. It is, however, possible to include a suite of multiple outcomes in an intervention study, while controlling the false positive rate, provided the criterion for rejecting the null hypothesis specifies that N or more of the outcomes reach an agreed level of statistical significance, where N depends on the total number of outcome measures included in the study, and the correlation between them. I present simulations that explore the case when between 2 and 12 outcome measures are included in a suite, with the average correlation between measures ranging from zero to .8, and the true effect size ranging from 0 to .7. Two different methods of simulating outcome measures are compared, using a conventional null-hypothesis significance testing approach with alpha set at .05. In step 1, a table is created giving the minimum N significant outcomes (MinNSig) that is required for a given size of suite of outcome measures to control the false positive rate at 5%. These values are used in the Adjust Nvar approach adopted in step 2. For this step, data are simulated in which the MinNSig values are used for each size of suite of correlated outcomes and the resulting proportion of significant results is computed at different sample sizes and effect sizes. The Adjust Nvar approach can achieve a more efficient ratio between power and false positive rate than use of a single outcome when the suite includes 6 or more moderately intercorrelated outcome variables, and is considerably better in this regard than data reduction by extraction of a single principal component. Where it is feasible to have a suite of correlated outcome measures, then this might be a more efficient approach than reliance on a single primary outcome measure. In effect, it builds in an internal replication to the study.

## The case against multiple outcomes

The CONSORT guidelines for clinical trials (Moher et al. 2010) are very clear on the importance of having a single primary outcome:  
*All RCTs assess response variables, or outcomes (end points), for which the groups are compared. Most trials have several outcomes, some of which are of more interest than others. The primary outcome measure is the pre-specified outcome considered to be of greatest importance to relevant stakeholders (such a patients, policy makers, clinicians, funders) and is usually the one used in the sample size calculation. Some trials may have more than one primary outcome. Having several primary outcomes, however, incurs the problems of interpretation associated with multiplicity of analyses and is not recommended.*

This advice often creates a dilemma for the researcher: in many situations there are multiple measures that could plausibly be used to index the outcome. A common solution is to apply a Bonferroni correction to the alpha level used to test significance of individual measures, but this is over-conservative if, as is usually the case, the different outcomes are intercorrelated. Alternative methods are to adopt some process of data reduction, such as extracting a principal component from the measures that can be used as the primary outcome, or using a permutation test to derive exact probability of an observed pattern of results. Here I explore a further, very simple, option which I term the “Adjust Nvar” approach. The idea is that if one has a suite of outcomes, instead of adjusting the alpha level, one can adjust the number of outcomes that are required to achieve significance at the conventional alpha level of .05 to maintain an overall false positive rate of 1 in 20 or less.

To illustrate the idea with a realistic example, consider a behavioural intervention that is designed to improve language and literacy, and there are 6 measures where we might plausibly expect to see some benefit. The average intercorrelation between measures is .4. Suppose we find that none of the outcomes achieves Bonferroni-adjusted significance criterion of p < .008, but two of them reach significance at p < .05. Should we dismiss the trial as showing no benefit? We can use the binomial theorem to check the probability of obtaining this result if the null hypothesis is true and the measures are independent: it is 0.033, clearly below the 5% alpha level. But what if the measures are intercorrelated? A thought experiment helps here. Suppose we had six measures that were intercorrelated at .95 - in effect they would all be measures of the same thing, and so we would expect the probability of a false positive to approximate that obtained for a single measure. Extending this logic in a more graded way, the higher the correlation between the measures, the more measures would need to reach the original significance criterion to maintain the overall significance level below .05.

A simulation script was developed to test these intuitions and to obtain estimates of:  
-(A) the minimum number of outcome variables in a suite that would maintain the overall false positive rate at 1 in 20, if each individual measure was evaluated at the significance criterion of .05. This we term MinNSig.  
-(B) the power to detect a true effect, if the criterion for rejection the null hypothesis was based on the value of MinNSig identified at step A.

## Methods

Correlated variables were simulated using in the R programming language (R Core Team 2020). The script to generate and analyse simulated data is available on <https://github.com/oscci/MinSigVar>, and more technical details are provided in the Appendix. Two approaches to modeling correlated variables were compared, Method M and Method L.

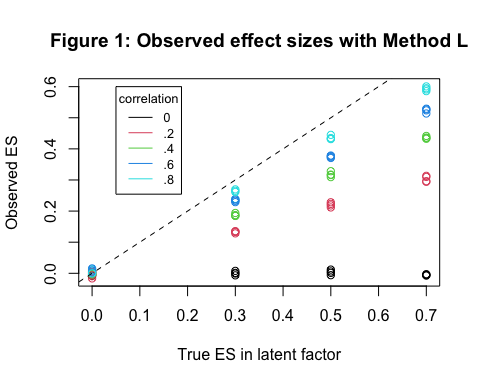
### Method M

Method M uses the *mvrnorm* function of the *MASS* package to generate a set of 12 outcome variables with a specified covariance matrix. For simplicity, all variables were simulated as random normal deviates with SD of 1, and the covariance matrix had a prespecified correlation, r, in all off-diagonal elements. The correlation varied across runs from 0 to .8 in steps of .2, and the number of simulated cases varied from 20 to 110 in steps of 30. Outcomes for Intervention (I) and Control (C) groups differed only in terms of the mean, which was always zero for the group C, and a given effect size, e, for group I.

### Method L

Method M is simple but potentially unrealistic as a representation of real data, because it implies that an intervention affects the whole set of outcome measures independently to a similar extent. With Method M it is possible to have a set of outcomes that are independent of one another yet all having the same effect size, which is unlike real-world data. Method L was designed to give a more realistic simulation, in which the set of 12 outcome measures are seen as indicators of an underlying latent variable, which mediates the intervention effect. In Method L, a latent variable is first simulated, with an effect size of either zero, for group C, or e for group I. Observed outcome measures are then simulated as having a specific correlation with the latent variable - i.e. the correlation determines the extent to which the outcomes act as indicators of the latent variable. This is achieved using the formula:  
- r \* latentvar + sqrt(1-r^2)\*err

where r is the correlation between latentvar and each outcome, and latentvar is a vector of random normal deviates that is the same for each outcome variable, while err is a vector of random normal deviates that differs for each outcome variable. Note that when outcome variables are generated this way, the mean intercorrelation between them will be r2. Thus if we want a set of outcome variables with mean intercorrelation of .4, we need to specify r in the formula above as sqrt(r) = .632.



For Method L, the observed effect size for the outcome measures will be lower than the effect size specified for the latent variable, because they are imperfect indicators. As shown in Figure 1, in the limiting case where the correlation is specified as zero, the observed effect size for observed outcomes will be zero, regardless of the effect size of the latent variable, because they do not share any variance with the latent variable. As the correlation between outcomes increases, the observed effect size approaches the true effect size in the latent variable. This made it potentially problematic when comparing power of a single outcome measure with a suite of measures, as this requires us to specify an effect size, and this depends on the intercorrelation between outcoms. The average observed effect size for all measures in a given condition was computed, and used as the basis for comparisons of efficiency between single and multiple measure scenarios.

### Data reduction

The size of the suite of outcome variables entered into later analysis ranged from 2 to 12 in steps of 2. For both Method M and L, principal components were computed for all outcome variables, using the base R function *prcomp* from the *stats* package. Thus PC2 is a principal component based on the first two outcome measures, PC4 based on the first four outcome measures, and so on.  
Power of analyses based on the principal components was compared with power obtained using the Adjust Nvar approach, as specified below.

### Simulation parameters

1000 simulations were run for each combination of:  
- sample size per group, ranging from 20 to 110 in steps of 30  
- correlation between outcome variables, ranging from .2 to .8 in steps of .2  
- true effect size, taking values of 0, .3, .5, or .7.  
- Method to generate data, either M or L

The data generated from each combination of conditions was used to derive results for different sizes of suites of outcome variables, ranging from 2 to 12 in steps of 2. Thus, the analysis was first conducted on the first 2 outcome measures, then on the first 4 outcome measures, and so on.

For each set of conditions, on each run, a one-tailed t-test was conducted to obtain a p-value for the comparison between C and I groups, assuming C would be lower. The p-values for outcome measures were rank ordered for each run and each suite size.

