DIMACS Technical Report 2004-20
April 2004

SUBJECT ALLOCATION AND STUDY CURTAILMENT FOR FIXED EVENT COMPARATIVE POISSON TRIALS[1]

By

Donald R Hoover
Statistics Department
Rutgers University
New Jersey 08854
Email: drhoover@stat.rutgers.edu

DIMACS is a collaborative project of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, NEC Laboratories America and Telcordia Technologies, as well as affiliate members Avaya Labs, HP Labs, IBM Research and Microsoft Research. DIMACS was founded as an NSF Science and Technology Center.

1. This work is supported by grant EIA 02-05116 from the National Science Foundation.

ABSTRACT

Comparative Poisson trials of prophylactic interventions, such as vaccines, can be lengthy and costly. We evaluate two easily implemented approaches to reduce numbers of disease cases and person years of follow up (N_{u+t}) for comparative Poisson trials with fixed numbers of cases (T); i) altering k the portion of N_{u+t} allocated to treatment, and ii) curtailed stopping before T cases if numbers of cases in the treatment or control group indicate that H_o has already been rejected or will not be rejected at T cases. Normal and arcsine approximations as well as discrete exact tests are evaluated. For studies not stopped early, allocating about $1/(1+\sqrt{r_a})$ of person years to treatment roughly minimizes T needed for given size and power (where r_a is the alternative hypothesized relative disease incidence in treated subjects used to power the study). This reduces T moderately Vs equal allocation (k=0.5); by 2-3 cases in our examples with exact tests. However, the common practice of allocating k=0.5 of the person years to treatment may be the overall best strategy to minimize N_{u+t} for studies that are not stopped early. For studies analyzed by exact test and planned with a one sided α ranging from 0.005 to 0.025, β from 0.1 to 0.2 and r_a from 0.2 to 0.5, curtailed stopping reduces both number of disease cases and N_{u+t} by 6% to 40% depending on true treatment benefit. With curtailed stopping, allocating k=0.5 person years to treatment approximately minimizes numbers of cases and person years under most conditions, although k as large as 0.6 often performs comparably well. If a specific localized k is selected to minimize disease cases for a study analyzed by exact test, the study may be underpowered should the final allocation deviate even slightly from that k when it is conducted. □

Key Words – Comparative Poisson, Curtailed Stopping, Power, Sample Size, Subject Allocation

1. INTRODUCTION

Studies of prophylactic interventions against uncommon diseases compare incidence between groups of treated and untreated persons. Often incidence is Poisson with rates λ_o per person year for untreated and $r^*\lambda_o$ for treated subjects or binomial with a very low probability that is closely approximated by a Poisson[1,2]. The hypothesis test is H_o: $r^* = 1$ Vs H_a: $r^* = r_a$ for a specified $r_a<1$. We define the person years of follow up as N_u for untreated subjects, N_t for treated subjects and $N_{u+t} = N_u + N_t$ total. If the study design involves a Poisson approximation to a binomial with low rates, then N_t , N_u and N_{u+t} would perhaps be persons rather than person-years; we loosely refer to both settings simultaneously in the remainder of the paper. Sometimes N_{u+t} is predetermined[1]. If the study is stopped at a planned N_{u+t} person years (or persons) and comparisons made irrespective of number of disease cases, it will be underpowered if λ_o is overestimated during study planning or unnecessarily long if λ_o is underestimated. The alternate approach studied here which ensures power is to continue follow up (with a person year ratio $k= N_t / N_{u+t}$ maintained) until a predetermined fixed number of disease events", T, occurs in both groups[1]. H_o is rejected if $\leq C$ events are in the treated group for C chosen to maintain a specified Type 1 error. Now N_{u+t} is a random; to prevent the study from lasting too long, it can also be stopped when N_{u+t} reaches a given number, but this reduces power.

Fixed event comparative Poisson trials can be long and costly. For example, a study of a vaccine against early onset Alzheimer's disease in subjects with an untreated incidence of $\lambda_o = 0.0001$ per person year would need considerable follow up to obtain sufficient cases to detect treatment group differences. Thus, minimizing N_{u+t} and/or T may be advantageous. For any given size and power, a study stopped at $T = 20$ Alzheimer's diseases cases might be less expensive and ethically preferable to one stopped at T=30, while a study with $N_{u+t} = 100{,}000$ person years should be cheaper than one with $N_{u+t} =$ 150,000. Two strategies that may reduce N_{u+t} and T are: 1) optimizing k, the portion of person years under treatment[2-7]; and 2) early (or curtailed) stopping once the number of

cases in the treated or untreated group indicates the null hypothesis already has been rejected or accepted[8-10].

Allocation Ratio of Subjects - While often $k = 0.5$ is used to plan studies, other k may result in smaller T and N_{u+t}. For the related problem of comparing binomial proportions between two samples, allocating more subjects to groups with lower rates under H_a often maximizes power,[3,4] particularly for superiority hypotheses[5,6]. Dragalin et. al. (2002)[7] evaluated subject randomization ratios for comparative Poisson designs similar to those considered here; their null hypothesis was H_o: $r^* = r_o$ (for $r_o > 1$) rather than H_o: $r^* = 1$, and found that unequal subject allocation often decreased required T and expected N_{u+t} below that for equal subject allocation[7]. It should be noted, however, that logistic ability of studies to maintain a planned allocation k varies due to random dropout or discovery of ineligibility if subjects are followed for short lengths of time, and also due to differential subject follow up for studies of persons followed indefinitely until the date of analysis. These problems may particularly manifest if one group has a higher dropout rate.

Curtailed Stopping - Blinded studies are conducted until all T events occur. But as many have noted[7-10], if treatment group of cases is known as they occur, studies can be stopped once (i) C+1 cases occur in the treatment group and H_o accepted or (ii) T-C cases occur in the untreated subjects and H_o rejected. Consider a study with k=0.5 allocation and a one sided Type-1 error of α=0.025 that will be stopped at T=72 events in both groups. By the exact test described in Section 2 for α=0.025, H_o is rejected if and only if the number of events in the treated group is C=27 or less. Thus, if 60 cases have occurred in both groups with 28 in treated subjects, by (i) above, H_o can no longer be rejected so there is no need to continue to 72 cases. But if 45 cases are in treated subjects, H_o has already been rejected, since at 72 events at most 27 could be in treated subjects. We now refer to stopping the study at the earliest of (i) and (ii) above as "Curtailed Stopping Time" and to studies that are not stopped early in this fashion as "Uncurtailed Studies".

This paper studies benefits of altering the subject allocation ratio, k, and using curtailed stopping to minimize T and N_{u+t} in comparative Poisson trials planned for fixed T. In

many cases, these depend on the true effect of the treatment r^*, which is unknown so we focus on 3 likely and spanning settings; $r^*=1$ (H_o true), $r^*=r_a$ (specified H_a true) and $r^*=0$ (treatment completely effective). Preliminary formulas and notation are presented in Section 2. Section 3 evaluates selection of k to minimize T in uncurtailed studies. Section 4 studies the impact of curtailed stopping on number of cases and the k which minimizes the expected number of cases. Section 5 evaluates minimization of expected person years of follow up, $E[N_{u+t}]$, in uncurtailed and curtailed designs, and is followed by the Discussion. Appendix A presents the special case of $r_a=0$ while Appendix B describes problems with exact test power estimates when k and T cannot be completely controlled.

2. PRELIMINARIES

For given N_u and N_t, the Poisson parameters for distribution of disease cases in the untreated and treated subjects are $\lambda_u = \lambda_o N_u$ and $\lambda_t = r^* \lambda_o N_t$, respectively. Thus, if X_t and X_u are numbers of events that occur among treated and untreated subjects, respectively, for studies with N_u, N_t and $k= N_t / N_{u+t}$; Przyborski and Wilenski (1939)[11] showed conditional on $X_t + X_u = T$, X_t is binomial with parameters $\{p_t = r^*k/[1 - k(1- r^*)], T\}$. The uniformly most powerful test[12] of $r^*=1$ Vs $r^*= r_a$ in two Poisson distributions conditions on T and tests H_o: $p_t = k$ against H_a: $p_t = r_a k/[1 - k(1-r_a)]$. Formulas to estimate power for $k=0.5$ and $k\neq 0.5$ can be based on exact tests as now described:

Exact Test: For a given T and one sided Type-1 error α, the exact test rejects H_o if X_t is less than or equal to the largest integer C for which $\sum_{i=0}^{C}[T!/(i!(T-i)!)]k^i(1-k)^{T-i} \leq \alpha$. The exact power to detect a reduction to r^* from the treatment is denoted $\Phi^{r^*}(E,T) = \sum_{i=0}^{C}\frac{T!}{i!(T-i)!}\left(\frac{r^*k}{1-k(1-r^*)}\right)^i\left(\frac{1-k}{1-k(1-r^*)}\right)^{T-i}$. The Minimal T for which $\Phi^{r^*}(E,T) \geq (1-\beta)$ under the specified H_a depends on k, r_a, α and β and is denoted here

by $T(E)$. The corresponding rejection region limit also depends on *k*, r_a, α and β and is denoted $C(E)$. We calculated $T(E), \Phi^{r^*}(E,T)$ and $C(E)$ with PASS [13].

