Wald-Type Testing and Estimation
Methods for Asymmetric Comparisons
of Poisson Rates

Brianna C. BRIGHT and Julia N. SOULAKOVA

This article considers the problem of asymmetric comparisons, that is, instances where one treatment serves as the compelling standard treatment for a certain disease or health condition, where the asymmetric comparisons are performed in terms of the Poisson rates. We propose asymptotic Wald-type tests and confidence intervals suitable for demonstrating superiority and noninferiority for a given margin. The testing methods for the difference and ratio are based on the unconstrained (MLE) and constrained (CMLE) maximum likelihood estimator. The resulting CMLE-based tests for asymmetric comparisons are equivalent to the standard (i.e., those for symmetric comparisons) CMLE-based tests. The asymmetric tests as well as the standard MLE-based tests are evaluated via simulations. The CMLE-based asymmetric tests are shown to adequately control the Type I error in most settings, while the MLE-based asymmetric tests are shown to have this property only in settings with relatively large sample sizes and means. In all settings, the MLE-based asymmetric tests are more powerful than both the asymmetric CMLE-based and the standard MLE-based tests. We also propose asymptotic confidence intervals that can be used to estimate the difference or ratio of the two rates (the presentation of this content includes supplementary material that is available online), and discuss how power and sample size estimation can be done to aid in study planning. Supplementary materials for this article are available online.

Key Words: Confidence intervals; Hypothesis testing; Poisson distribution; Variance estimation.

1. Introduction

The Poisson distribution is commonly encountered in practice, for example, in transplantation studies (Rochon, Aprile, and Cardella 1992). Therefore, not surprisingly, many statistical methods concerning testing or estimation of the Poisson rate have been developed and widely discussed (Pearson and Hartley 1966; Hall 1982; Schwertman and Martinez 1994b; Brown, Cai, and Dasgupta 2001; Barker 2002; Byrne and Kabaila 2005; Johnson, Kemp, and Kotz 2005; Swift 2009; Khamkong 2012; Patil and Kuklarni 2012). In addition, methods for the difference of the Poisson rates (Miettinen and Nurminen 1985; Weissfeld, Laurent, and Moulo 1991; Schwertman and Martinez 1994a; Stanney and Hamilton 2006; Liu et al. 2006; Ng, Gu, and Tang 2007; Li et al. 2011; Krishnamoorthy and Lee 2013; Bright and Soulakova, in press) and ratio of the Poisson rates (Nelson 1970; Sahai and Khurshid 1993; Price and Bonett 2000; Liu 2004; Tang and Ng 2004; Ng and Tang 2005; Barker and Cadwell 2008; Gu et al. 2008; Krishnamoorthy and Lee 2010; Li, Tang, and Wong 2014) have also been discussed. Among different methods for the difference and ratio of Poisson rates, the Wald-type hypothesis testing methods are one of the most commonly addressed in the literature.
(Ng and Tang 2005; Ng, Gu, and Tang 2007; Gu et al. 2008), as well as Wald-type interval-estimating methods (Weissfeld, Laurent, and Moultou 1991; Schwertman and Martinez 1994b; Price and Bonett 2000; Graham, Mengersen, and Morton 2003; Tang and Ng 2004; Stamely and Hamilton 2006; Liu et al. 2006; Barker and Cadwell 2008; Krishnamoorthy and Lee 2010; Krishnamoorthy and Lee 2013). The Wald-type and other statistical methods for functions of Poisson rates find wide application in many medical areas, for example, breast cancer research (Colditz et al. 1990; Liu 2004), coronary heart disease research (Stampfer and Willett 1985; Gu et al. 2008), drug research (Pitt et al. 1997; Liu et al. 2006), and infectious disease research (Hargarten and Baker 1985; Weissfeld, Laurent, and Moultou 1991).