### Identifying MinNSig

To obtain MinNSig, the results were filtered to include only the runs where the null hypothesis was true, i.e. effect size = 0. Then, the proportion of p-values less than .05 was calculated for each rank for each number of outcome variables, to find the highest rank at which the overall proportion was less than .05. This is the MinNSig.  
Table 0 gives a toy example of the logic, using the case where we have either 2 or 4 outcome measures. Columns V1 to V2 show p-values for the t-test comparing the two groups each of the 4 outcome measures. Columns r2.1 and r2.2 show the same p-values rank ordered for just the first two measures; columns r4.1 to r4.4 show the p-values rank ordered for all 4 outcomes. We can then count the number of p-values that are below .05 for all 1000 runs for each ranked position. With 2 outcomes, if we take just the first ranked (lowest) p-value, the probability of it being lower than .05 is around .10. For the 2nd ranked p-value, the probability drops below .05, to .002. Thus we set MinNSig to 2. We can then turn to the case where we have four outcomes: the probability of the lowest p-value being below .05 is .185; the probability of the second lowest being below .05 is .014. Thus again, we set MinNSig to 2. As noted above, when the correlation between variables is zero, we can use the binomial theorem to compute values in the final row; however, when variables are intercorrelated, more p-values will be below .05, and so MinNSig may be higher.  
Because MinNSig moves in quantum steps, the effective false positive rate is often lower than .05. For instance, in the example above with a suite of four outcome measures, MinNSig is set to 2 when there is no correlation between outcomes, but this gives p = .014, rather than .05.

Table 1: Demonstration of how MinNSig is determined

| run | V1 | V2 | V3 | V4 | r2.1 | r2.2 | r4.1 | r4.2 | r4.3 | r4.4 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 0.793 | 0.549 | 0.229 | 0.846 | 0.549 | 0.793 | 0.229 | 0.549 | 0.793 | 0.846 |
| 2 | 0.274 | 0.859 | 0.091 | 0.951 | 0.274 | 0.859 | 0.091 | 0.274 | 0.859 | 0.951 |
| 3 | 0.139 | 0.45 | 0.306 | 0.64 | 0.139 | 0.45 | 0.139 | 0.306 | 0.45 | 0.64 |
| 4 | 0.036 | 0.111 | 0.429 | 0.521 | 0.036 | 0.111 | 0.036 | 0.111 | 0.429 | 0.521 |
| 5 | 0.729 | 0.671 | 0.876 | 0.047 | 0.671 | 0.729 | 0.047 | 0.671 | 0.729 | 0.876 |
| 6 | 0.091 | 0.885 | 0.308 | 0.678 | 0.091 | 0.885 | 0.091 | 0.308 | 0.678 | 0.885 |
| 7 | 0.475 | 0.043 | 0.781 | 0.944 | 0.043 | 0.475 | 0.043 | 0.475 | 0.781 | 0.944 |
| 8 | 0.382 | 0.255 | 0.577 | 0.988 | 0.255 | 0.382 | 0.255 | 0.382 | 0.577 | 0.988 |
| 9 | 0.337 | 0.684 | 0.684 | 0.552 | 0.337 | 0.684 | 0.337 | 0.552 | 0.684 | 0.684 |
| 10 | 0.032 | 0.588 | 0.641 | 0.309 | 0.032 | 0.588 | 0.032 | 0.309 | 0.588 | 0.641 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 999 | 0.057 | 0.364 | 0.745 | 0.188 | 0.057 | 0.364 | 0.057 | 0.188 | 0.364 | 0.745 |
| 1000 | 0.752 | 0.971 | 0.735 | 0.724 | 0.752 | 0.971 | 0.724 | 0.735 | 0.752 | 0.971 |
| N < .05 | . | . | . | . | 100 | 2 | 185 | 14 | 0 | 0 |
| p < .05 | . | . | . | . | 0.1 | 0.002 | 0.185 | 0.014 | 0 | 0 |

### Computing power using Adjust Nvar

For each run of the simulation, and each number of outcome measures (2, 4, 6, 8, 10 or 12), we take the value of MinNSig from the previous step and compute the proportion of p-values below .05, depending on the effect size, sample size and correlation between measures. For effect sizes above zero, this proportion corresponds to the statistical power.

Power using Adjust Nvar can be compared to:  
- power obtained with a single outcome measure for the same effect size and sample size  
- power obtained by using the principal component extracted for this set of outcome measures

## Results

### MinNSig

Table 2 shows results from a simulation of the Adjust Nvar approach, with the values in the body of the table showing MinNSig, the minimum number of measures that would maintain the overall false positive rate at 1 in 20, if each individual measure was evaluated at the significance criterion of .05. It can be seen that the requirements for MinNSig are rather more stringent for Method L then for Method M. Because the t-test statistic used to determine p-values is adjusted for sample size, these values are independent of numbers of subjects. In principle, researchers could use Table 2 to specify in their research protocol the minimum number of outcomes that would reach their significance level in order for the null hypothesis to be rejected, and it is recommended that the values for Method L be preferred, as the assumptions behind this method seem more realistic than those of Method M.

Table 2A: Values of MinNSig from Method M

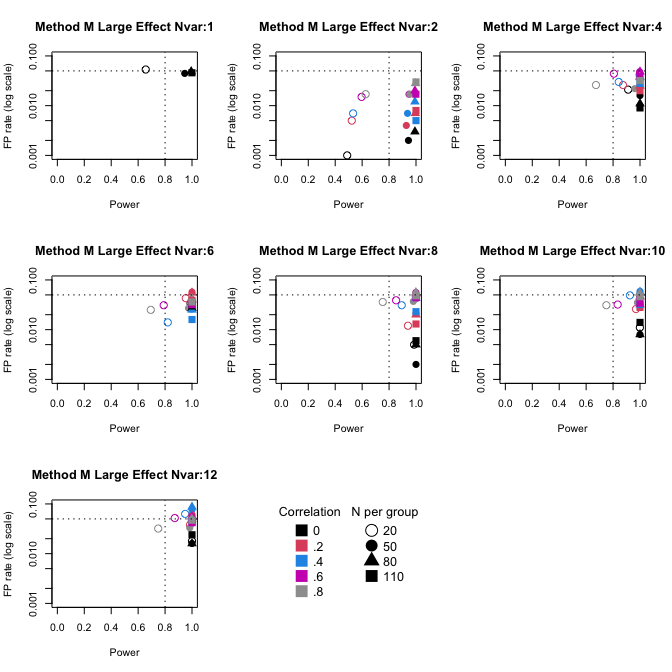
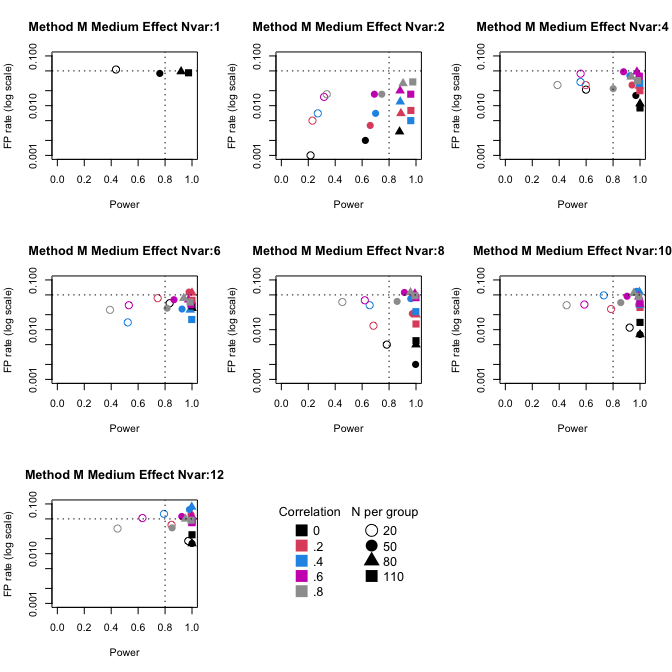
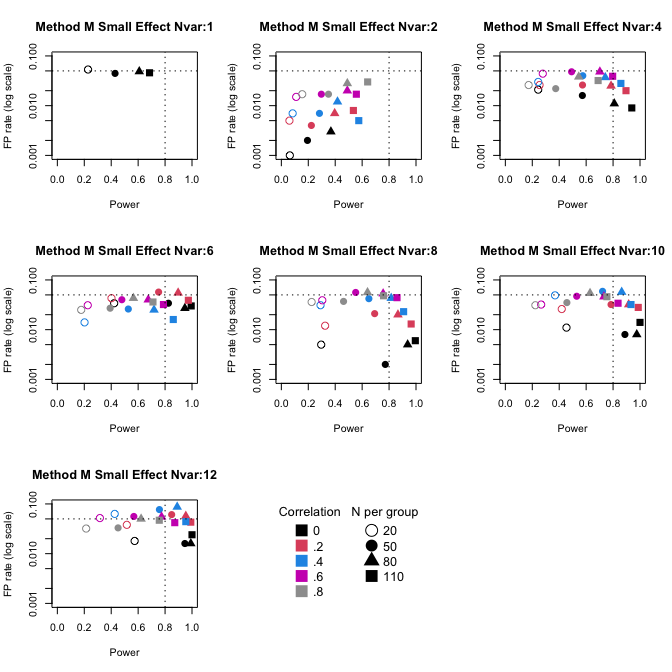
| corr | N2 | N4 | N6 | N8 | N10 | N12 |
| --- | --- | --- | --- | --- | --- | --- |
| 0.0 | 2 | 2 | 2 | 3 | 3 | 3 |
| 0.2 | 2 | 2 | 2 | 3 | 3 | 3 |
| 0.4 | 2 | 2 | 3 | 3 | 3 | 3 |
| 0.6 | 2 | 2 | 3 | 3 | 4 | 4 |
| 0.8 | 2 | 3 | 4 | 4 | 5 | 6 |