Due to discreteness of binomials, actual Exact Tests sizes are < α. While (with fixed *k*, r_a, α and β) exact power usually increases when *T* increases, decreases in exact power when *T* increases are possible. As noted in Appendix B, with other parameters held constant, $T(E)$ fluctuates with *k* including discontinuities that complicate identification of more general patterns. We therefore also use two asymptotic approximations to the binomial that do not have discrete fluctuations to aid in identifying patterns and general rules. While several asymptotic approximations to the binomial have been proposed[14-17], we use here the constrained normal and arcsine approximations which are perhaps among the most commonly used for the purpose of study planning and collectively gave results similar to exact tests in the examples we studied.

A Common Normal Approximation to the sample proportion[14-16] estimates $T(E)$ by

$$T(N) = \left[\frac{Z_\alpha \sqrt{k(1-k)} + Z_\beta \sqrt{r_a k(1-k)}/(1-k(1-r_a)))}{k - r_a k/(1-k(1-r_a))} \right]^2 \qquad (1)$$

where Z_α and Z_β are upper α and β standard normal quantiles. H_o is rejected for $X_t <$ $C(N) = k\ T(N) - Z_\alpha \sqrt{(T\ N)(k\ -)k}$. When $T(N)$ and $C(N)$ are not integers, they are rounded conservatively.

The arcsine approximation to the binomial[1,14,16,17] estimates $T(E)$ by $T(A)$;

$$T(A) = \{(Z_\alpha + Z_\beta)/[4(\text{arcsine}\sqrt{k} - \text{arcsine}\sqrt{r_a k / 1 - k(1-r_a)}]\}^2 \qquad (2)$$

H_o is rejected if $X_t < C(A) = T(A)[\sin(\text{arcsine}(1-k) - Z_\alpha/(2\sqrt{T(A)}))]^2$. In practice, $T(A)$ and $C(A)$ are also rounded conservatively.

Expected Person Years for Uncurtailed and Curtailed Studies Using Exact Tests: We first consider expected person years for studies being analyzed by exact tests and will

extend the ideas to those analyzed using normal and arcsine approximations later. The time to observe $T(E)$ cases in both groups for an uncurtailed study has a gamma distribution with parameters $T(E)$ and $\lambda_o[1-k(1-r^*)]$ and thus:

$$E[N_{u+t}] = T(E)/[\lambda_o[1-k(1-r^*)] \quad (3);$$

A curtailed study analyzed by exact tests stops once $C(E)+1$ cases occur in the treated group with H_o accepted or $T(E)$-$C(E)$ cases occur in the control group with H_o rejected. Thus, the number of cases that occur in this study, T_{CS}, is random, and from Schmitz (1993)[9] the expected number of cases is:

$$E[T_{CS}] = (C(E)+1)p^{(C(E)+1)} \sum_{i-0}^{(T(E)-C(E))-1} \left(\frac{(C(E)+1+i)!}{(C(E)+1)!(i)!}\right)(1-p)^i$$

$$+(T(E)-C(E))(1-p)^{(T(E)-C(E))} \sum_{i-0}^{C(E)} \left(\frac{((T(E)-C(E))+i)!}{(T(E)-C(E))!(i)!}\right)p^i \quad (4),$$

where p is the probability that a case occurs from the treated (rather than the control) group and equals $r^*k/[1-(1-k)r^*]$. Since for Poisson processes, the time needed to observe a combined number of cases from two sources (i.e., the control and treatment groups) is independent of the portion contributed by each source, the expected follow up time in a curtailed study is

$$E[N_{u+t}] = E[T_{CS}]/[\lambda_o(1-k(1-r^*)] \quad (5).$$

If r^*=0, then a study using curtailed stopping always stops when T-C cases occur in the controls so reworking (3), $E[N_{u+t}]$ also equals $[T(E)-C(E)]/[(1-k)\lambda_o]$.

Illustration of Expected Person Years for Studies Analyzed by Exact Tests: In the earlier example using exact tests with k=0.5, one sided α=0.025, where $T(E)=72$, $C(E)+1$ = 28 and $T(E)-C(E)$=45; let λ_o=0.0001 and r^*=0.5. Then by (3), for r^*=1, r^*= 0.5 and r^*= 0, it takes $E[N_{u+t}]$ = 720,000, 960,000 and 1,440,000 person years, respectively, to stop the study at $T(E)$ cases in both groups. The more effective a treatment is, the slower cases occur in the treated group making it take longer to recruit $T(E)$ cases in both groups. If the same study is stopped at the earliest of $C(E)+1$ cases in the control

group and $T(E)-C(E)$ cases in the treated group using curtailed stopping, then for r^*=1, r^*= 0.5 and r^*= 0 from (4) the expected number of cases at stoppage is ; $E[T_{CS}]$ = 55.9, 65.7 and 45 respectively. Based on the rates of accrual for these cases, from (5) this corresponds to expected person years of $E[N_{u+t}]$ = 559,396, 842,840 and 900,000, respectively, using curtailed stopping for r^*=1, r^*= 0.5 and r^*= 0.

Expected Person Years for Studies Analyzed by Normal and Arcsine Approximations: For uncurtailed studies analyzed by normal and arcsine approximations, (3) extends directly substituting $T(N)$ or $T(A)$ for T(E). The setting is more complicated for curtailed studies analyzed by normal and arcsine approximations. Since asymptotically, stoppage occurs at $C(N)$ cases in the treated group by approximations, versus $C(E)$+1 cases using exact methods (4) must be modified replacing the $C(E)+1$ terms with, $C(N)$, the $(T(E)-C(E))$ terms with $(T(N)-C(N))$, and the stand alone $C(E)$ terms with $C(N)-1$; and similarly for the arcsine approximation. Furthermore, since $C(N)$, $T(N)-C(N)$, $C(A)$ and $T(A)-C(A)$ are not integers, (4) cannot be applied directly, but can be applied through interpolation of respective upper and lower integer values as described in Section 5.3 to obtain $E[T_{CS}]$ which are then used in (5) to get $E[N_{u+t}]$.

3. MINIMIZATION OF *T* WITH RESPECT TO *k* IN UNCURTAILED STUDIES

Maximizing $\text{arcsine}\left(\sqrt{k}\right)-\text{arcsine}\left(\sqrt{r_a k/1-k(1-r_a)}\right)$ minimizes $T(A)$ in (2). This happens when $\frac{1}{\sqrt{k(1-k)}}-\frac{1-k(1-r_a)}{\sqrt{r_a k}\sqrt{1-k}}[\frac{r_a}{1-k(1-r_a)}+\frac{r_a k(1-r_a)}{(1-k(1-r_a))^2}]=0$ which (from first and second derivatives) for $r\neq 1$ occurs when

$$k=1/(1+\sqrt{r_a}) \qquad (6)$$

Thus, irrespective of the values of α and β, for r_a = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6 and 0.7 respectively, $T(A)$ is minimized with k = 0.76, 0.69, 0.65, 0.61, 0.59, 0.56 and 0.54 of the person years treated.

Rearranging the terms in the square root of (1) and taking first and second derivatives of the logarithm show $T(N)$ is minimized when

$$k = \left[Z_\alpha + Z_\beta\sqrt{r_a}\right] / \left[(1+r_a)Z_\alpha + 2Z_\beta\sqrt{r_a}\right] \quad (7)$$

When $\alpha=\beta$, (7) reduces to (6). For $\alpha<\beta$, $T(N)$ is minimized with more person years in the treatment group than minimizes $T(A)$, while for $\alpha>\beta$, $T(N)$ is minimized with less person years in the treatment group than minimizes $T(A)$. For example, with r_a=0.2; if α=0.025 and β= 0.2 then $T(N)$ is minimized by k=0.75, while if α=0.2 and β= 0.025, $T(N)$ is minimized by k=0.62. Since for most analyses done in practice $\alpha<\beta$, the k which minimizes $T(N)$ places even more to the treatment group than the k which minimizes $T(A)$.

