# A Case-Only Approach to Estimating the RR in the CR-TND

Suzanne Dufault

March 2, 2018

## Contents



# 1 Set Up

Existing methods for estimating the RR from CR-TND data require the use of test-positive (case) and test-negative (control) counts from each cluster. Because the treatment is randomized, it is possible to estimate the intervention RR using only the test-positive counts. This approach is explored here.

Suppose there are $2m$ clusters in the study with $m$ assigned to the intervention and the remaining $m$ to control. Let $j \in I$ specify a cluster $j$ in the intervention arm and $j \in C$ be a cluster $j$ in the control arm. In each cluster, the number of test-positive individuals $n_{D_j}$ is recorded. A proposed test-statistic is then:

$$T = \sum_{j \in I} n_{D_j} - \sum_{j \in C} n_{D_j} \tag{1}$$

$$= n_{D_I} - n_{D_C} \tag{2}$$

# 2 At the Null

## 2.1 Total

At the null, we expect half of all cases ($n_D$) to fall in each arm:

$$E[n_{D_I}] = E[n_{D_C}] = \frac{n_D}{2} \tag{3}$$

Hence, at the null, $T = 0$

$$E[T] = E[n_{D_I}] - E[n_{D_C}] \tag{4}$$

$$= E[n_{D_I}] - (n_D - E[n_{D_I}]) \tag{5}$$

$$= 2E[n_{D_I}] - n_D \tag{6}$$

$$= 2\frac{n_D}{2} - n_D \tag{7}$$

$$= 0 \tag{8}$$

## 2.2 Variance

The variance of $T$ is estimated as follows:

$$Var(T) = Var(2 \times [n_{D_I}] - n_D) \tag{9}$$
$$= 2^2 Var(n_{D_I}) \tag{10}$$
$$= 2^2 \frac{(2m)^2}{m} \frac{m}{2m-1} \frac{\sigma^2}{m} \tag{11}$$
$$= \frac{16m}{2m-1} \sigma^2 \tag{12}$$

where $\sigma^2$ is the population variance of the $2m$ test-positive counts.

## 2.3 Standardized Statistic

Hence, the standardized test statistic is $\frac{T}{\sqrt{\frac{16m}{2m-1}\sigma^2}} \sim N(0,1)$.

# 3 Recovering the Intervention RR

When an intervention is applied to reduce the number of cases in the intervention arm, we would like to estimate the intervention effect through the RR ($\lambda$). In the test-positive setting, this can be accomplished through a simple ratio of the sums in the two arms:

$$\lambda = \frac{n_{D_I}}{n_{D_C}} \tag{13}$$
$$= \frac{n_{D_I}}{n_D - n_{D_I}} \tag{14}$$

## 3.1 Variance of the Intervention RR

In order to estimate the variance of the intervention RR $\lambda$, we can use the delta method. We know the properties of $n_{D_I}$, and $\lambda$ is a simple transformation of $n_{D_I}$.

Hence if $x = n_{D_I}$:

$$g(x) = \frac{x}{n_D - x} \tag{15}$$
$$g'(x) = \frac{n_D}{(n_D - x)^2} \tag{16}$$

Implying:

$$\hat{\lambda} \approx N\left(\lambda, \left[\left(\frac{n_D}{(n_D - x)^2}\right)^2 Var(x)\right]\right) \tag{17}$$
$$= N\left(\lambda, \left[\left(\frac{n_D}{(n_D - n_{D_I})^2}\right)^2 Var(n_{D_I})\right]\right) \tag{18}$$
$$= N\left(\lambda, \left[\left(\frac{n_D}{(n_D - n_{D_I})^2}\right)^2 \frac{4m^2}{m} \frac{m}{2m-1} \frac{\sigma^2}{m}\right]\right) \tag{19}$$
$$= N\left(\lambda, \left[\left(\frac{n_D}{(n_D - n_{D_I})^2}\right)^2 \frac{4m}{2m-1} \sigma^2\right]\right) \tag{20}$$

# 4 Simulation Work

We know $Var(T)$, but in order to estimate it, we need to replace the population variance $\sigma^2$ with an estimate.

$$\hat{Var}(T) = 2^2 \hat{Var}(n_{D_I}) \tag{21}$$
$$= 4(2m)^2 \frac{1}{2} \frac{s^2}{m} \tag{22}$$
$$= 8ms^2 \tag{23}$$

where $s^2 = \frac{1}{m-1} \sum_{j \in I, j=1}^{m} (n_{D_j} - \bar{n}_{D_J})^2$. Hence, we can compare $T/\sqrt{8ms^2}$ to a T distribution with 2(m-1) degrees of freedom.

The same is true for estimating $Var(\lambda)$.

$$Var(\lambda) = \left( \frac{n_D}{(n_D - n_{D_I})^2} \right)^2 Var(n_{D_I}) \tag{24}$$
$$\hat{Var}(\lambda) = \left( \frac{n_D}{(n_D - n_{D_I})^2} \right)^2 \hat{Var}(n_{D_I}) \tag{25}$$
$$= \left( \frac{n_D}{(n_D - n_{D_I})^2} \right)^2 4m^2 \frac{1}{2} \frac{s^2}{m} \tag{26}$$
$$= \left( \frac{n_D}{(n_D - n_{D_I})^2} \right)^2 2ms^2 \tag{27}$$