Table 2B: Values of MinNSig from Method L

| corr | N2 | N4 | N6 | N8 | N10 | N12 |
| --- | --- | --- | --- | --- | --- | --- |
| 0.0 | 2 | 2 | 2 | 2 | 3 | 3 |
| 0.2 | 2 | 2 | 3 | 3 | 3 | 4 |
| 0.4 | 2 | 2 | 3 | 3 | 3 | 4 |
| 0.6 | 2 | 3 | 3 | 4 | 4 | 5 |
| 0.8 | 2 | 3 | 3 | 4 | 4 | 5 |

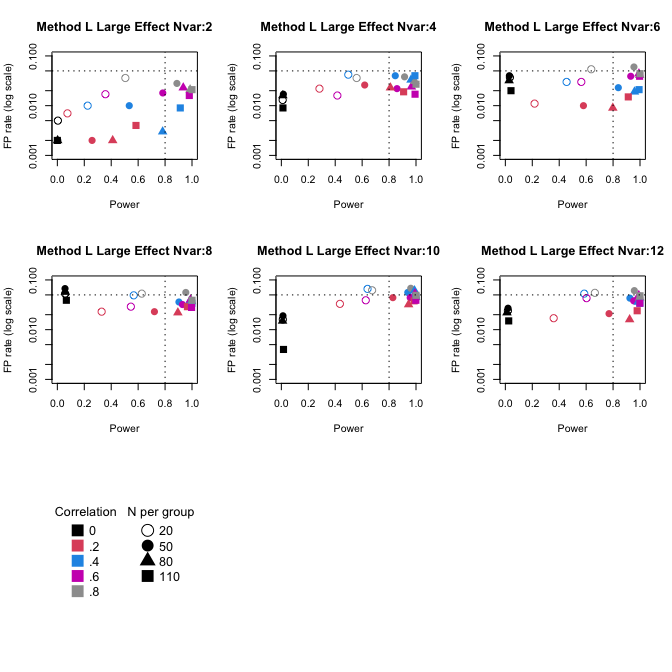
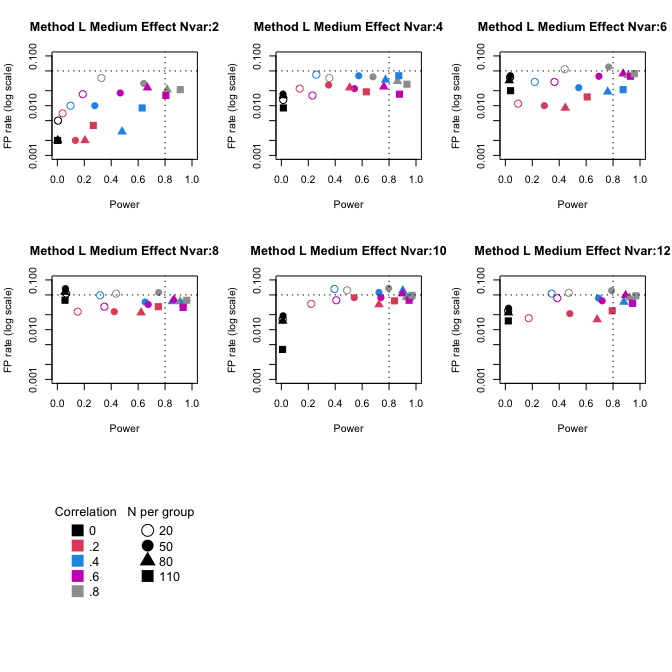
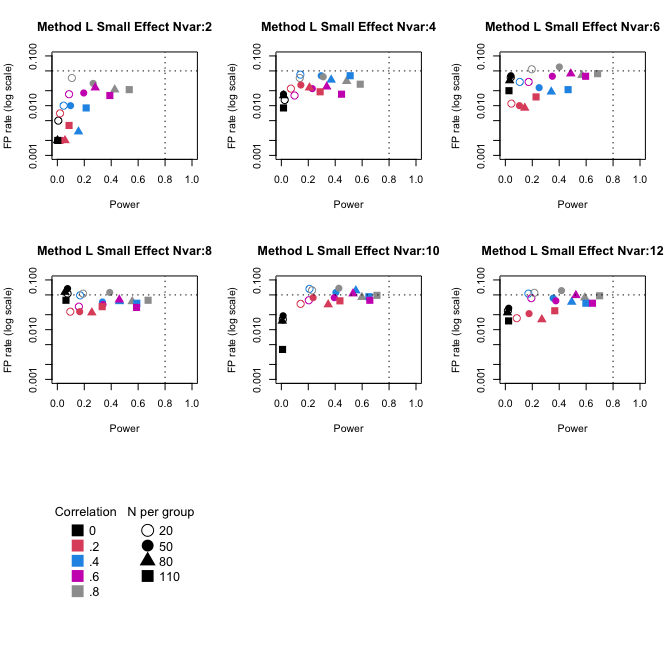
### Power of Adjust Nvar approach

Full tables of results for all combinations of parameters are provided in the Appendix. Figure 2 plots power vs false positive rate for different sizes of suite of outcome measures, including the case where a single outcome measure is used for comparison, using data from method M.



For these plots, the small, medium and large effect sizes correspond to Cohen’s d of .3, .5 and .7. An efficient method is one that gives power of .8 or above, and a false positive rate of .05 or less, i.e. the results should cluster in the bottom right quadrant. Power is dependent on sample size, as is evident from these plots where the unfilled figures, corresponding to the smallest sample size of 20 per group,fail to achieve power above .8, except at the highest effect size of .7. It is evident from inspection that when method M is used, the Adjust NVar approach compares favourably with the single outcome (Nvar 1) case, when there are 4 or more outcome measures in a suite. The power tends to be at least as high as for the single variable case, with lower false positive rates. Note, however, that the highest power is seen for the case where the variables are uncorrelated (black symbols), and this is an unrealistic scenario, especially for large suites of outcome measures.

Figure 3 gives an equivalent plot for method L.



As noted above, method L is more realistic, but because the effect size is computed from a latent variable, with outcome measures as indirect indicators, the observed effect sizes for individual variables in a suite will vary, depending on the correlation between the variables. Here, the designations of effect size as small, medium and large corresponds to the effect on the unobserved latent variable used to generate the outcome measures (as shown in Figure 1).

Method L provides power estimates that are more in keeping with the single outcome case (see previous Figure), but achieves this level of power while providing a more stringent control of false positives. The Adjust Nvar approach works well when there are between 4 and 6 outcome measures that are moderately intercorrelated (r between .4 and .6).

Finally, the Principal Components from each suite of outcomes were inspected, but they are not plotted here, as it was evident that they are reliably lower than the power obtained with the Adjust Nvar approach, regardless of the method used to generate the data (see Appendix for values).