To study associations between T and k under arcsine and normal approximations, Figure 1 plots $T(A)$ and $T(N)$ for (one sided α=0.025, β=0.2) and (one sided α=0.005, β=0.1), for r_a=0.2 and 0.5, and 0.1 $\leq k \leq$0.9. Exact $T(E)$ are also given (denoted by E) for k = 0.1, 0.2, …, 0.9. The values of α and β presented in Figure 1 and elsewhere span the range commonly used; one sided α=0.025 and 0.005 correspond to two sided α=0.05 and 0.01 respectively. From here on, all α presented are one sided and we drop the "one sided" designation. Similarly, r_a>0.5 may often not be a worthwhile reduction and r_a=0.2 might be a maximal effectiveness powered even if investigators had more confidence in the intervention. Table 1 presents $T(A)$, $T(N)$ and $T(E)$ from Figure 1 and corresponding $C(A)$, $C(N)$ and $C(E)$ for (k = 0.1, 0.2, …., 0.9); with non-integer values of $T(A)$, $T(N)$, $C(A)$ and $C(N)$ rounded to larger integers.

INSERT FIGURE 1 AND TABLE 1 HERE

For example, in the curves of Figure 1A (r_a=0.2, α=0.025 and β=0.2) for k=0.5, $T(A)$ was 14.74 while $T(N)$ was 15.06 and the exact $T(E)$ was 17. In Table 1A for (r_a=0.2, α=0.025, β=0.2 and k=0.5), $T(A)$ and $T(N)$ are rounded up to 15 and 16 respectively. The rejection regions for H_o with (r_a=0.2, α=0.025, β=0.2 and k=0.5) are $\leq$3 of the 15,

$\leq$4 of the 16 and $\leq$4 of the 17 cases in the treated group, respectively, by the rounded up arcsine and normal approximations and the exact test.

In Table 1, $T(E)$ was minimized at k=0.7 for r_a=0.2 and at k=0.6 for r_a=0.5 and reductions in $T(E)$ from optimal k Vs k=0.5 in Table 1 tended to be 2-3 cases. In Figure 1, $T(N)$ tends to be closer to $T(E)$ for k between 0.2 and 0.5 while $T(A)$ is closer for k $\geq$0.6. Among the decimal k's in Table 1, $T(E)$ was minimized at fractions closer to the arcsine minimum (6) than to the normal minimum (7). However, i n Figures 1, $T(A)$, $T(N)$ and $T(E)$ are level in regions around their respective optimal k suggesting convenient fractions near the optimums perform nearly as well to minimize T. For example, using the arcsine approximation with r_a=0.2, α=0.025 and β=0.2, $T(A)$=12.77 for the optimal k=0.69 (from (6)) Vs 12.92 for k=0.65 and 12.94 for k=0.75.

4. EFFECT OF CURTAILED STOPPING ON MINIMAL CASES AND OPTIMAL k

While $E[T_{CS}]$ for a curtailed study is always less than T from the same design that is not curtailed, (4) does not lend itself to easy analysis to determine the impact of k and r^* on $E[T_{CS}]$. To observe these patterns directly, Table 3 presents $T(E)$ and $E[T_{CS}]$ from the same designs in Table 1 and Figure 1 analyzed by exact test. For example, looking at k=0.4 in Table 2.B. (with r_a=0.2, α=0.005, and β=0.1) we see $T(E)$ from an uncurtailed study is 34; the same value it is in Table 1.B. But with curtailed stopping applied to the same design, from (4) to the nearest decimal $E[T_{CS}]$ = 17.5, 31.0 and 28 for $r^*=1$, $r^*=r_a=0.2$ and $r^*=0$, respectively. Note for $r^*=0$, $E[T_{CS}]$ is more directly obtained as $T(E)$ - $C(E)$ or from Table 1.B or 34 – 6.

INSERT TABLE 2 HERE

While for all r^* the minimal $E[T_{CS}]$ was smaller than the minimal T for the same α and β combination, there was not one value of k which simultaneously minimized $E[T_{CS}]$ for all values of r^*. When r^*=1, $E[T_{CS}]$ was consistently minimized by k=0.5 (although k = 0.6 and sometimes k=0.7) did nearly as well, and was ~40% smaller than the optimal

uncurtailed $T(E)$ from the corresponding uncurtailed study when r_a=0.2 and ~20% smaller when r_a=0.5. For $r^*=r_a$ (and the study now curtailed mostly at $T(E)$ - $C(E)$ cases being treated rather than at $C(E)$ + 1 cases in the control group) the optimal k rises to 0.8 for r_a=0.2 and 0.6-0.7 for r_a=0.5 with $E[T_{CS}]$ being 10-20% smaller than the optimal $T(E)$ in the table. For $r^*=0$, with the study now always curtailed at $T(E)$ - $C(E)$ cases treated, the optimal k shift towards putting as many subjects as possible in the treated group; $E[T_{CS}]$ was optimized at the largest k, 0.9, in Tables 2 with reductions of > 70% over the optimal $T(E)$ in the table; and $E[T_{CS}]$ continues to decreases with k for k>0.9.

5. MINIMIZING PERSON YEARS IN UNCURTAILED AND CURTAILED STUDIES

5.1 Some Observations on Uncurtailed Studies

We begin with some analytical inferences on $E[N_{u+t}]$ for uncurtailed studies. From (3), $E[N_{u+t}]$ for $r^*=1$ is minimized by the same k that minimizes T (i.e., $T(E)$, $T(A)$ or $T(N)$) which by (6) and (7) and Table 1 tends to put more person years in treatment. However, a clinical trial is only undertaken if substantial evidence already exists that the null hypothesis is false, thus values of k which minimize person years when $r^*<1$ may have more relevant. For $r^*<1$ in uncurtailed studies, substituting the arcsine approximation $T(A)$ from (2) for $T(E)$ into (3) gives for specified α, β, r_a and k.

$$E[N_{u+t}] \approx \frac{1}{\lambda_o(1-k(1-r^*))}\left[\frac{(Z_\alpha + Z_\beta)}{4(\arcsin\sqrt{k} - \arcsin\sqrt{r_a k/1-k(1-r_a)}}\right]^2 \qquad (8)$$

We could not analytically minimize (8) with respect to k for $r^*=r_a$ or $r^*=0$. But for $0 \le r_1^* < r_2^* \le 1$ the ratio of $E[N_{u+t}]$ for r_2^* in (8) to $E[N_{u+t}]$ for r_1^* in (8) is; $[(1-k(1-r_1^*)]$ / $[(1-k(1-r_2^*)]$. The first derivative of this ratio with respect to k is negative which means that relative to r_2^*, $E[N_{u+t}]$ is minimized at a smaller k for r_1^*. Thus, from (6), and (8) using the arcsine approximation, $k_{r^*-0} < k_{r^*-r_a} < 1/(1+\sqrt{r_a})$ where k_{r^*-0} minimizes E[N_{u+t}] for $r^*=0$ and $k_{r^*-r_a}$ minimizes E[N_{u+t}] for $r^*=r_a$. While allocating more person

years to treatment minimizes E[N_{u+t}] in uncurtailed studies when r^*=1, t as r^* decreases, the optimal k to minimize E[N_{u+t}] may shift back towards less person years treated.

5.2 Some Observations on Curtailed Studies

Table 2 showed that for $r^*=r_a$, and $r^*=0$, $E[T_{CS}]$ was minimized by $k > 0.5$. But this may not transfer to minimizing E[N_{u+t}]. With effective treatments, curtailed studies mostly stop with cases in the control group exceeding $T(E)-C(E)$ (or $T(A)-C(A)$ or $T(N)-C(N)$). Since control groups are allocated (1-k) of the person-years, for $r^*=0$, E[N_{u+t}] equals $E[T_{CS}]/[\lambda_o(1-k)]$ and for $r^*=r^a$, E[N_{u+t}] is a little smaller than $E[T_{CS}]/[\lambda_o(1-k)]$. With the ratio $E[N_{u+t}]/E[T_{CS}]$ inverse (or roughly inverse) to (1-k), this suggests the value of k which minimizes E[N_{u+t}] is smaller (shifts back to more untreated subjects) than the k which minimizes $E[T_{CS}]$ for $r^*=r_a$, and $r^*=0$.

Table 2 also showed that for $r^*=1$, placing k=0.5 subjects in treatment tended to minimize $E[T_{CS}]$. Since, with H_o true, the study will most often (i.e., (1-α) of the time) stop from the number of cases from the treatment group exceeding $C(E)+1$ (or $C(A)$ or $C(N)$) and k of the person years are allocated to treatment, by an argument similar to that in the previous paragraph, the value of k which minimizes E[N_{u+t}] when $r^*=1$ should be larger (i.e., put more person years into treatment) than the value of k which minimizes $E[T_{CS}]$ which based on Table 2 suggests this k will be larger than 0.5.

