Leibniz
Universität Hannover

Reports of the Institute of Biostatistics

No 09 / 2007

Leibniz University of Hannover

Natural Sciences Faculty

Approximate one-sided two-sample confidence limits for the comparison of a treatment versus a near-zero spontaneous rate in control

Sill, M.

1 Introduction

Non-inferiority or equivalence studies are commonly used in the evaluation of new treatments. E.g new drugs, vaccines or laboratory assays are promised to offer better administration, lower costs or diagnostic safety compared to the standard treatment. To ensure whether effectiveness of a new treatment is at least as good as compared to an active control or standard treatment, clinical equivalence or non-inferiority trials are conducted. In this report the evaluation of two arm non-inferiority trials with binary endpoints is discussed. Particularly the emphasis is put on situations with proportions near zero which are common in toxicology. For the comparison of proportions in two-by-two tables three measures of dissimilarity can be used: the difference of proportions, the relative risk and the odds ratio. Non-inferiority hypothesis can be formulated for all three measures and corresponding testing procedures do exist (Wellek, 2005). For the difference of proportions the one-sided non-inferiority hypotheses are: $H_0 : \pi_1 - \pi_0 < -\delta$ and $H_A : \pi_1 - \pi_0 \geq -\delta$. The $-\delta$ is the pre-specified non-inferiority margin. In order to derive an inference decision about these hypotheses confidence intervals are a reasonable tool. Confidence intervals have the advantage to combine statistical significance with the idea of clinical relevance. Moreover in terms of difference of proportions the bounds of the interval are on the scale of the clinical endpoint and allow for immediate interpretation. The intervals can be presented in a figure together with the equivalence margin. Here we emphasis on approximate one-sided confidence intervals for the difference of proportions. Asymptotic intervals are based on the normal approximation, that can fail if sample sizes are low and if the data is skewed. The latter occurs in situations when measured proportions are close to zero or one. To overcome this two problems adjusted intervals based on pseudo observations are investigated. The idea of this adjustments is to add some successes and failures to the observed proportions resulting in a shift of the mid point of the interval towards 0.5 and an increased variance (Agresti and Caffo, 2000). By shifting the mid point towards 0.5 possible skewed binomial data is aligned to the normal distribution. This report presents some simulation studies that reveal the behavior of adjusted approximate confidence intervals for situations where measured proportions are close to zero.

2 Methods

In a randomized clinical trial a new therapy or compound is investigated regarding its non-inferiority to a standard treatment. The outcome of the study is binary: y_{j0} are the $j = 1, \cdots, n_0$ observations in the control group with sample size n_0 and y_{j1} are the observations in the treatment group with sample size n_0. Assuming y_{j0} and y_{j1} to follow two independent binomial distributions $Bin(n_0, \pi_0)$ and $Bin(n_1, \pi_1)$ respectively, the maximum likelihood estimators are $\hat{\pi}_0 = \sum_{j=1}^{n_0} y_{j0}/n_0$ and $\hat{\pi}_0 = \sum_{j=1}^{n_1} y_{j1}/n_1$. The lower bound of the Wald interval for the difference of proportions is:

$$\hat{\pi_1} - \hat{\pi}_0 - z_{1-\alpha}\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_0(1-\hat{\pi}_0)}{n_0}} \tag{1}$$

Following the ideas of Agresti and Caffo (2000), to add pseudo observations to the point estimates, the estimators for the Add-1 interval adding one pseudo observation to each proportion are $\tilde{\pi}_0 = (\sum_{j=1}^{n_0} y_{j0}) + 0.5/n_0 + 1$ and $\tilde{\pi}_1 = (\sum_{j=1}^{n_1} y_{j1}) + 0.5/n_1 + 1$. The lower bound of the Add-1 interval is:

$$\tilde{\pi}_1 - \tilde{\pi}_0 - z_{1-\alpha}\sqrt{\frac{\tilde{\pi}_1(1-\tilde{\pi}_1)}{n_1+1} + \frac{\tilde{\pi}_0(1-\tilde{\pi}_0)}{n_0+1}} \tag{2}$$

Another possibility is to add two pseudo observations to each proportion. Then the estimators are $\bar{\pi}_0 = (\sum_{j=1}^{n_0} y_{j0}) + 1/n_0 + 2$ and $\bar{\pi}_1 = (\sum_{j=1}^{n_1} y_{j1}) + 1/n_1 + 2$. The lower bound of this so called Add-2 interval is:

$$\bar{\pi}_1 - \bar{\pi}_0 - z_{1-\alpha}\sqrt{\frac{\bar{\pi}_1(1-\bar{\pi}_1)}{n_1+2} + \frac{\bar{\pi}_0(1-\bar{\pi}_0)}{n_0+2}} \tag{3}$$

The idea for this intervals for the difference of proportions are derived from the Wilson Score interval for a single proportion (Wilson, 1927). Based on the Wilson Score interval Newcombe (1998) proposed an interval for the difference of proportions:

$$\hat{\pi}_1 - \hat{\pi}_0 - z_{1-\alpha}\sqrt{l_0(1-l_0)/n_0 + u_1(1-u_1)/n_1} \tag{4}$$

l_0 is the lower bound of the Wilson Score interval for $\hat{\pi}_0$ and u_1 the upper bound of the Wilson Score interval for $\hat{\pi}_1$. The formula for the two-sided Score interval is:

$$\left[\frac{\hat{\pi}_i + \frac{1}{2n_i}z^2_{1-\alpha/2} \pm z_{1-\alpha/2}\sqrt{\frac{1}{n_i}\left[\hat{\pi}_i\left(1-\hat{\pi}_i\right) + \frac{1}{4n_i}z^2_{1-\alpha/2}\right]}}{1 + \frac{1}{n_i}z^2_{1-\alpha/2}}\right]$$

where $i = 0, 1$.

3 Simulation

To investigate the one-sided coverage probabilities of the intervals presented in Section 2 a simulation study was performed. Therefore the whole parameter space with $\pi_0 = 0.01, \cdots, 0.99$ and $\pi_1 = 0.01, \cdots, 0.99$ was considered for investigation. For each possible combination of π_0 and π_1 10000 random samples were generated from a binomial distribution and coverage probability, namely the probability that the interval includes the true difference of proportions $\pi_0 - \pi_1$, was calculated. The results of this simulations are presented Figure 1 to 11. A coverage probability close to the nominal level is desired. The main focus is on small proportions, e.g. a parameter space with $\pi_0 = 0.01, \cdots, 0.20$ and $\pi_1 = 0.01, \cdots, 0.20$.