## Discussion

The logic of conventional multiple testing is turned on its head with the Adjust Nvar approach, in that instead of adjusting the p-value used for significance (as in the Bonferroni correction, or methods based on False Discovery Rate), we adjust the number of individual outcome measures that we need to reach the intended significance criterion. This value can be easily computed using the binomial theorem for a given suite size of outcomes if the measures are uncorrelated, but in the context of intervention trials that is an unrealistic assumption.

One advantage of this approach is that it is more compatible with trials of interventions that are expected to affect a range of related processes, as is common in some fields such as education or speech and language therapy. In such cases, the need to specify a single primary outcome tends to create difficulties, because it is often unclear which of a suite of outcomes is likely to show an effect. Note that the Adjust Nvar approach does not give the researcher free rein to engage in p-hacking: the larger the suite of measures included in the study, the higher the value of MinNSig will be. It does, however, remove the need to put all one’s eggs in one basket by pre-specifying one measure as the primary outcome.

A second advantage of this approach is that once the values of MinNSig have been computed, it is very simple to apply, and could be used a priori to specify the criterion that will be used in a protocol, assuming the researcher already has a rough idea of the degree of intercorrelation between outcome measures. This may help guard against a tendency to explore different kinds of correction for multiple hypothesis testing only after viewing the data (Lazic 2021).

A third advantage is that in effect, by including multiple outcome measures, one can improve the efficiency of a study, in terms of the trade-off between power and false positives. If several outcome measures are seen as imperfect proxy indicators of an underlying latent construct, then we are in effect building in a degree of within-study replication if we require that more than one measure shows the same effect in the same direction before we reject the null hypothesis.

The results obtained with this approach depend crucially on assumptions embodied in the simulation that is used to derive predictions. Outcome measures simulated here are normally distributed, and quite uniform in their covariance structure. In Method L, which seems a more realistic approach than Method M, the set of correlated measures was generated by first creating a random normal deviate that was treated as the latent variable, and then creating outcome measures with a specific degree of correlation with the latent variable. It would be possible to generate datasets with different underlying covariance structures to be tested in the same way, but that is beyond the scope of this paper.

# Appendix

Computed power for Method M

Power table for Method M (where ES = 0, values are false positive rate)