5.3 Examples of Effect of k and Curtailed Stopping Times on E[N_{u+t}]

Figures 2A – 2D present E[N_{u+t}] needed by k for $r^*=1$, $r^*=r_a$ and $r^*=0$ based on the arcsine approximation (2) incorporated into (3) – (5) as described in Section 2). The exact test would not give useful results over a continuous range of k in Figures 2 due to discrete fluctuation. Also, due to the number of plots, it was unfeasible to use both the arcsine and normal approximations here. However, the results of both approximations were not qualitatively different in Figures 1 or for the estimates in Figures 2. Finally, values in Figure 2 are standardized as ratios to the E[N_{u+t}] for an uncurtailed study with the same α, β, r_a, λ_o and k=0.5. These standardized ratios are invariant to λ_o.

For each combination of α, β and r_a, the left side figures compare standardized E[N_{u+t}] (by k) when the null hypothesis is true, $r^*=1$, for uncurtailed studies (solid line) to (dashed line) standardized $E[T_{CS}]$ for the same (i.e, same α, β, r_a and λ_o) curtailed studies. As noted in Section 2 for curtailed studies, since $T(A)-C(A)$ and $C(A)$ do not take on integer values and thus cannot be directly entered into (4), interpolated values were obtained from the upper and lower integers. For example, if $T(A)-C(A)=20.3$ (upper integer 21 and lower integer 20) and $C(A)$=9.6, (upper integer 10 and lower integer 9) then the interpolated $E[T_{SC}]$ would equal (0.7*0.4) * Q_1 + (0.7*0.6) * Q_2 + (0.3*0.4) * Q_3 + (0.3*0.6) * Q_4 where Q_1, Q_2, Q_3 and Q_4 are (4) obtained with [$T(A)-C(A)$), $C(A)$] equal to [20,9], [20,10], [21,9] and [21,10], respectively.

The right side plots compare standardized E[N_{u+t}] of uncurtailed studies for two settings of interest when the null hypothesis is not true, $r^*=0$ (top solid line) and $r^*=r_a$ (second highest solid line) to standardized E[T_{CS}] for curtailed studies when $r^*=0$ (higher dashed line) and $r^*=r_a$ (lower dashed line). For Figures 2B – 2D, (given the scales of the plots) the two dashed lines for curtailed studies are somewhat overlaid on each other. As described in the previous paragraph, interpolation was used to obtain E[T_{CS}] from (4) for the non-integer values of $T(A)-C(A)$ and $C(A)$.

INSERT FIGURE 2 HERE

In the left plot of Figure 2A, (r_a=0.2, α=0.025 and β=0.2,), the highest solid line is 0.90 at k=0.7. Thus if $r^*=1$ (H_o is true) and a study is uncurtailed , E[N_{u+t}] is 0.90 times what it is using the same uncurtailed design with k=0.5. If for $r^*=1$, the uncurtailed study lasted an expected 10,000 person years with k=0.5, then the uncurtailed study would need an expected 9,000 person years for $r^*=1$ when k=0.7. The lower dashed line in the left plot of Figure 2A is 0.53 at k=0.5; it takes an expected 0.53 times the person years to stop a (interpolated) curtailed study for k=0.5 than for the same design with uncurtaled stoppage for $r^*=1$ and k=0.5. At k=0.7, the value of the dashed (curtailed study) line in the left plot of Figure 2A is 0.51, barely lower than for k=0.5.

In the left plots of Figures 2A & 2B reflecting studies planned with r_a=0.2; to minimize E[N_{u+t}] when r^*=1, the optimal k for uncurtailed studies was close to 0.7 with 10-14% reduction in E[N_{u+t}] versus k=0.5 in the same uncurtailed design. But, as Section 5.1 suggests might happen when using curtailed sampling for r^*=1, the optimal k decreases to (0.5 - 0.6), with reductions of close to 50% in E[N_{u+t}] versus the same uncurtailed design with k=0.5. In the left plots of Figures 2C & 2D reflecting studies planned with r_a=0.5, to minimize E[N_{u+t}] when r^*=1, the optimal k for uncurtailed studies was close to 0.6 with only ~2% reduction in E[N_{u+t}] versus k=0.5. But, as Section 5.1 suggests would happen, when using a curtailed study with r^*=1, the optimal k shifts closer to 0.5, with reductions of ~ 25% in E[N_{u+t}] compared to the same uncurtailed design with k=0.5.

In the right plot of Figure 2A, (α=0.025, β=0.2, r_a=0.2), the top solid line is 2.89 and the lower solid line is 1.97 for k=0.7. If k= 0.7 allocation is used for an uncurtailed study, it requires ~2.89 times as many person years if r^*=r_a=0.2 and 1.97 times as many if r^*=0 than does the same uncurtailed study design with k= 0.5 and r^*=1. These ratios are > 1 mostly because if treatments work, cases accrue more slowly in treated groups. At k=0.7, the top dashed line in the right plot of Figure 2A is 1.64 and the lower dashed line is 1.51 which means that if k=0.7 and (interpolated) curtailed stopping is used, it will require ~1.64 times and ~1.51 times as many person years for r^*=1 and r^*=r_a=0.2 , respectively, than it would for the same uncurtailed design when r^*=1 and k= 0.5.

The right plots of Figures 2A & 2B reflecting studies planned with r_a=0.2 when H_o is false, show with uncurtailed stopping; 1) If r^*=0, the optimal k to minimize E[N_{u+t}] occurs on (0.4 - 0.5), and requires about twice as many person years Vs uncurtailed stopping at k=0.5 and r^*=1. 2). However, for uncurtailed stopping with r^*=r_a=0.2, the optimal k to minimize E[N_{u+t}] was close to 0.5 and required ~67% more person years Vs uncurtailed stopping with r^*=1 and k=0.5. But, as Section 5.2 suggests might happen, when using curtailed stopping for these designs, with r^*=0 and r^*=r_a=0.2, (the upper and lower dashed lines, respectively, of Figures 2A & 2B), the k to minimize E[N_{u+t}] increases to (0.5 – 0.6) and requires ~45% more person years Vs uncurtailed stopping for k=0.5 and r^*=1.

The right plots of Figures 2C & 2D for studies planned with r_a=0.5 when H_o is false, show with uncurtailed stopping: 1) If r^*=0, then k close to 0.4 minimizes E[N_{u+t}] with almost twice as many person years being needed Vs uncurtailed stopping for this design with k=0.5 when r^*=1; 2) If r^*=r_a=0.5, k very close to 0.5 minimizes E[N_{u+t}] with ~40% more person years needed Vs uncurtailed stopping for k=0.5 when r^*=1. When using curtailed stopping for these designs, k~ (0.5-0.6) minimizes E[N_{u+t}] with ~30% more person years needed versus uncurtailed stopping for k=0.5 when r^*=1.

Table 3 presents expected N_{u+t} to complete uncurtailed and curtailed studies analyzed by exact tests for k = 0.4, 0.5, 0.6 and 0.7 which encompasses the optimal values in Figure 2. These were obtained from (5) applied to $T(E)$ in Table 1 or $E[T_{CS}]$ in Table 2 for the corresponding study. Incidence of λ_o=0.0001 per person year was assumed. But, as with Figure 2, for each combination of (α, β and r_a), the ratios of numbers in Table 2 are invariant to λ_o. As done previously, results are given for (α = 0.005, β=0.1) and (α = 0.025, β=0.2) and for r_a=0.2 and r_a=0.5.

INSERT TABLE 3 HERE

The first 4 rows of Table 3 give E[N_{u+t}] to finish comparative Poisson studies analyzed by exact tests with fixed T designed for r_a=0.2, α = 0.025 and β=0.2 when incidence under the null is λ_o =0.0001 per person year. For example, consider the study in the first row of Table 3 with k=0.4 of person years in treatment (from Table 1, the study is planned for $T(E)$=23 cases and H_o rejected if $C(E)$ = 4 or less of these are in the treated group). If this study is uncurtailed and lasts until all 23 cases occur, then E[N_{u+t}] is 230,000 when H_o (r^*=1) is true. This rises to 338,235 if r^*=r_a=0.2 and 383,333 if r^*=0 respectively as incidence of cases in the 40% of person years in the treated group is 0.00002 and 0 respectively (rather than being 0.0001 as is the case when r^*=1). Now suppose curtailed stopping is used with the above study stopped once either $C(E)+1$ = 5 cases occur in treated subjects or $T(E)-C(E)$ = 19 cases occur in controls. If r^*=1, and curtailed stopping is used, it takes an expected 124,263 person years to stop the study by reaching either 5 cases in treated subjects or 19 cases in the controls, while if r^*=r_a=0.2 and curtailed stopping is used, it takes an expected 305,520 person years to

reach this point. It takes an expected 316,667 person years to obtain 19 cases among untreated subjects; which for $r^*=0$ is E[N_{u+t}] for the curtailed stopping time.