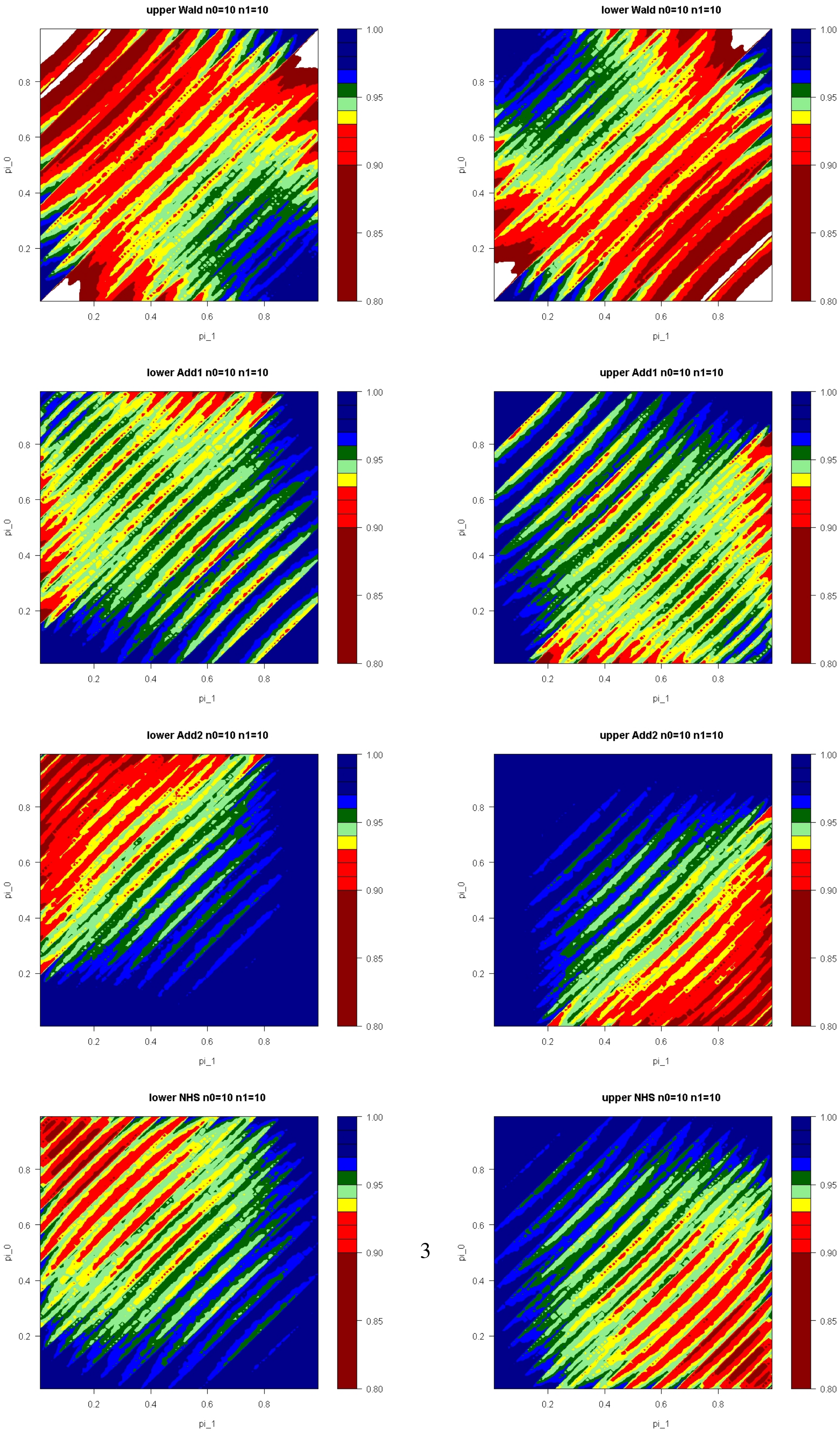

Figure 1: Coverage probabilities for n0=10 and n1=10

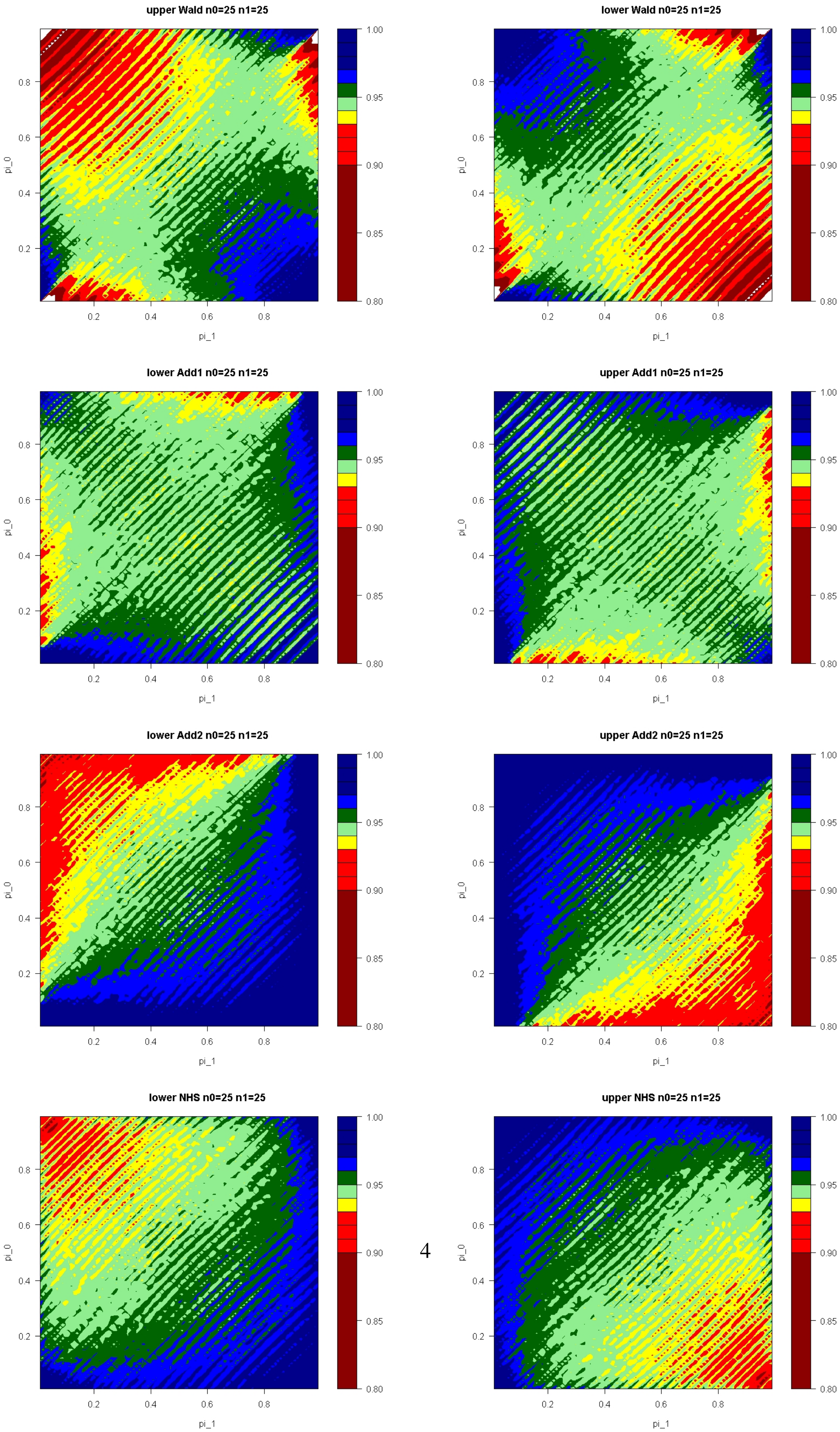

Figure 2: Coverage probabilities for n0=25 and n1=25

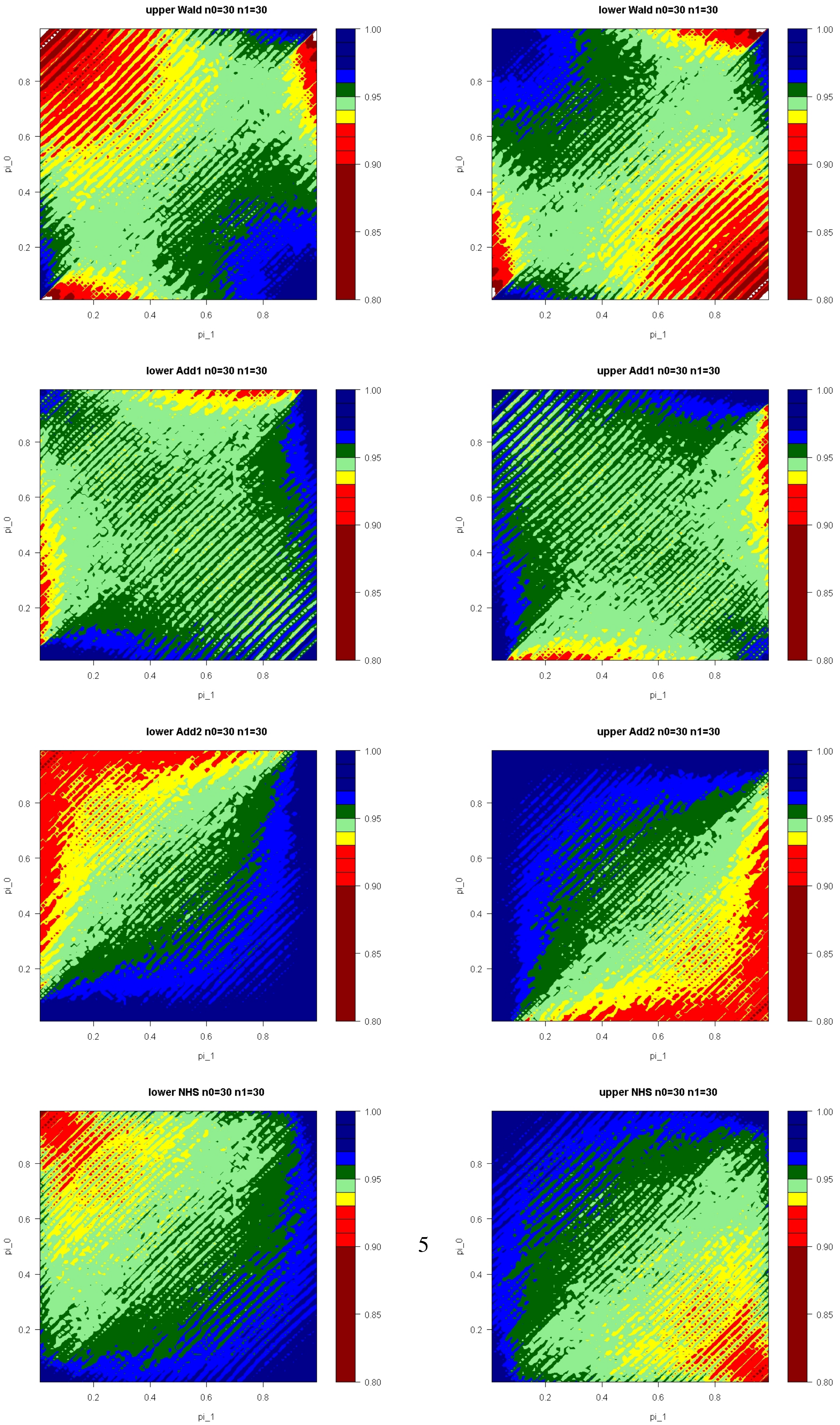

Figure 3: Coverage probabilities for n0=30 and n1=30

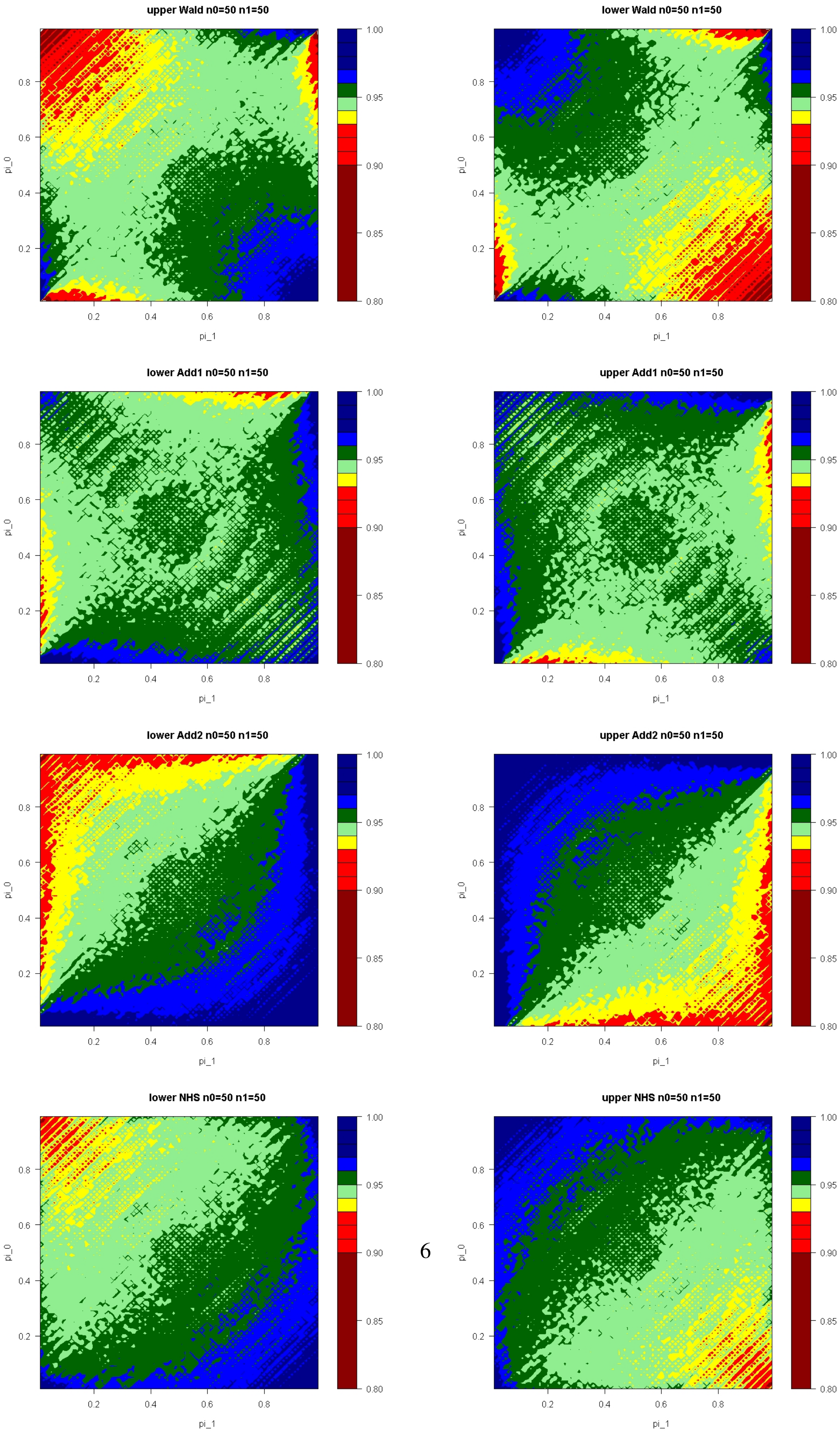

Figure 4: Coverage probabilities for n0=50 and n1=50

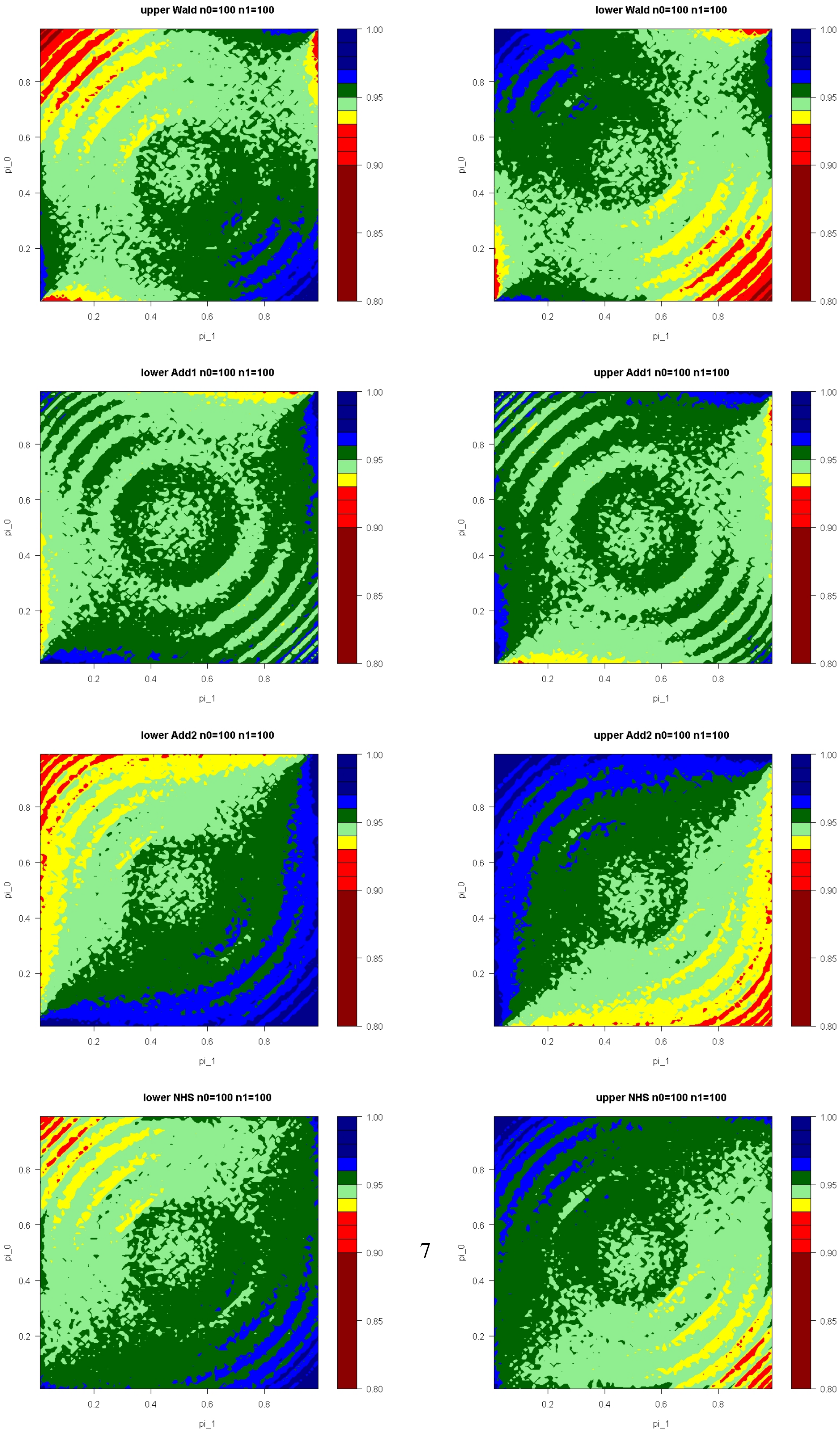