| ES | obsES | nsub | corr | N1 | N2 | N4 | N6 | N8 | N10 | N12 | PC2 | PC4 | PC6 | PC8 | PC10 | PC12 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.0 | 0.010 | 20 | 0.0 | 0.053 | 0.001 | 0.021 | 0.034 | 0.005 | 0.011 | 0.018 | 0.015 | 0.008 | 0.004 | 0.001 | 0.021 | 0.000 |
| 0.0 | -0.011 | 20 | 0.2 |  | 0.005 | 0.026 | 0.043 | 0.012 | 0.026 | 0.038 | 0.027 | 0.016 | 0.012 | 0.005 | 0.026 | 0.001 |
| 0.0 | -0.006 | 20 | 0.4 |  | 0.007 | 0.030 | 0.014 | 0.031 | 0.049 | 0.063 | 0.035 | 0.037 | 0.037 | 0.007 | 0.030 | 0.000 |
| 0.0 | 0.005 | 20 | 0.6 |  | 0.015 | 0.044 | 0.031 | 0.039 | 0.032 | 0.052 | 0.041 | 0.045 | 0.043 | 0.015 | 0.044 | 0.006 |
| 0.0 | -0.012 | 20 | 0.8 |  | 0.017 | 0.026 | 0.025 | 0.036 | 0.031 | 0.032 | 0.043 | 0.044 | 0.038 | 0.017 | 0.048 | 0.015 |
| 0.0 | -0.011 | 50 | 0.0 | 0.052 | 0.002 | 0.016 | 0.034 | 0.002 | 0.008 | 0.016 | 0.017 | 0.012 | 0.006 | 0.002 | 0.016 | 0.000 |
| 0.0 | 0.010 | 50 | 0.2 |  | 0.004 | 0.026 | 0.057 | 0.021 | 0.032 | 0.061 | 0.032 | 0.028 | 0.027 | 0.004 | 0.026 | 0.000 |
| 0.0 | 0.009 | 50 | 0.4 |  | 0.007 | 0.040 | 0.026 | 0.042 | 0.059 | 0.077 | 0.047 | 0.041 | 0.032 | 0.007 | 0.040 | 0.001 |
| 0.0 | -0.003 | 50 | 0.6 |  | 0.017 | 0.048 | 0.040 | 0.056 | 0.047 | 0.056 | 0.064 | 0.071 | 0.065 | 0.017 | 0.048 | 0.003 |
| 0.0 | -0.004 | 50 | 0.8 |  | 0.017 | 0.022 | 0.027 | 0.037 | 0.035 | 0.033 | 0.047 | 0.049 | 0.045 | 0.017 | 0.048 | 0.011 |
| 0.0 | -0.002 | 80 | 0.0 | 0.041 | 0.003 | 0.011 | 0.027 | 0.005 | 0.008 | 0.016 | 0.043 | 0.019 | 0.014 | 0.003 | 0.011 | 0.000 |
| 0.0 | -0.001 | 80 | 0.2 |  | 0.007 | 0.025 | 0.055 | 0.020 | 0.032 | 0.057 | 0.039 | 0.035 | 0.038 | 0.007 | 0.025 | 0.000 |
| 0.0 | 0.004 | 80 | 0.4 |  | 0.012 | 0.037 | 0.025 | 0.043 | 0.057 | 0.087 | 0.048 | 0.051 | 0.045 | 0.012 | 0.037 | 0.002 |
| 0.0 | -0.003 | 80 | 0.6 |  | 0.020 | 0.048 | 0.040 | 0.053 | 0.046 | 0.055 | 0.040 | 0.043 | 0.044 | 0.020 | 0.048 | 0.004 |
| 0.0 | 0.005 | 80 | 0.8 |  | 0.028 | 0.038 | 0.043 | 0.056 | 0.054 | 0.050 | 0.046 | 0.044 | 0.043 | 0.028 | 0.071 | 0.021 |
| 0.0 | -0.006 | 110 | 0.0 | 0.044 | 0.000 | 0.009 | 0.030 | 0.006 | 0.014 | 0.024 | 0.023 | 0.026 | 0.018 | 0.000 | 0.009 | 0.000 |
| 0.0 | -0.003 | 110 | 0.2 |  | 0.008 | 0.020 | 0.039 | 0.013 | 0.028 | 0.043 | 0.039 | 0.031 | 0.034 | 0.008 | 0.020 | 0.000 |
| 0.0 | -0.002 | 110 | 0.4 |  | 0.005 | 0.028 | 0.016 | 0.023 | 0.032 | 0.044 | 0.043 | 0.040 | 0.044 | 0.005 | 0.028 | 0.002 |
| 0.0 | -0.009 | 110 | 0.6 |  | 0.017 | 0.039 | 0.032 | 0.044 | 0.034 | 0.042 | 0.040 | 0.042 | 0.035 | 0.017 | 0.039 | 0.006 |
| 0.0 | 0.000 | 110 | 0.8 |  | 0.030 | 0.032 | 0.036 | 0.048 | 0.046 | 0.047 | 0.058 | 0.065 | 0.062 | 0.030 | 0.064 | 0.014 |
| 0.3 | 0.290 | 20 | 0.0 | 0.250 | 0.063 | 0.244 | 0.422 | 0.296 | 0.454 | 0.574 | 0.056 | 0.039 | 0.027 | 0.063 | 0.244 | 0.003 |
| 0.3 | 0.292 | 20 | 0.2 |  | 0.060 | 0.254 | 0.402 | 0.325 | 0.419 | 0.516 | 0.122 | 0.159 | 0.182 | 0.060 | 0.254 | 0.013 |
| 0.3 | 0.289 | 20 | 0.4 |  | 0.084 | 0.244 | 0.202 | 0.291 | 0.369 | 0.426 | 0.156 | 0.174 | 0.180 | 0.084 | 0.244 | 0.031 |
| 0.3 | 0.301 | 20 | 0.6 |  | 0.110 | 0.278 | 0.226 | 0.304 | 0.266 | 0.316 | 0.151 | 0.154 | 0.155 | 0.110 | 0.278 | 0.047 |
| 0.3 | 0.297 | 20 | 0.8 |  | 0.155 | 0.173 | 0.177 | 0.225 | 0.223 | 0.214 | 0.139 | 0.131 | 0.134 | 0.155 | 0.249 | 0.101 |
| 0.3 | 0.295 | 50 | 0.0 | 0.431 | 0.194 | 0.572 | 0.825 | 0.772 | 0.888 | 0.947 | 0.147 | 0.110 | 0.100 | 0.194 | 0.572 | 0.036 |
| 0.3 | 0.308 | 50 | 0.2 |  | 0.223 | 0.573 | 0.752 | 0.693 | 0.785 | 0.849 | 0.349 | 0.396 | 0.444 | 0.223 | 0.573 | 0.078 |
| 0.3 | 0.313 | 50 | 0.4 |  | 0.283 | 0.575 | 0.527 | 0.650 | 0.722 | 0.758 | 0.310 | 0.332 | 0.340 | 0.283 | 0.575 | 0.143 |
| 0.3 | 0.300 | 50 | 0.6 |  | 0.297 | 0.492 | 0.479 | 0.552 | 0.530 | 0.568 | 0.292 | 0.306 | 0.309 | 0.297 | 0.492 | 0.192 |
| 0.3 | 0.295 | 50 | 0.8 |  | 0.350 | 0.374 | 0.392 | 0.462 | 0.457 | 0.451 | 0.249 | 0.248 | 0.244 | 0.350 | 0.521 | 0.258 |
| 0.3 | 0.305 | 80 | 0.0 | 0.592 | 0.368 | 0.809 | 0.947 | 0.938 | 0.976 | 0.991 | 0.229 | 0.192 | 0.152 | 0.368 | 0.809 | 0.131 |
| 0.3 | 0.297 | 80 | 0.2 |  | 0.396 | 0.785 | 0.897 | 0.866 | 0.916 | 0.954 | 0.448 | 0.480 | 0.481 | 0.396 | 0.785 | 0.203 |
| 0.3 | 0.300 | 80 | 0.4 |  | 0.417 | 0.742 | 0.718 | 0.816 | 0.864 | 0.890 | 0.409 | 0.419 | 0.436 | 0.417 | 0.742 | 0.248 |
| 0.3 | 0.302 | 80 | 0.6 |  | 0.490 | 0.702 | 0.675 | 0.756 | 0.732 | 0.774 | 0.383 | 0.400 | 0.400 | 0.490 | 0.702 | 0.346 |
| 0.3 | 0.295 | 80 | 0.8 |  | 0.490 | 0.545 | 0.565 | 0.640 | 0.629 | 0.622 | 0.334 | 0.342 | 0.342 | 0.490 | 0.656 | 0.410 |
| 0.3 | 0.288 | 110 | 0.0 | 0.710 | 0.522 | 0.939 | 0.994 | 0.995 | 1.000 | 1.000 | 0.233 | 0.204 | 0.166 | 0.522 | 0.939 | 0.268 |
| 0.3 | 0.295 | 110 | 0.2 |  | 0.535 | 0.896 | 0.973 | 0.964 | 0.987 | 0.991 | 0.495 | 0.503 | 0.513 | 0.535 | 0.896 | 0.343 |
| 0.3 | 0.300 | 110 | 0.4 |  | 0.575 | 0.858 | 0.861 | 0.908 | 0.933 | 0.954 | 0.464 | 0.482 | 0.486 | 0.575 | 0.858 | 0.404 |
| 0.3 | 0.290 | 110 | 0.6 |  | 0.557 | 0.797 | 0.788 | 0.859 | 0.837 | 0.872 | 0.427 | 0.434 | 0.446 | 0.557 | 0.797 | 0.436 |
| 0.3 | 0.294 | 110 | 0.8 |  | 0.641 | 0.689 | 0.710 | 0.759 | 0.754 | 0.757 | 0.435 | 0.434 | 0.434 | 0.641 | 0.778 | 0.555 |