In general, the patterns (as reflected by taking ratios of expected person years) for exact tests in Table 3 are similar to those in Figures 2. When $r^*=1$ and r_a = 0.2; 1) for uncurtailed studies, allocation of k=0.7 person years (versus k=0.5) to treatment reduced E[N_{u+t}] by 10-12% (see column 2); but 2) for curtailed studies the optimal allocations were k=0.5-0.6 (see column 5) and reduced E[N_{u+t}] about 40% below the optimums for uncurtailed studies. When $r^*=1$ and r_a = 0.5; 1) for uncurtailed studies, allocation of k=0.6 to treatment reduced E[N_{u+t}] by 1-3% versus equal allocation; and 2) but for curtailed studies, k=0.5 was optimal and reduced E[N_{u+t}] about 20% below the optimums for uncurtailed studies.

When $r^*= r_a$, for both r_a=0.2 and r_a=0.5, allocation of k=0.5 minimized E[N_{u+t}] for both uncurtailed (column 3) and curtailed (column 6) studies, while the minimal E[N_{u+t}] for curtailed studies were about 6-13% lower than the minimal E[N_{u+t}] for uncurtailed studies. When $r^*=0$; 1) E[N_{u+t}] was usually minimized at k=0.4 for uncurtailed studies (see column 4); 2) but as column 7 shows, irrespective of r_a, E[N_{u+t}] was consistently minimized at k=0.5 for curtailed studies; and 3) the minimal E[N_{u+t}] was ~25% less for curtailed than for uncurtailed studies when r_a=0.2 and ~40% less for curtailed than uncurtailed studies when r_a=0.5.

6. DISCUSSION

For comparative Poisson trials with fixed events, if the objective is to minimize T, and the study is blinded until all cases occur, then from Figure 1, Table 1 and (6), letting k be a convenient fraction close to $k = 1/(1+\sqrt{r_a})$ is close to optimal and may reduce the number of cases by 2-3 Vs equal allocation. If minimization of $E[N_{u+t}]$ is of most importance in an uncurtailed study, then allocating k=0.5 to treatment may be the best compromise as it is close to optimal when $r^*=0$ or $r^*=r_a$ and not too far from optimal if $r^*=1$. Allocating k>0.6 to treatment risks substantially larger $E[N_{u+t}]$ for completely effective treatments ($r^*=0$) while allocating k<0.5 makes $E[N_{u+t}]$ larger when the

treatment has no effect (r^*=1). However, if it is important to allocate more subjects into treatment for other reasons, k as high as 0.6 often performs nearly as well as k=0.5 in uncurtailed studies.

If logistical and/or regulatory concerns do not prevent this, curtailed stopping can reduce cases and follow up time considerably. From Table 2, with an optimal k, curtailed sampling can reduce expected number of cases by 10-20% when $r^*=r_a$, 20-40% when r^*=1, and by 60% or more (see also Appendix A) when r^*=0. From Table 3 (and Figure 2), with an optimal k, curtailed sampling can reduce expected person years by about 20-40% when r^*=1, 10% when $r^*=r_a$, and 25%-40% when r^*=0.

For curtailed studies k ranging from 0.5-0.7 minimizes expected cases when r^*=1 while k ranging from 0.7-0.8 minimizes expected cases $r^*=r_a$. But if one strongly believes r^*=0 and reducing cases had paramount importance (as might be the case for a fatal disease) then setting k much larger may make sense for a curtailed study; see Appendix A for further discussion on this. However, to minimize the expected number of person years when curtailed stopping is used, allocating k=0.5 to treatment was roughly optimal for the different r^* evaluated here, but larger values up to k=0.6 also performed comparably well.

Logistics of large multicenter studies often mandate k be a convenient fraction with a small denominator. Since T and $E[N_{u+t}]$ were often level in overlapping regions near the optimal k across the different values of r^* studied here, this may not present a big problem. However, due to the discreteness of Exact Tests, power may be either greater (or smaller) in very localized regions of k. As Appendix B shows, caution should be taken in using a specific localized value of k to minimize T if the investigators cannot guarantee maintaining that k during the study due to logistics or drop out.

It should be noted that besides the curtailed approach, other methods for early stopping of binomial studies have been proposed These methods include Armitage's[18] and Spicer's[19] stopping rules for sequential binomials and group sequential methods[9,10], and they may warrant further study in terms of overall reduction of cases and person-years

and for optimal k. We focused on hypothesis tests H_o: r=1 Vs H_a: r<1. However, for comparative Poisson studies of adverse events from interventions, the null hypothesis is H_o: r=q for some q>1 Vs H_a: q=1. Now the treated group has higher risk under H_o instead of having lower risk under H_a as was the case for this paper. This likely flips the direction of k to optimize results. For example, using the arcsine transformation, a similar result to (6) now gives $k = 1/(1+\sqrt{q})$ person years to treatment minimizes $T(E)$ for adverse event studies, which places a smaller portion in treatment. Others have also observed that $T(E)$ and E[N_{u+t}] for comparative Poisson studies of adverse events tend to be minimized with more subjects in the untreated group[7].

In conclusion, this study suggests that while curtailed sampling has not been frequently used in fixed event comparative Poisson studies of interventions, curtailment could result in substantial savings of time and reduction of cases. While the common practice of allocating k=0.5 of the cases to the treatment group was often optimal in terms of minimizing cases and person years needed for uncurtailed and curtailed studies, depending on the setting and minimization goal, allocation of a convenient k as large as 0.6 of the person years to treatment often performed well. Thus, if allocating more subjects to treatment provides additional incentive for persons to enroll, this should be considered. However, due to very localized minimums with exact tests one should not intensively search for a specific k that minimizes cases or follow up time unless that allocation can be logistically guaranteed.

REFERENCES

1. Gail M. 'Power considerations for designing comparative Poisson trials' *Biometrics* **30**, 231-7 (1974).

2. Brown CC & Green SB. 'Additional power comparisons for designing comparative Poisson trials' *Am J Epidemiol* **115(5)**, 752-8 (1982).

3. Brittain E & Schlesselman JJ 'Optimal Allocation for the Comparison of Proportions' *Biometrics* **38**,1003-9 (1982).

4. Duchateau L, McDermott B, Rowlands GJ. 'Power evaluations of small drug and vaccine experiments with binary outcomes' *Stat Med* **17**, 111-20 (1998).

5. Blackwelder WC. 'Sample size and power for prospective analysis of relative risk' *Stat Med* **12** 691-698 (1993).

6. Hoover DR, Blackwelder WC. 'Allocation Of subjects to test null relative risks smaller than one'. *Stat Med* **20** 3071-82 (2001).

7. Dragalin V, Fedorov V, Cheuvart B. 'Statistical approaches to establishing vaccine safety' *Stat Med* **21**, 877-93 (2002).

8. Alling DW. 'Closed sequential tests for binomial probabilities' *Biometrika* **53**,73-84 (1966).

9. Schmitz N. *Optimal Sequentially Planned Decision Procedures*. Lecture Notes in Statistics **79**, New York: Springer-Verlag (1993).

10. Jennison C & Turnbull B. *Group Sequential Methods With Applications to Clinical Trials*: London: Chapman and Hall CRC (2000).

11. Przyboroski J & Wilenski H. 'Homogeneity of results in testing samples from Poisson series'. *Biometrika* **31**, 313-23 (1939).

12. Lehman EL. '*Testing Statistical Hypotheses'*. New York: Wiley 140-2 (1959).

13. NCSS Statistical Software. PASS 2000. '*Power analysis and sample size software'*. NCSS Statistical Software, Kaysville: Utah (2000).

14. Pocock SJ *Clinical Trials a Practical Approach* Chinchester: Wiley (1983)

15. Sahai H, Khurshid A. Formulas and tables for the determination of sample sizes and power in clinical trials for testing differences in proportions for the matched pair design: a review. *Fundam Clin Pharmacol* 10(6):554-63 (1996)

16. Dixon WJ & Massey FM. '*Introduction to Statistical Analysis 2nd Edition'* New York City: McGraw Hill, 274-7 (1969).

17. Sillito GP. 'Note on approximations to the power function of the 2 x 2 comparative trial' *Biometrika* **36**, 347-52 (1948).