Figure 5: Coverage probabilities for n0=100 and n1=100

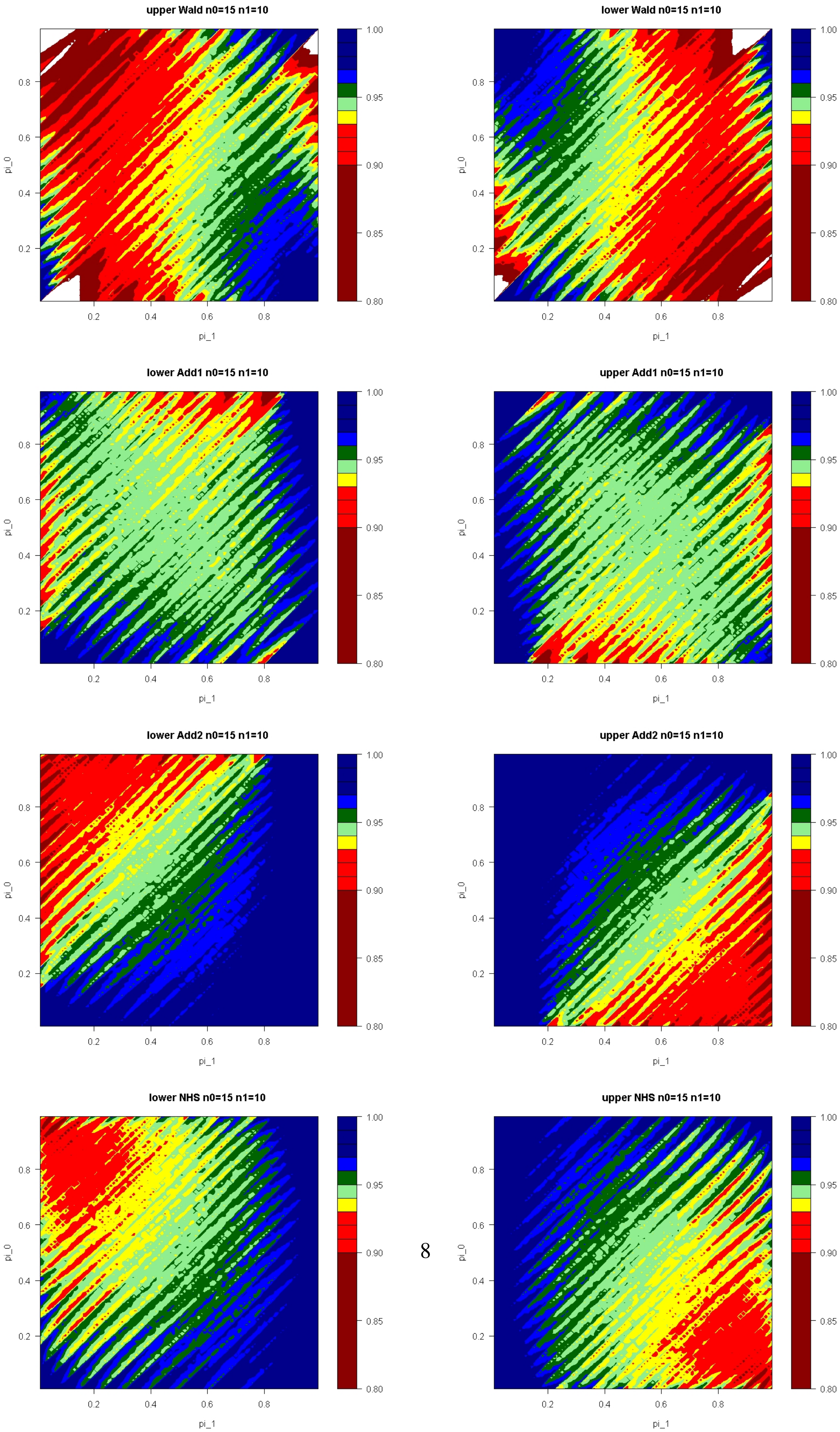

Figure 6: Coverage probabilities for n0=15 and n1=10

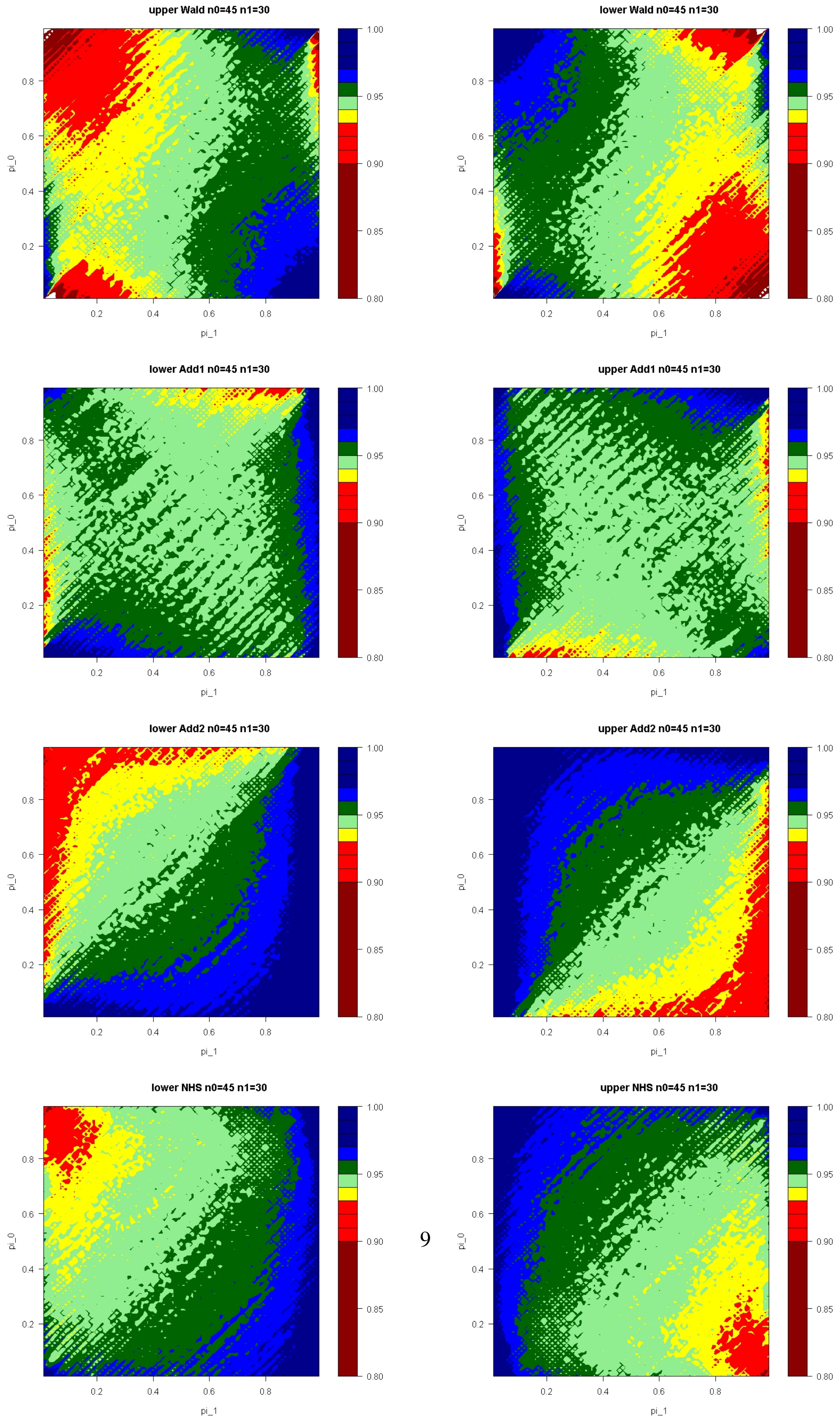

Figure 7: Coverage probabilities for n0=45 and n1=30

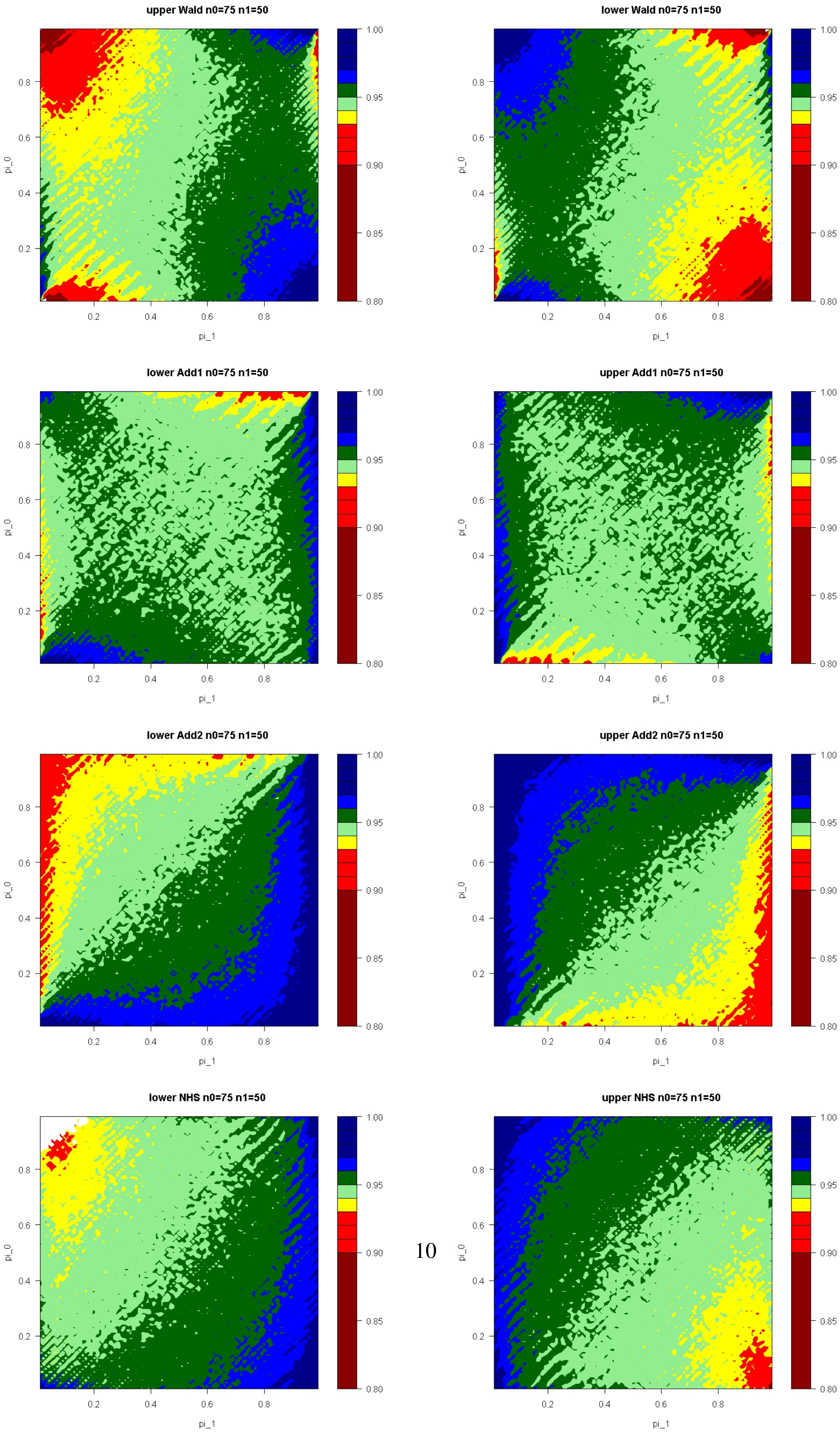

Figure 8: Coverage probabilities for n0=75 and n1=50

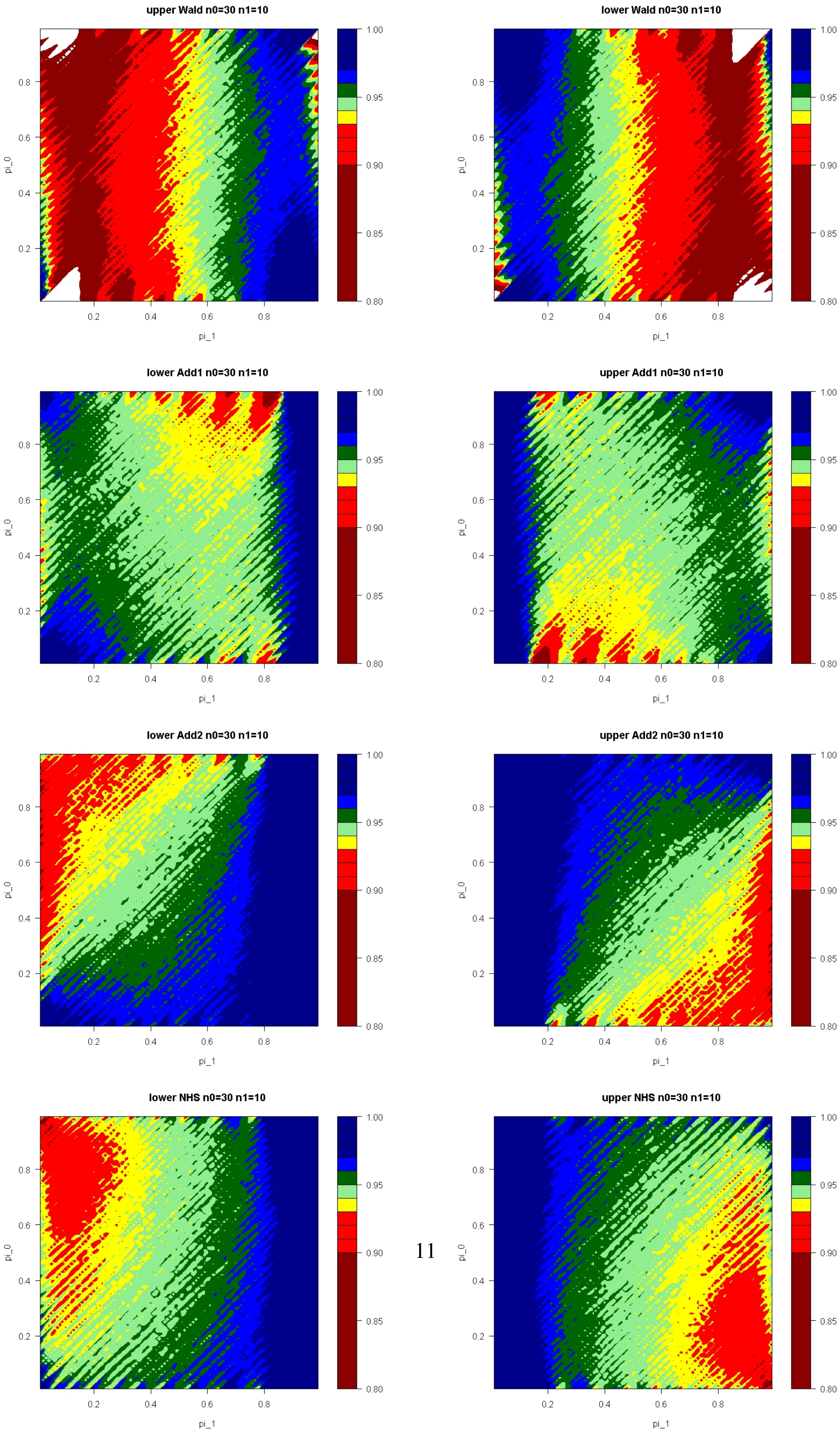

Figure 9: Coverage probabilities for n0=30 and n1=10