| 0.5 | 0.477 | 20 | 0.0 | 0.460 | 0.217 | 0.599 | 0.834 | 0.783 | 0.923 | 0.972 | 0.125 | 0.090 | 0.071 | 0.217 | 0.599 | 0.049 |
| 0.5 | 0.490 | 20 | 0.2 |  | 0.231 | 0.597 | 0.745 | 0.684 | 0.785 | 0.848 | 0.273 | 0.329 | 0.377 | 0.231 | 0.597 | 0.081 |
| 0.5 | 0.484 | 20 | 0.4 |  | 0.271 | 0.558 | 0.523 | 0.656 | 0.731 | 0.792 | 0.334 | 0.364 | 0.375 | 0.271 | 0.558 | 0.130 |
| 0.5 | 0.476 | 20 | 0.6 |  | 0.317 | 0.559 | 0.531 | 0.620 | 0.587 | 0.632 | 0.301 | 0.323 | 0.319 | 0.317 | 0.559 | 0.205 |
| 0.5 | 0.470 | 20 | 0.8 |  | 0.337 | 0.387 | 0.391 | 0.453 | 0.455 | 0.447 | 0.240 | 0.243 | 0.246 | 0.337 | 0.485 | 0.256 |
| 0.5 | 0.474 | 50 | 0.0 | 0.813 | 0.624 | 0.969 | 0.997 | 0.997 | 1.000 | 1.000 | 0.257 | 0.209 | 0.202 | 0.624 | 0.969 | 0.386 |
| 0.5 | 0.482 | 50 | 0.2 |  | 0.660 | 0.941 | 0.978 | 0.973 | 0.991 | 0.993 | 0.497 | 0.516 | 0.513 | 0.660 | 0.941 | 0.472 |
| 0.5 | 0.493 | 50 | 0.4 |  | 0.700 | 0.925 | 0.926 | 0.961 | 0.974 | 0.981 | 0.472 | 0.486 | 0.495 | 0.700 | 0.925 | 0.536 |
| 0.5 | 0.489 | 50 | 0.6 |  | 0.691 | 0.878 | 0.866 | 0.914 | 0.904 | 0.924 | 0.458 | 0.466 | 0.472 | 0.691 | 0.878 | 0.569 |
| 0.5 | 0.485 | 50 | 0.8 |  | 0.746 | 0.801 | 0.815 | 0.860 | 0.857 | 0.852 | 0.419 | 0.431 | 0.431 | 0.746 | 0.869 | 0.672 |
| 0.5 | 0.483 | 80 | 0.0 | 0.950 | 0.878 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.304 | 0.291 | 0.252 | 0.878 | 1.000 | 0.781 |
| 0.5 | 0.490 | 80 | 0.2 |  | 0.888 | 0.995 | 1.000 | 1.000 | 1.000 | 1.000 | 0.506 | 0.516 | 0.518 | 0.888 | 0.995 | 0.783 |
| 0.5 | 0.483 | 80 | 0.4 |  | 0.883 | 0.985 | 0.987 | 0.994 | 0.996 | 0.998 | 0.470 | 0.473 | 0.472 | 0.883 | 0.985 | 0.791 |
| 0.5 | 0.480 | 80 | 0.6 |  | 0.882 | 0.977 | 0.975 | 0.992 | 0.991 | 0.998 | 0.490 | 0.495 | 0.493 | 0.882 | 0.977 | 0.821 |
| 0.5 | 0.485 | 80 | 0.8 |  | 0.904 | 0.934 | 0.939 | 0.959 | 0.954 | 0.952 | 0.486 | 0.485 | 0.488 | 0.904 | 0.963 | 0.861 |
| 0.5 | 0.486 | 110 | 0.0 | 0.976 | 0.948 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.359 | 0.334 | 0.313 | 0.948 | 1.000 | 0.905 |
| 0.5 | 0.486 | 110 | 0.2 |  | 0.962 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 0.479 | 0.474 | 0.474 | 0.962 | 0.999 | 0.931 |
| 0.5 | 0.488 | 110 | 0.4 |  | 0.961 | 0.998 | 0.998 | 0.999 | 1.000 | 1.000 | 0.513 | 0.515 | 0.515 | 0.961 | 0.998 | 0.933 |
| 0.5 | 0.477 | 110 | 0.6 |  | 0.961 | 0.995 | 0.992 | 0.999 | 0.999 | 0.999 | 0.509 | 0.509 | 0.509 | 0.961 | 0.995 | 0.936 |
| 0.5 | 0.489 | 110 | 0.8 |  | 0.976 | 0.982 | 0.986 | 0.992 | 0.991 | 0.993 | 0.519 | 0.519 | 0.519 | 0.976 | 0.993 | 0.958 |
| 0.7 | 0.659 | 20 | 0.0 | 0.707 | 0.489 | 0.912 | 0.989 | 0.986 | 0.999 | 1.000 | 0.213 | 0.160 | 0.153 | 0.489 | 0.912 | 0.230 |
| 0.7 | 0.665 | 20 | 0.2 |  | 0.523 | 0.874 | 0.954 | 0.940 | 0.970 | 0.985 | 0.378 | 0.425 | 0.446 | 0.523 | 0.874 | 0.291 |
| 0.7 | 0.664 | 20 | 0.4 |  | 0.533 | 0.842 | 0.820 | 0.894 | 0.925 | 0.949 | 0.444 | 0.462 | 0.472 | 0.533 | 0.842 | 0.375 |
| 0.7 | 0.651 | 20 | 0.6 |  | 0.596 | 0.806 | 0.790 | 0.852 | 0.834 | 0.872 | 0.425 | 0.433 | 0.438 | 0.596 | 0.806 | 0.455 |
| 0.7 | 0.659 | 20 | 0.8 |  | 0.626 | 0.673 | 0.694 | 0.753 | 0.750 | 0.749 | 0.401 | 0.404 | 0.409 | 0.626 | 0.775 | 0.528 |
| 0.7 | 0.655 | 50 | 0.0 | 0.972 | 0.944 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.338 | 0.284 | 0.255 | 0.944 | 1.000 | 0.887 |
| 0.7 | 0.664 | 50 | 0.2 |  | 0.928 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.482 | 0.476 | 0.486 | 0.928 | 1.000 | 0.882 |
| 0.7 | 0.660 | 50 | 0.4 |  | 0.937 | 0.994 | 0.997 | 1.000 | 1.000 | 1.000 | 0.528 | 0.525 | 0.527 | 0.937 | 0.994 | 0.901 |
| 0.7 | 0.662 | 50 | 0.6 |  | 0.948 | 0.997 | 0.997 | 0.999 | 0.997 | 0.999 | 0.489 | 0.491 | 0.491 | 0.948 | 0.997 | 0.911 |
| 0.7 | 0.663 | 50 | 0.8 |  | 0.954 | 0.961 | 0.974 | 0.980 | 0.981 | 0.983 | 0.466 | 0.468 | 0.470 | 0.954 | 0.982 | 0.932 |
| 0.7 | 0.660 | 80 | 0.0 | 0.996 | 0.992 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.355 | 0.335 | 0.329 | 0.992 | 1.000 | 0.988 |
| 0.7 | 0.661 | 80 | 0.2 |  | 0.995 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.502 | 0.497 | 0.494 | 0.995 | 1.000 | 0.990 |
| 0.7 | 0.659 | 80 | 0.4 |  | 0.992 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.494 | 0.491 | 0.492 | 0.992 | 1.000 | 0.990 |
| 0.7 | 0.659 | 80 | 0.6 |  | 0.992 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.492 | 0.492 | 0.492 | 0.992 | 1.000 | 0.983 |
| 0.7 | 0.663 | 80 | 0.8 |  | 0.996 | 0.997 | 0.997 | 0.999 | 0.999 | 0.999 | 0.513 | 0.513 | 0.513 | 0.996 | 0.999 | 0.991 |
| 0.7 | 0.660 | 110 | 0.0 | 0.997 | 0.997 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.402 | 0.375 | 0.353 | 0.997 | 1.000 | 0.997 |
| 0.7 | 0.662 | 110 | 0.2 |  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.485 | 0.502 | 0.498 | 1.000 | 1.000 | 1.000 |
| 0.7 | 0.654 | 110 | 0.4 |  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.492 | 0.491 | 0.490 | 1.000 | 1.000 | 0.998 |
| 0.7 | 0.651 | 110 | 0.6 |  | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.513 | 0.513 | 0.513 | 0.999 | 1.000 | 0.999 |
| 0.7 | 0.659 | 110 | 0.8 |  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.535 | 0.535 | 0.535 | 1.000 | 1.000 | 0.999 |