18. Armitage P. 'Restricted sequential procedures', *Biometrika* **44**, 9-26 (1957).

19. Spicer C. 'Some new closed sequential designs for clinical trials' *Biometrics* **18**, 203-11 (1962).

20. Barnard G.A. 'On alleged gains in power from lower p-values' *Stat Med* **8**, 1469-77 (1989).

21. Whitehead J. 'Overrunning and underrunning in sequential clinical trials' *Control Clin Trials* **13**, 106-21 (1992).

Figure 1 - Number Of Disease Cases (T) Needed For Different Fixed Event Comparative Poisson Studies

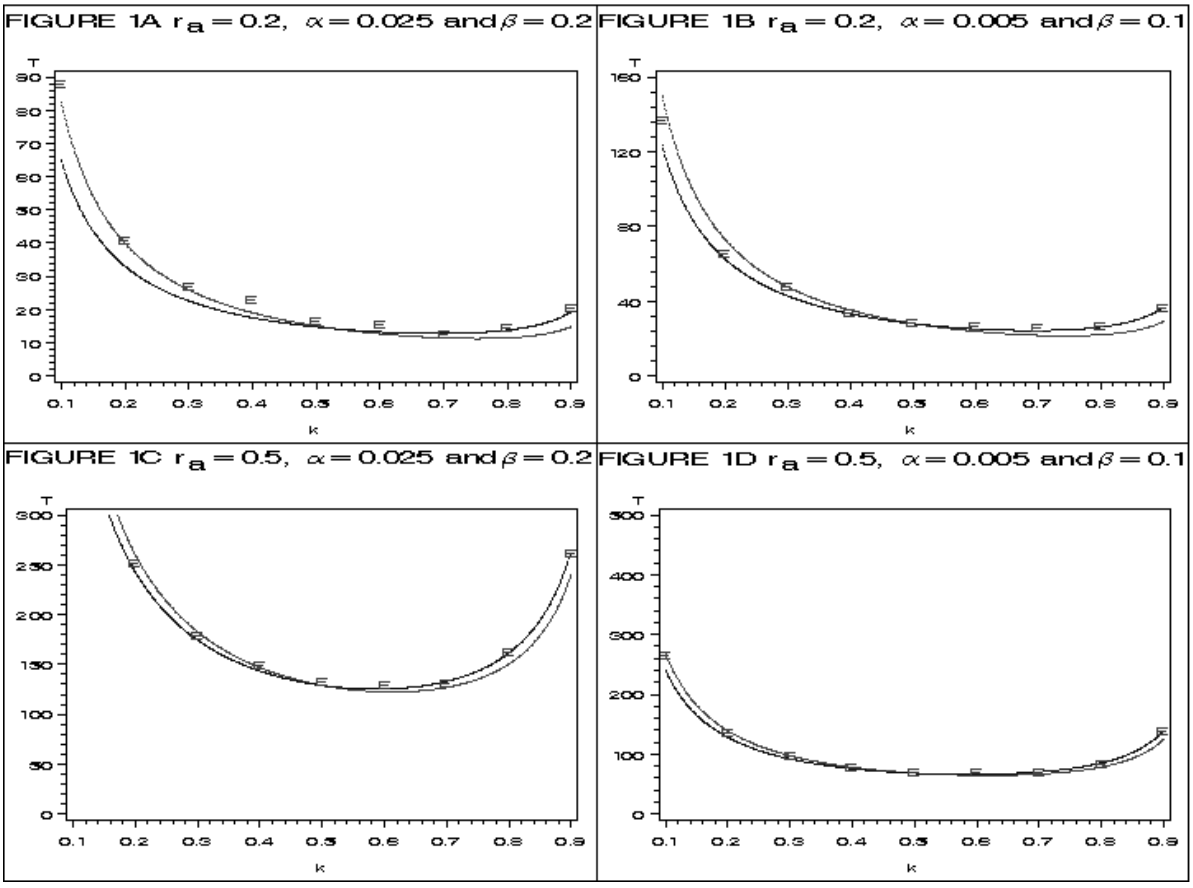
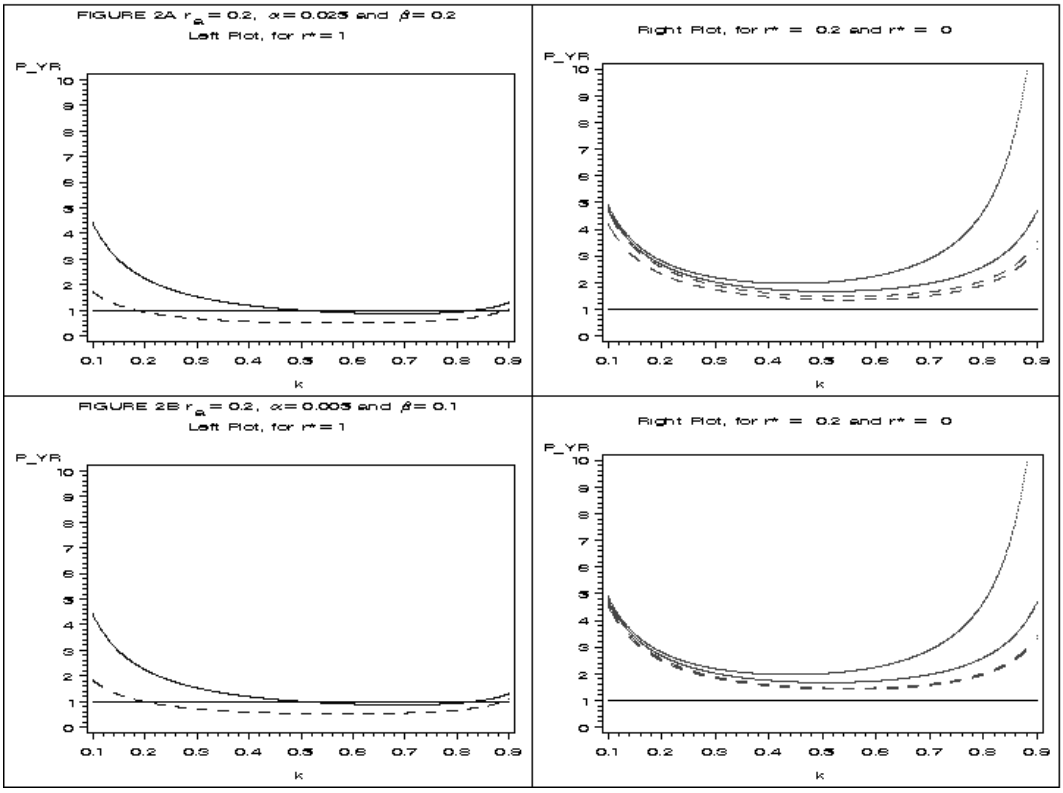


Figure 2 – Standardized[1] Expected Person Years (N_{u+t}) Needed by Arcsine Approximation for Different Fixed Event Comparative Poisson Studies



LEFT PLOTS – Standardized Times When H_o True

For an Uncurtailed Study to Stop at $T(E)$ Cases When r^*=1

For a Curtailed Study to Stop at T_{CS} Cases When r^*=1

RIGHT PLOTS – Standardized Times When H_o False

For an Uncurtailed Study to Stop at $T(E)$ Cases:

When $r^*=r_a$

When r^*=0

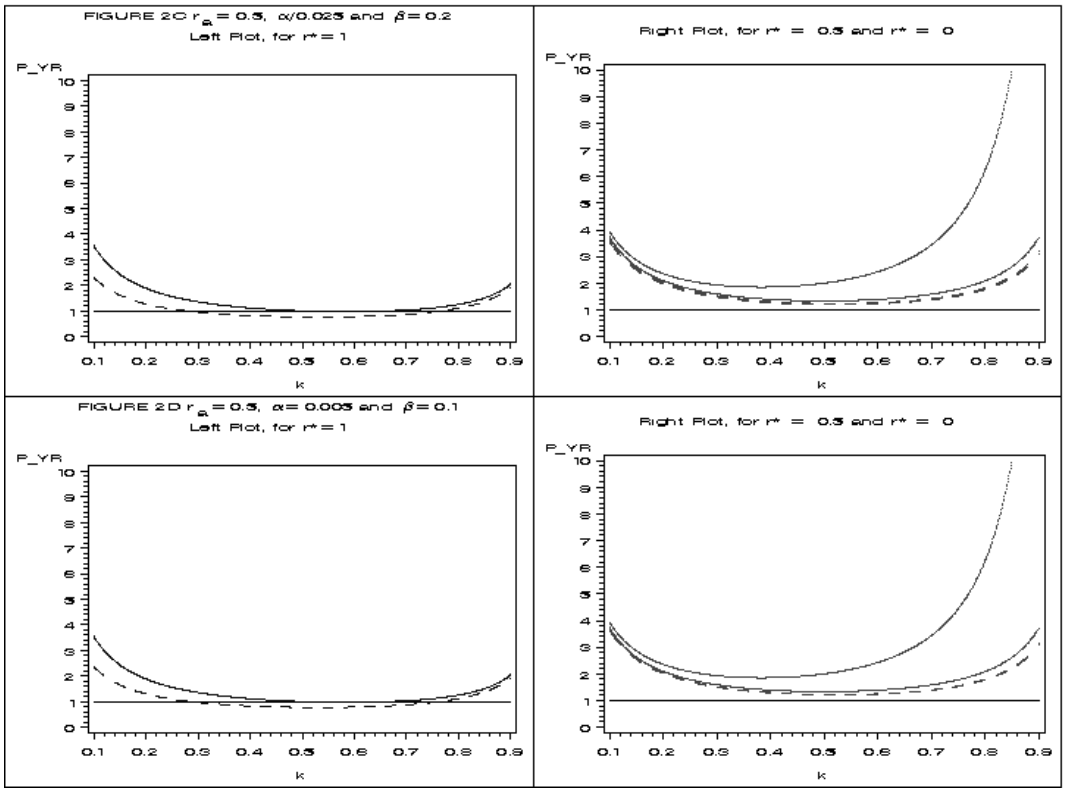
For a Curtailed Study to Stop at T_{CS} Cases

When $r^*=r_a$

When r^*=0

[NOTE – The curtailed study lines for $r^*=r_a$ and r^*=0 are overlaid for Figures 2B, 2C and 2D

1. Standardized as the ratio of the $E[N_{u+t}]$ for a study with the given conditions and stopping rule to: $E[N_{u+t}]$ for an uncurtailed study with k=0.5 and H_o: (r^*=1) true that has same α, β, r_a and untreated incidence λ_o



LEFT PLOTS – Standardized Times When H_o True

For an Uncurtailed Study to Stop at $T(E)$ Cases When r^*=1

For a Curtailed Study to Stop at T_{CS} Cases When r^*=1

RIGHT PLOTS – Standardized Times When H_o False

For a Uncurtailed Study to Stop at $T(E)$ Cases:

When $r^*=r_a$

When r^*=0

For a Curtailed Study to Stop at T_{CS} Cases

When $r^*=r_a$

When r^*=0

[NOTE – The curtailed study lines for $r^*=r_a$ and r^*=0 are overlaid for Figures 2B, 2C and 2D

1. Standardized as the ratio of the $E[N_{u+t}]$ for a study with the given conditions and stopping rule to: $E[N_{u+t}]$ for an uncurtailed study with k=0.5 and H_o: (r^*=1) true that has same α, β, r_a and untreated incidence λ_o

Table 1 – *T* and C For Given Type 1 & Type 2 Errors, r_a and Decimal Allocation Ratios *k*

A. r_a= 0.2, α = 0.025, and β = 0.2

Method	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	*k*=0.2	*k*=0.3	*k*=0.4	*k*=0.5	*k*=0.6	*k*=0.7	*k*=0.8	*k*=0.9
Arcsine Approx [1] $T(A)$	65	34	23	18	15	14	13	14	20 (14)
$(C(A))$	(2)	(2)	(3)	(3)	(3)	(4)	(5)	(7)	
Normal Approx [1] $T(N)$	83	40	26	20	16	13	12	12	15
$(C(N))$	(2)	(3)	(3)	(3)	(4)	(4)	(5)	(6)	(11)
Exact $T(E)$	88	41	27	23	17	16	15	15	21
$(C(E))$	(3)	(3)	(3)	(4)	(4)	(5)	(6)	(8)	(15)

B. r_a= 0.2, α = 0.005, and β = 0.1

Method	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	*k*=0.2	*k*=0.3	*k*=0.4	*k*=0.5	*k*=0.6	*k*=0.7	*k*=0.8	*k*=0.9
Arcsine Approx [1] $T(A)$	124	63	43	34	28	26	25	27	37
$(C(A))$	(5)	(5)	(6)	(6)	(7)	(9)	(11)	(15)	(27)
Normal Approx [1] $T(N)$	151	74	48	36	29	24	22	22	30
$(C(N))$	(5)	(6)	(6)	(6)	(7)	(8)	(9)	(12)	(22)
Exact $T(E)$	137	66	48	34	29	27	26	27	37
$(C(E))$	(5)	(5)	(6)	(6)	(7)	(9)	(11)	(15)	(27)

C. r_a= 0.5, α = 0.025, and β = 0.2

Method	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	*k*=0.2	*k*=0.3	*k*=0.4	*k*=0.5	*k*=0.6	*k*=0.7	*k*=0.8	*k*=0.9
Arcsine Approx [1] $T(A)$	241	129	92	76	68	67	71	86	140
$(C(A))$	(15)	(17)	(19)	(22)	(25)	(32)	(41)	(61)	(118)
Normal Approx [1] $T(N)$	269	140	98	78	69	65	67	79	125
$(C(N))$	(17)	(17)	(20)	(22)	(26)	(31)	(39)	(56)	(105)
Exact $T(E)$	267	137	98	80	72	70	72	85	140
$(C(E))$	(17)	(18)	(20)	(23)	(27)	(33)	(42)	(60)	(118)

D. r_a= 0.5, α = 0.005, and β = 0.1

Method	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	k=0.2	k=0.3	k=0.4	k=0.5	k=0.6	k=0.7	k=0.8	k=0.9
Arcsine Approx [1] $T(A)$	457	243	175	144	129	126	133	162	264
$(C(A))$	(30)	(33)	(37)	(43)	(50)	(61)	(79)	(115)	(224)
Normal Approx [1] $T(N)$	500	260	183	147	129	123	127	151	240
$(C(N))$	(33)	(35)	(39)	(43)	(50)	(60)	(75)	(108)	(204)
Exact $T(E)$	476	252	180	150	133	131	132	163	262
$(C(E))$	(31)	(34)	(38)	(44)	(51)	(64)	(78)	(116)	(222)

Table 2 – T(E) and E[T_{CS}] for Uncurtailed and Curtailed Studies Analyzed by Exact Test

A. ***r_a*** **= 0.2,** α **= 0.025 AND** β **= 0.2**

STOPPING RULE	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	k=0.2	k=0.3	k=0.4	k=0.5	k=0.6	k=0.7	k=0.8	k=0.9
Uncurtailed Study $T(E)$	88	41	27	23	17	16	15	15	21
Curtailed Study E[T_{CS}]									
$r^* = 1$	39.7	19.8	13.2	12.4	9.92	9.94	9.96	11.2	17.7
$r^* = r_a$	83.9	38.4	24.8	20.8	14.9	13.7	12.5	11.9	15.8
$r^* = 0$	85	38	24	24	13	11	9	7	6

B. ***r_a*** **= 0.2,** α **= 0.005 AND** β **= 0.1**

STOPPING RULE	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	k=0.2	k=0.3	k=0.4	k=0.5	k=0.6	k=0.7	k=0.8	k=0.9
Uncurtailed Study $T(E)$	137	66	48	34	29	27	26	27	37
Curtailed Study E[T_{CS}]									
$r^* = 1$	59.9	30.0	29.7	17.5	16.0	16.7	17.1	20.0	31.1
$r^* = r_a$	132.7	62.7	41.2	31.0	25.8	23.0	21.6	21.2	27.4
$r^* = 0$	132	61	38	28	22	18	15	12	10

C. ***r_a*** **= 0.5,** α **= 0.025 AND** β **= 0.2**

STOPPING RULE	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	k=0.2	k=0.3	k=0.4	k=0.5	k=0.6	k=0.7	k=0.8	k=0.9
Uncurtailed Study $T(E)$	267	137	98	80	72	70	72	85	140
Curtailed Study E[T_{CS}]									
$r^* = 1$	179.4	94.7	69.8	59.8	55.9	56.5	61.2	76.1	131.9
$r^* = r_a$	257.9	130.5	92.1	74.1	65.7	62.9	63.2	72.9	117.5
$r^* = 0$	250	119	78	57	45	37	30	25	22