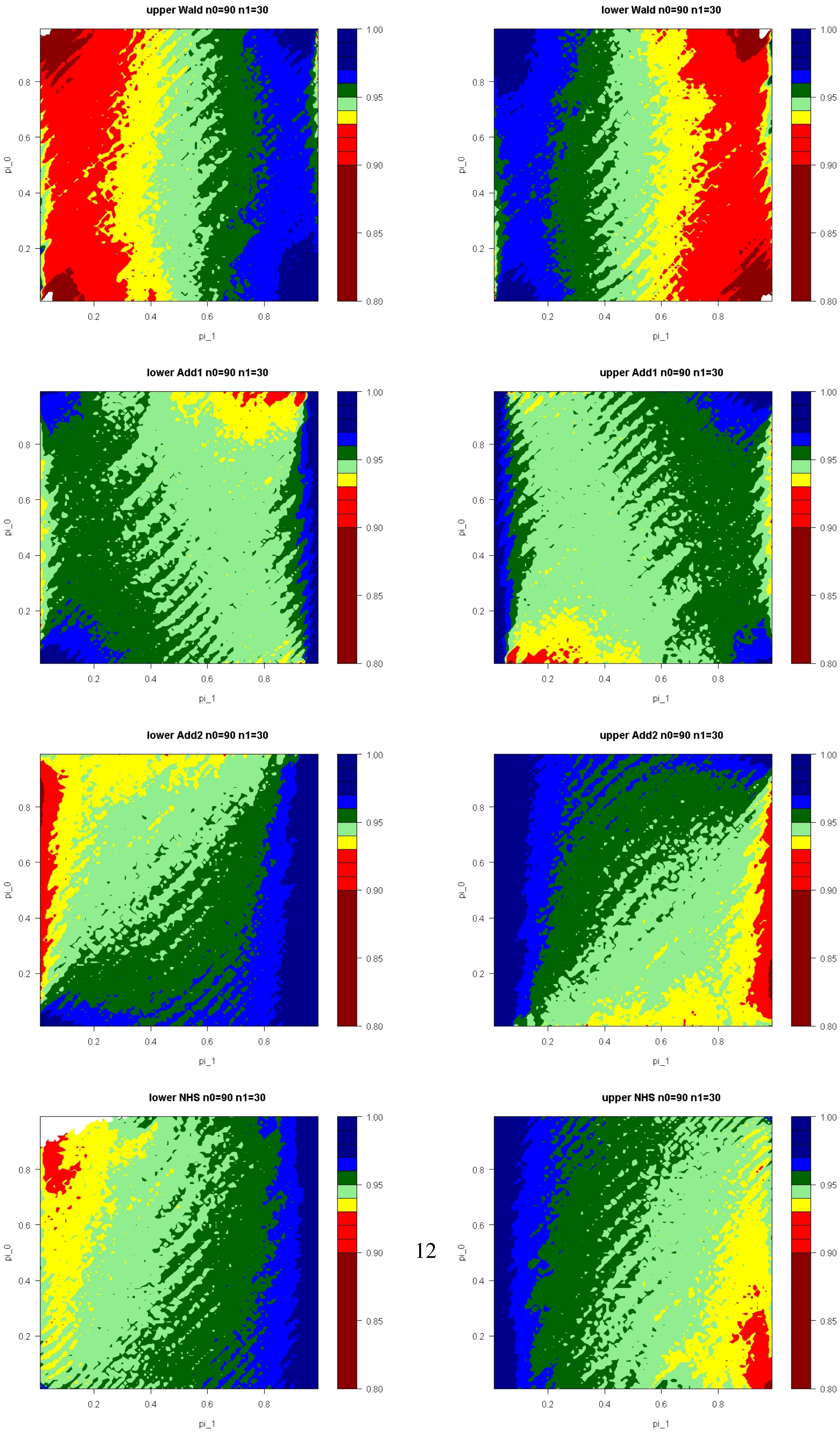

Figure 10: Coverage probabilities for n0=90 and n1=30

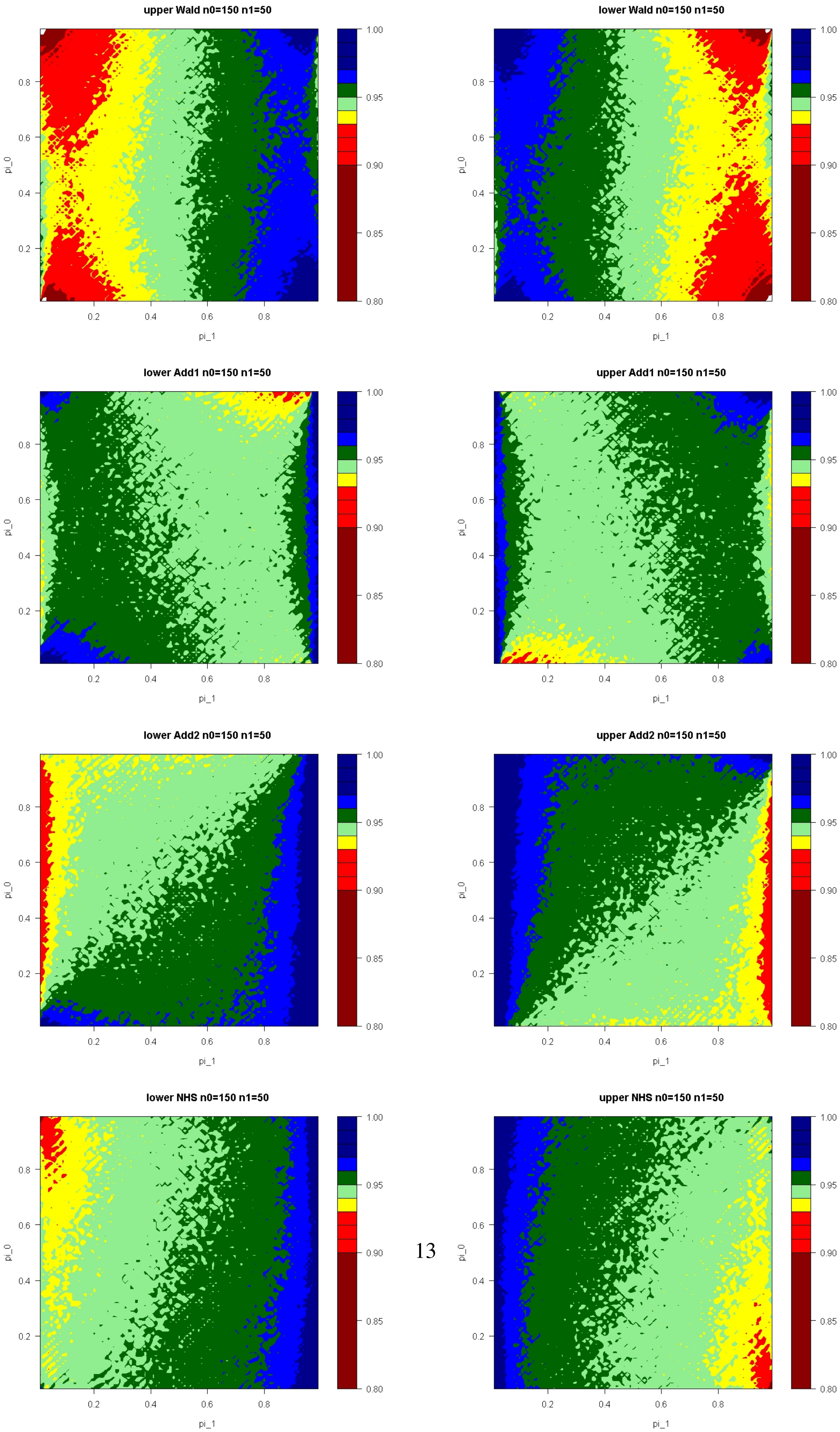

Figure 11: Coverage probabilities for n0=150 and n1=50

4 Discussion

As known, the Wald interval has too liberal coverage probability when sample sizes are small and in cases when the proportions are close to zero or one. The Add-1 interval has reasonable coverage probabilities for moderate and large sample sizes