Power table for Method L (where ES = 0, values are false positive rate)

| ES | obsES | nsub | corr | N1 | N2 | N4 | N6 | N8 | N10 | N12 | PC2 | PC4 | PC6 | PC8 | PC10 | PC12 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.0 | 0.004 | 20 | 0.0 | 0.049 | 0.005 | 0.013 | 0.037 | 0.053 | 0.016 | 0.024 | 0.017 | 0.004 | 0.003 | 0.005 | 0.013 | 0.000 |
| 0.0 | -0.016 | 20 | 0.2 |  | 0.007 | 0.022 | 0.011 | 0.023 | 0.033 | 0.017 | 0.025 | 0.020 | 0.023 | 0.007 | 0.022 | 0.000 |
| 0.0 | 0.000 | 20 | 0.4 |  | 0.010 | 0.042 | 0.030 | 0.049 | 0.066 | 0.053 | 0.037 | 0.035 | 0.027 | 0.010 | 0.042 | 0.001 |
| 0.0 | 0.005 | 20 | 0.6 |  | 0.017 | 0.016 | 0.030 | 0.029 | 0.039 | 0.043 | 0.035 | 0.045 | 0.044 | 0.017 | 0.043 | 0.006 |
| 0.0 | 0.013 | 20 | 0.8 |  | 0.036 | 0.036 | 0.054 | 0.053 | 0.062 | 0.055 | 0.059 | 0.056 | 0.063 | 0.036 | 0.058 | 0.017 |
| 0.0 | -0.008 | 50 | 0.0 | 0.070 | 0.002 | 0.017 | 0.040 | 0.067 | 0.019 | 0.027 | 0.028 | 0.013 | 0.015 | 0.002 | 0.017 | 0.000 |
| 0.0 | -0.009 | 50 | 0.2 |  | 0.002 | 0.026 | 0.010 | 0.023 | 0.044 | 0.021 | 0.040 | 0.036 | 0.040 | 0.002 | 0.026 | 0.000 |
| 0.0 | -0.003 | 50 | 0.4 |  | 0.010 | 0.040 | 0.023 | 0.036 | 0.056 | 0.043 | 0.056 | 0.048 | 0.051 | 0.010 | 0.040 | 0.001 |
| 0.0 | 0.016 | 50 | 0.6 |  | 0.018 | 0.022 | 0.039 | 0.032 | 0.044 | 0.038 | 0.047 | 0.054 | 0.059 | 0.018 | 0.051 | 0.008 |
| 0.0 | 0.009 | 50 | 0.8 |  | 0.028 | 0.038 | 0.060 | 0.056 | 0.068 | 0.061 | 0.062 | 0.065 | 0.066 | 0.028 | 0.069 | 0.010 |
| 0.0 | -0.006 | 80 | 0.0 | 0.056 | 0.002 | 0.016 | 0.032 | 0.057 | 0.015 | 0.022 | 0.024 | 0.021 | 0.009 | 0.002 | 0.016 | 0.000 |
| 0.0 | 0.004 | 80 | 0.2 |  | 0.002 | 0.023 | 0.009 | 0.022 | 0.032 | 0.016 | 0.038 | 0.039 | 0.044 | 0.002 | 0.023 | 0.000 |
| 0.0 | -0.006 | 80 | 0.4 |  | 0.003 | 0.033 | 0.019 | 0.037 | 0.061 | 0.036 | 0.049 | 0.050 | 0.051 | 0.003 | 0.033 | 0.000 |
| 0.0 | 0.009 | 80 | 0.6 |  | 0.023 | 0.024 | 0.044 | 0.040 | 0.053 | 0.049 | 0.055 | 0.057 | 0.059 | 0.023 | 0.060 | 0.006 |
| 0.0 | 0.002 | 80 | 0.8 |  | 0.021 | 0.031 | 0.041 | 0.037 | 0.045 | 0.044 | 0.046 | 0.043 | 0.043 | 0.021 | 0.047 | 0.011 |
| 0.0 | -0.001 | 110 | 0.0 | 0.050 | 0.002 | 0.009 | 0.020 | 0.039 | 0.004 | 0.015 | 0.034 | 0.021 | 0.016 | 0.002 | 0.009 | 0.000 |
| 0.0 | -0.004 | 110 | 0.2 |  | 0.004 | 0.019 | 0.015 | 0.029 | 0.038 | 0.024 | 0.038 | 0.050 | 0.043 | 0.004 | 0.019 | 0.000 |
| 0.0 | -0.002 | 110 | 0.4 |  | 0.009 | 0.040 | 0.021 | 0.034 | 0.046 | 0.034 | 0.043 | 0.036 | 0.035 | 0.009 | 0.040 | 0.001 |
| 0.0 | 0.002 | 110 | 0.6 |  | 0.016 | 0.017 | 0.039 | 0.028 | 0.039 | 0.034 | 0.054 | 0.052 | 0.051 | 0.016 | 0.053 | 0.001 |
| 0.0 | 0.000 | 110 | 0.8 |  | 0.021 | 0.027 | 0.044 | 0.039 | 0.049 | 0.048 | 0.056 | 0.050 | 0.048 | 0.021 | 0.060 | 0.010 |
| 0.3 | -0.001 | 20 | 0.0 | 0.051 | 0.007 | 0.025 | 0.040 | 0.076 | 0.011 | 0.019 | 0.021 | 0.009 | 0.002 | 0.007 | 0.025 | 0.000 |
| 0.3 | 0.135 | 20 | 0.2 |  | 0.020 | 0.071 | 0.045 | 0.096 | 0.143 | 0.085 | 0.061 | 0.081 | 0.076 | 0.020 | 0.071 | 0.001 |
| 0.3 | 0.184 | 20 | 0.4 |  | 0.047 | 0.141 | 0.107 | 0.169 | 0.209 | 0.171 | 0.126 | 0.144 | 0.145 | 0.047 | 0.141 | 0.012 |
| 0.3 | 0.234 | 20 | 0.6 |  | 0.087 | 0.098 | 0.173 | 0.160 | 0.204 | 0.193 | 0.184 | 0.197 | 0.193 | 0.087 | 0.192 | 0.043 |
| 0.3 | 0.239 | 20 | 0.8 |  | 0.108 | 0.137 | 0.196 | 0.191 | 0.228 | 0.215 | 0.203 | 0.211 | 0.215 | 0.108 | 0.207 | 0.076 |
| 0.3 | -0.007 | 50 | 0.0 | 0.058 | 0.002 | 0.016 | 0.041 | 0.076 | 0.013 | 0.025 | 0.028 | 0.011 | 0.004 | 0.002 | 0.016 | 0.000 |
| 0.3 | 0.134 | 50 | 0.2 |  | 0.034 | 0.145 | 0.105 | 0.166 | 0.236 | 0.176 | 0.184 | 0.213 | 0.231 | 0.034 | 0.145 | 0.006 |
| 0.3 | 0.194 | 50 | 0.4 |  | 0.098 | 0.296 | 0.251 | 0.335 | 0.405 | 0.355 | 0.308 | 0.358 | 0.359 | 0.098 | 0.296 | 0.034 |
| 0.3 | 0.238 | 50 | 0.6 |  | 0.196 | 0.230 | 0.349 | 0.339 | 0.393 | 0.375 | 0.385 | 0.400 | 0.411 | 0.196 | 0.384 | 0.119 |
| 0.3 | 0.262 | 50 | 0.8 |  | 0.266 | 0.311 | 0.401 | 0.389 | 0.427 | 0.417 | 0.413 | 0.414 | 0.417 | 0.266 | 0.420 | 0.186 |
| 0.3 | 0.001 | 80 | 0.0 | 0.051 | 0.001 | 0.017 | 0.032 | 0.057 | 0.007 | 0.016 | 0.027 | 0.023 | 0.016 | 0.001 | 0.017 | 0.000 |
| 0.3 | 0.128 | 80 | 0.2 |  | 0.057 | 0.209 | 0.144 | 0.256 | 0.348 | 0.271 | 0.258 | 0.319 | 0.343 | 0.057 | 0.209 | 0.008 |
| 0.3 | 0.187 | 80 | 0.4 |  | 0.157 | 0.370 | 0.341 | 0.462 | 0.553 | 0.493 | 0.412 | 0.468 | 0.497 | 0.157 | 0.370 | 0.067 |
| 0.3 | 0.234 | 80 | 0.6 |  | 0.281 | 0.337 | 0.486 | 0.460 | 0.532 | 0.524 | 0.544 | 0.565 | 0.570 | 0.281 | 0.516 | 0.174 |
| 0.3 | 0.270 | 80 | 0.8 |  | 0.426 | 0.485 | 0.569 | 0.556 | 0.598 | 0.590 | 0.580 | 0.588 | 0.594 | 0.426 | 0.578 | 0.339 |
| 0.3 | 0.008 | 110 | 0.0 | 0.062 | 0.003 | 0.016 | 0.027 | 0.064 | 0.009 | 0.024 | 0.033 | 0.024 | 0.022 | 0.003 | 0.016 | 0.000 |
| 0.3 | 0.133 | 110 | 0.2 |  | 0.087 | 0.288 | 0.228 | 0.333 | 0.435 | 0.367 | 0.324 | 0.426 | 0.452 | 0.087 | 0.288 | 0.019 |
| 0.3 | 0.185 | 110 | 0.4 |  | 0.214 | 0.511 | 0.466 | 0.590 | 0.650 | 0.598 | 0.545 | 0.610 | 0.615 | 0.214 | 0.511 | 0.094 |
| 0.3 | 0.229 | 110 | 0.6 |  | 0.390 | 0.447 | 0.596 | 0.588 | 0.657 | 0.647 | 0.659 | 0.687 | 0.698 | 0.390 | 0.623 | 0.265 |
| 0.3 | 0.265 | 110 | 0.8 |  | 0.534 | 0.586 | 0.685 | 0.674 | 0.710 | 0.700 | 0.694 | 0.698 | 0.713 | 0.534 | 0.692 | 0.450 |
| 0.5 | 0.006 | 20 | 0.0 | 0.053 | 0.005 | 0.014 | 0.035 | 0.066 | 0.010 | 0.021 | 0.022 | 0.010 | 0.008 | 0.005 | 0.014 | 0.001 |
| 0.5 | 0.212 | 20 | 0.2 |  | 0.039 | 0.136 | 0.095 | 0.151 | 0.222 | 0.174 | 0.120 | 0.151 | 0.176 | 0.039 | 0.136 | 0.008 |