D. ***r_a*** **= 0.5,** α **= 0.005 AND** β **= 0.1**

STOPPING RULE	ALLOCATION RATIO $k=N_t/N_{u+t}$								
	k=0.1	k=0.2	k=0.3	k=0.4	k=0.5	k=0.6	k=0.7	k=0.8	k=0.9
Uncurtailed Study $T(E)$	476	252	180	150	133	131	132	163	262
Curtailed Study E[T_{CS}]									
$r^* = 1$	319.9	174.9	130.0	112.5	104.0	108.3	112.8	146.2	247.7
$r^* = r_a$	465.7	243.1	170.9	140.1	121.9	116.5	115.9	139.8	218.1
$r^* = 0$	445	218	142	106	82	67	54	47	40

TABLE 3 – Expected Person Years to Complete Study by r^*, k and Stopping Times When Exact Tests are Used and Incidence in Untreated Subjects is $\lambda_o = 0.0001$ per Person Year[1]

ALLOCATION RATIO $k=N_t / N_{u+t}$	UNCURTAILED STOPPING TIMES			CURTAILED STOPPNG TIMES		
	$r^*=1$	$r^*=r_a$	$r^*=0$	$r^*=1$	$r^*=r_a$	$r^*=0$
A. For $r_a=0.2$, $\alpha=0.025$ and $\beta=0.2$						
0.4	230,000	338,235	383,333	124,263	305,520	316,667
0.5	170,000	283,333	340,000	99,962	248,888	260,000
0.6	160,000	307,692	400,000	99,432	263,140	275,000
0.7	150,000	340,909	500,000	99,572	284,030	300,000
B. For $r_a=0.2$, $\alpha=0.005$ and $\beta=0.1$						
0.4	340,000	500,000	566,667	174,808	456,408	466,667
0.5	290,000	483,333	580,000	159,859	430,430	440,000
0.6	270,000	519,230	675,000	166,514	443,035	450,000
0.7	260,000	590,909	866,667	171,231	459,843	466,667
C. For $r_a=0.5$, $\alpha=0.025$ and $\beta=0.2$						
0.4	800,000	1,000,000	1,333,333	598,251	925,772	950,000
0.5	720,000	960,000	1,440,000	559,396	842,840	900,000
0.6	700,000	1,000,000	1,750,000	565,438	898,531	925,000
0.7	720,000	1,107,692	2,400,000	612,738	972,376	1,000,000
D. For $r_a=0.2$, $\alpha=0.005$ and $\beta=0.1$						
0.4	1,500,000	1,875,000	2,500,000	1,124,650	1,751,550	1,766,667
0.5	1,330,000	1,733,333	2,660,000	1,039,670	1,625,880	1,640,000
0.6	1,310,000	1,871,429	3,275,000	1,082,286	1,664,760	1,675,000
0.7	1,320,000	2,030,769	4,400,000	1,111,408	1,809,810	1,833,333

1. Based on corresponding $T(E)$ and $C(E)$ reported in Table 1, and $E[T_{CS}]$ in Table 2

Appendix A – The Special Case of $r_a = 0$

We use α=0.05 and 0.01 rather than α=0.025 and 0.005 here since for r_a=0 only one sided rejection can occur; the one sided Type 1 error is doubled to attain the same overall Type 1 error. Since cases never occur among treated subjects under the alternative, power is always 100% so β is ignorable. Taken to the limit of $r_a = 0$, (6) and (7) place all subjects in the treated group which perhaps makes sense if one uses disease in historical controls against no disease ever in the study. As at least one case must occur for a comparison to be made, the minimum for T is 1. Since under the null hypothesis, (1-k) is the probability that this case would occur among untreated subjects; allocation of k=(1-α) person years to treatment achieves 100% power for $r^* = r_a$=0 with Type 1 error of α. For example, if k=0.95 person years are allocated to treatment and the study is stopped at T=1 with H_o: rejected if the case was untreated, the size is 0.05 and power is 1.00. This also minimizes E[N_{u+t}] under H_o.

Minimizing E[N_{u+t}] under H_a (whether or not accumulation of cases is curtailed) is equivalent to minimizing the ratio $T/[\lambda_o(1-k)]$ under the constraint that for H_o true, the probability all T cases are from the untreated group $(1-k)^T = \alpha$ or $T = \log(\alpha)/\log(1-k)$. Since λ_o is constant, assuming for now T is continuous, we wish to minimize $T/(1-k)$ $= \log(\alpha)/[(1-k)\log(1-k)]$. For all α, this occurs when $(1-k)\log(1-k)]$ is maximized. Taking logarithms and first and second derivatives, $(1-k)\log(1-k)]$ is maximized when $-1/(1-k)[1+1/\log(1-k)] = 0$ which for $0 \le k \le 1$ occurs when log(1-k) = -1 or k = 0.63 and $T = \log(\alpha) / \log(.37)$. For α=0.05, log(.05)/log(.37) almost exactly equals 3. With k=0.63 of the person years allocated to treatment, T=3 and the null hypothesis rejected if all 3 cases were untreated, E[N_{u+t}] under H_a is $3/[(1-0.63)\ \lambda_o] = 8.1/\lambda_o$ compared to $1/(.05\lambda_o) = 20/\lambda_o$ had k=0.95, the portion that minimizes T and E[N_{u+t}] under H_o. For α=0.01, $\log(\alpha) / \log(.37) = 4.63$ which going to closest integer values means either the allocation k which makes T=4 or that which makes T=5 should be used. Coincidentally, with k=0.6, $(1-0.6)^5$=0.01 and E[N_{u+t}] is $12.5/\lambda_o$ and with k=.68, $(1-0.68)^4$=0.01 and E[N_{u+t}] is also $12.5/\lambda_o$ meaning both allocations work equally well for this purpose.

Appendix B – Local Maximums Of Exact Tests

As noted by others, due to the discreteness of exact tests, size and power can vary in counterintuitive and problematic fashions that cannot be predicted by asymptotic results[20]. For example, Table B shows exact powers ($\Phi^{r^*}(E,T)$) for T= 13, 14 and 15 based on rejection regions $C(E)$ with one sided α=0.025 and k ranging from 0.67 to 0.71 for H_o: r^*=1 against H_a: r^*=r_a=0.2.

Table B – Powers for Exact Tests with One Sided α=0.025 for H_o: r^*=1 Vs H_a: r^* = r_a=0.2 and k Ranging from 0.67 to 0.71

$T(E)$	$k=N_t/N_{u+t}$				
	0.67	**0.68**	**0.69**	**0.70**	**0.71**
13	69%	66%	82%	79%	77%
14	81%	79%	76%	73%	70%
15	75%	87%	85%	83%	81%

As shown in Figure 1 and Table 1, for k =0.70, $T(E)$=15 was the minimal $T(E)$ that assured at least 80% power (or β=0.2). However, for k just one hundredth lower k=0.69, $T(E)$ could be reduced to 13 and obtain 82% power by the exact test. This would be a wise strategy assuming both (i) that the allocation ration of 0.69 could be maintained and (ii) that no more data will be collected once T=13 events occur.

However, in most large scale prophylactic intervention studies, neither (i) nor (ii) is possible. Random dropout, ineligibility, etc will occur in both intervention groups and in some settings the different treatment groups may have different dropout or follow up rates. This means that even if the planned allocation ratio was k=0.69, at the time $T(E)$ = 13 events occurs, the actual study ratio could be N_t/N_{u+t}=0.68. Even though the planned allocation to treatment was 0.69, which had a rejection region of $\leq$ 5 events in the treated group, the exact test analysis will be performed using the actual ratio of 0.68 and the corresponding rejection region $C(E)$ for that ratio, which

is ≤ 4 events in the treated group to ensure the nominal Type 1 error of 0.025. This change in the rejection region reduces power from 82% as planned to only 66%.

Also, even if the allocation ratio can be maintained at k=0.69, there is the possibility that events will occur after the study is "officially stopped" due to reaching T= 13 events but before the data is analyzed. Study subjects will have to be monitored after the official close out and scientifically / ethically any new cases observed cannot be ignored; a phenomenon that has been denoted "overrunning"[21]. Suppose because of one more case occurring after the "official closeout" a study with k=0.69 planned for $T(E)$=13 ends up being evaluated with k=0.69 and $T(E)$= 14 events. In spite of the additional event, the rejection region must remain at $C(E) \leq 5$ as increasing it to 6 will exceed the Type 1 error of 0.025 under the null. However, under the alternative leaving this rejection threshold at 5 event in the treated group in spite of one more event overall, reduces the power down to 76% below the desired threshold of 80%.

Therefore, when using exact methods to plan comparative Poisson trials that will be stopped at $T(E)$ events, one may need to consider both the powers for other values of k around the targeted allocation and for $T(E)$ that slightly exceed the planned value.