but for proportions close to zero or one, still strong variations from the nominal level occur. For the Add-2 interval this variations are stronger and even for large sample sizes present. Considering the main focus of lower one-sided confidence intervals for cases were both proportions are small, the NHS shows a well performance. In cases were one proportion is close to one and the other close to zero the NHS interval is also problematic. In cases of unbalanced sample sizes the regions of reasonable, liberal and conservative coverage probability are shifted depending on the difference between n_0 and n_1. For the evaluation of toxicological equivalence trails were small proportions occur we recommend to use the NHS interval. In the future further approximate intervals should be investigated. All investigated intervals are implemented for the R software environment in the binMto package.

References

[1] Agresti, A. and Caffo, B. (2000). Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *American Statistician* 54:280-288.

[2] Brown, L. Li, X. (2005). Confidence intervals for two sample binomial distributions. *Journal of Statistical Planning and Inference* 130:359-375.

[3] Chan, SF. (2003). Proving non-inferiority or equivalence of two treatments with dichotomous endpoints using exact methods. *Stat Methods Med Res* 12:37.

[4] Newcombe, RG. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Statistics in Medicine* 17:873-890.

[5] Wellek, S. (2005). Statistical methods for the analysis of two-arm non-inferiority trials with binary outcomes. *Biometrical Journal* 47:48-61.

[6] Wilson, EB. (1927). Probable inference, the law of succession and statistical inference. *Journal of the American Statistical Association* 22:209-212.