| 0.5 | 0.328 | 20 | 0.4 |  | 0.098 | 0.260 | 0.219 | 0.318 | 0.394 | 0.344 | 0.267 | 0.291 | 0.308 | 0.098 | 0.260 | 0.036 |
| 0.5 | 0.374 | 20 | 0.6 |  | 0.189 | 0.231 | 0.365 | 0.349 | 0.408 | 0.386 | 0.368 | 0.396 | 0.402 | 0.189 | 0.396 | 0.102 |
| 0.5 | 0.445 | 20 | 0.8 |  | 0.327 | 0.356 | 0.441 | 0.437 | 0.488 | 0.471 | 0.458 | 0.462 | 0.458 | 0.327 | 0.469 | 0.233 |
| 0.5 | 0.012 | 50 | 0.0 | 0.053 | 0.002 | 0.012 | 0.037 | 0.060 | 0.013 | 0.024 | 0.018 | 0.019 | 0.016 | 0.002 | 0.012 | 0.000 |
| 0.5 | 0.228 | 50 | 0.2 |  | 0.134 | 0.351 | 0.290 | 0.423 | 0.541 | 0.479 | 0.385 | 0.454 | 0.512 | 0.134 | 0.351 | 0.030 |
| 0.5 | 0.316 | 50 | 0.4 |  | 0.278 | 0.575 | 0.545 | 0.652 | 0.725 | 0.692 | 0.584 | 0.621 | 0.648 | 0.278 | 0.575 | 0.145 |
| 0.5 | 0.371 | 50 | 0.6 |  | 0.467 | 0.543 | 0.695 | 0.675 | 0.739 | 0.719 | 0.714 | 0.737 | 0.746 | 0.467 | 0.714 | 0.321 |
| 0.5 | 0.431 | 50 | 0.8 |  | 0.641 | 0.680 | 0.767 | 0.752 | 0.797 | 0.789 | 0.777 | 0.787 | 0.792 | 0.641 | 0.777 | 0.556 |
| 0.5 | 0.001 | 80 | 0.0 | 0.044 | 0.003 | 0.014 | 0.029 | 0.056 | 0.010 | 0.023 | 0.023 | 0.019 | 0.008 | 0.003 | 0.014 | 0.000 |
| 0.5 | 0.223 | 80 | 0.2 |  | 0.205 | 0.508 | 0.445 | 0.622 | 0.725 | 0.682 | 0.530 | 0.636 | 0.667 | 0.205 | 0.508 | 0.054 |
| 0.5 | 0.317 | 80 | 0.4 |  | 0.481 | 0.774 | 0.760 | 0.855 | 0.902 | 0.880 | 0.786 | 0.821 | 0.825 | 0.481 | 0.774 | 0.300 |
| 0.5 | 0.378 | 80 | 0.6 |  | 0.669 | 0.761 | 0.875 | 0.866 | 0.901 | 0.893 | 0.887 | 0.913 | 0.912 | 0.669 | 0.873 | 0.547 |
| 0.5 | 0.435 | 80 | 0.8 |  | 0.816 | 0.863 | 0.913 | 0.912 | 0.929 | 0.923 | 0.909 | 0.926 | 0.928 | 0.816 | 0.912 | 0.760 |
| 0.5 | -0.006 | 110 | 0.0 | 0.045 | 0.002 | 0.015 | 0.039 | 0.057 | 0.009 | 0.022 | 0.029 | 0.022 | 0.019 | 0.002 | 0.015 | 0.000 |
| 0.5 | 0.218 | 110 | 0.2 |  | 0.267 | 0.631 | 0.608 | 0.750 | 0.840 | 0.794 | 0.639 | 0.733 | 0.758 | 0.267 | 0.631 | 0.098 |
| 0.5 | 0.309 | 110 | 0.4 |  | 0.630 | 0.872 | 0.875 | 0.935 | 0.960 | 0.946 | 0.847 | 0.869 | 0.874 | 0.630 | 0.872 | 0.467 |
| 0.5 | 0.377 | 110 | 0.6 |  | 0.806 | 0.877 | 0.929 | 0.934 | 0.949 | 0.943 | 0.948 | 0.959 | 0.964 | 0.806 | 0.936 | 0.738 |
| 0.5 | 0.432 | 110 | 0.8 |  | 0.913 | 0.933 | 0.959 | 0.961 | 0.971 | 0.969 | 0.965 | 0.969 | 0.972 | 0.913 | 0.961 | 0.875 |
| 0.7 | -0.007 | 20 | 0.0 | 0.057 | 0.004 | 0.010 | 0.034 | 0.058 | 0.011 | 0.019 | 0.014 | 0.006 | 0.001 | 0.004 | 0.010 | 0.000 |
| 0.7 | 0.296 | 20 | 0.2 |  | 0.075 | 0.283 | 0.217 | 0.329 | 0.435 | 0.359 | 0.202 | 0.281 | 0.302 | 0.075 | 0.283 | 0.012 |
| 0.7 | 0.442 | 20 | 0.4 |  | 0.225 | 0.497 | 0.455 | 0.567 | 0.640 | 0.586 | 0.459 | 0.502 | 0.518 | 0.225 | 0.497 | 0.091 |
| 0.7 | 0.524 | 20 | 0.6 |  | 0.357 | 0.416 | 0.564 | 0.546 | 0.627 | 0.603 | 0.554 | 0.584 | 0.582 | 0.357 | 0.598 | 0.253 |
| 0.7 | 0.586 | 20 | 0.8 |  | 0.505 | 0.559 | 0.638 | 0.627 | 0.674 | 0.665 | 0.654 | 0.661 | 0.668 | 0.505 | 0.658 | 0.421 |
| 0.7 | -0.005 | 50 | 0.0 | 0.040 | 0.003 | 0.015 | 0.030 | 0.057 | 0.012 | 0.020 | 0.025 | 0.022 | 0.006 | 0.003 | 0.015 | 0.000 |
| 0.7 | 0.294 | 50 | 0.2 |  | 0.258 | 0.619 | 0.580 | 0.721 | 0.828 | 0.770 | 0.551 | 0.645 | 0.686 | 0.258 | 0.619 | 0.087 |
| 0.7 | 0.436 | 50 | 0.4 |  | 0.534 | 0.846 | 0.839 | 0.903 | 0.939 | 0.925 | 0.775 | 0.810 | 0.817 | 0.534 | 0.846 | 0.358 |
| 0.7 | 0.527 | 50 | 0.6 |  | 0.783 | 0.859 | 0.931 | 0.929 | 0.957 | 0.953 | 0.919 | 0.933 | 0.941 | 0.783 | 0.933 | 0.688 |
| 0.7 | 0.595 | 50 | 0.8 |  | 0.888 | 0.915 | 0.955 | 0.954 | 0.961 | 0.958 | 0.952 | 0.959 | 0.958 | 0.888 | 0.949 | 0.848 |
| 0.7 | -0.006 | 80 | 0.0 | 0.045 | 0.002 | 0.011 | 0.028 | 0.059 | 0.009 | 0.012 | 0.028 | 0.016 | 0.012 | 0.002 | 0.011 | 0.000 |
| 0.7 | 0.310 | 80 | 0.2 |  | 0.411 | 0.810 | 0.798 | 0.895 | 0.945 | 0.923 | 0.710 | 0.744 | 0.756 | 0.411 | 0.810 | 0.221 |
| 0.7 | 0.435 | 80 | 0.4 |  | 0.782 | 0.963 | 0.959 | 0.983 | 0.989 | 0.986 | 0.879 | 0.891 | 0.893 | 0.782 | 0.963 | 0.666 |
| 0.7 | 0.529 | 80 | 0.6 |  | 0.935 | 0.966 | 0.992 | 0.989 | 0.992 | 0.991 | 0.984 | 0.987 | 0.986 | 0.935 | 0.990 | 0.895 |
| 0.7 | 0.601 | 80 | 0.8 |  | 0.986 | 0.991 | 0.997 | 0.997 | 0.998 | 0.998 | 0.995 | 0.998 | 0.998 | 0.986 | 0.996 | 0.972 |
| 0.7 | -0.003 | 110 | 0.0 | 0.058 | 0.000 | 0.012 | 0.042 | 0.067 | 0.016 | 0.025 | 0.045 | 0.026 | 0.014 | 0.000 | 0.012 | 0.000 |
| 0.7 | 0.313 | 110 | 0.2 |  | 0.584 | 0.908 | 0.913 | 0.966 | 0.986 | 0.979 | 0.770 | 0.793 | 0.786 | 0.584 | 0.908 | 0.401 |
| 0.7 | 0.432 | 110 | 0.4 |  | 0.913 | 0.992 | 0.993 | 0.999 | 0.999 | 0.999 | 0.900 | 0.900 | 0.902 | 0.913 | 0.992 | 0.839 |
| 0.7 | 0.514 | 110 | 0.6 |  | 0.980 | 0.993 | 0.997 | 0.997 | 0.999 | 0.999 | 0.997 | 0.996 | 0.997 | 0.980 | 0.998 | 0.970 |
| 0.7 | 0.591 | 110 | 0.8 |  | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 |

# Notes

The script is available on <https://github.com/oscci/MinSigVar>.

# References

Lazic, Stanley E. 2021. “Why Multiple Hypothesis Test Corrections Provide Poor Control of False Positives in the Real World.” *arXiv:2108.04752 [q-Bio, Stat]*, August. <http://arxiv.org/abs/2108.04752>.

Moher, David, Sally Hopewell, Kenneth F. Schulz, Victor Montori, Peter C. Gøtzsche, P. J. Devereaux, Diana Elbourne, Matthias Egger, and Douglas G. Altman. 2010. “CONSORT 2010 Explanation and Elaboration: Updated Guidelines for Reporting Parallel Group Randomised Trials.” *BMJ (Clinical Research Ed.)* 340 (March): c869. <https://doi.org/10.1136/bmj.c869>.

R Core Team. 2020. “R: A Language and Environment for Statistical Computing.” R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.