All available methods regarding the difference and ratio of the Poisson rates are proposed for symmetric comparisons, that is, when “neither group represents a compelling standard for the other” (Falk and Koch 1998, p. 1604). However, there are settings where the reference treatment does serve as the compelling standard, for example, a study presented by Anbar (1983). In the latter case, the statistical method should reflect this “asymmetry” of the comparison, by giving more weight to information obtained from the reference group. If an asymmetric comparison is of interest, then the study should be planned accordingly, for example, the power and group sample sizes must be estimated for the asymmetric tests that will be used in the data analysis. Asymmetric comparisons have been discussed with respect to the difference of two proportions by Anbar (1983), Mee (1984), Falk and Koch (1998), Soulakova and Roy (2011), and Soulakova and Bright (2013). However, no such methods have been proposed for Poisson outcomes. In this article, we propose hypothesis testing and interval-estimation methods for asymmetric comparisons in terms of the difference and ratio of two Poisson rates.

The article is outlined as follows. The problem formulation and proposed Wald-type tests are presented in Section 2.1. Approximate power of the tests and sample size computing are addressed in Section 2.2. The performance of the proposed tests is evaluated and compared to that of the standard tests for symmetric comparisons via a simulation study presented in Sections 2.3, and applications are illustrated via an example presented in Section 2.4. Confidence intervals for asymmetric comparisons, as well as an evaluation of the performance of these confidence intervals, are provided in Section 3. The article concludes with a discussion, given in Section 4.

2. Wald-Type Tests for Asymmetric Comparisons

2.1 Problem Formulation and Proposed Wald-Type Tests

Suppose that the study goal is to determine whether the experimental treatment is more beneficial than (or noninferior to) some reference treatment, when the parameter of interest is the rate of occurrences of some phenomenon, for example, incidence of a disease. This problem is commonly stated as a hypothesis testing problem for the difference or ratio of the Poisson rates (Ng, Gu, and Tang 2007; Gu et al. 2008). In particular, if the goal is to assess whether the benefits of the experimental treatment exceed the benefits of the reference treatment by some fixed margin \( d \) (\( d \) could be any real number), then the hypothesis testing problem can be stated as \( H_{0,d} : \lambda_1 - \lambda_0 = d \) with the corresponding alternative hypothesis \( H_{1,d} : \lambda_1 - \lambda_0 > d \), where \( \lambda_1 \) and \( \lambda_0 \) denote the rate, that is, the true mean number of occurrences of the event of interest per some unit of time (or space), for the experimental and reference treatments, respectively.

Another formulation corresponds to demonstrating that the benefits of the experimental treatment exceed the benefits of the reference treatment by some specified percentage, that is, the null hypothesis \( H_{0,r} : \lambda_1 = r\lambda_0 \) is tested against the alternative hypothesis \( H_{1,r} : \lambda_1 > r\lambda_0 \), where \( r \) is the specified percentage (in decimal notation), also termed margin, such that \( r > 0 \). Note that when \( d = 0 \) and \( r = 1 \), the hypotheses for the difference and ratio reduce to the same set of hypotheses.

We consider a study that has two groups. Let \( X_i \) denote the total number of occurrences in the \( i \)th treatment group such that \( X_i \sim \text{Poisson} (\lambda_{1,i}, t_i) \), where \( t_i \) is the total number of time units in the \( i \)th group. If a study deals with the number of occurrences per subject, and a subject is treated as one unit, then \( t_i \) corresponds to the group sample size. For simplicity of wording, we will refer to \( t_i \) as the \( i \)th group sample size.

To test the above hypotheses, we construct Wald-type tests that differentiate between the experimental and reference treatments in the sense that more weight is given to the information obtained from the reference treatment group rather than the experimental treatment group. Let \( E(.) \) and \( V(.) \) denote the expectation and variance, respectively. First, we consider the random variable \( W_d = X_1/t_1 - X_0/t_0 - d \). Note that under the null hypothesis \( H_{0,d}, E(W_d) = 0 \) and \( V(W_d) = \lambda_1/t_1 + \lambda_0/t_0 \) can be written as \( V(W_d) = (\lambda_0 + d)/t_1 + \lambda_0/t_0 \). Second, we consider another random variable \( W_r = X_1/t_1 - rX_0/t_0 \). Under the null hypothesis \( H_{0,r}, E(W_r) = 0 \) and \( V(W_r) = \lambda_1/t_1 + r^2\lambda_0/t_0 \) can be expressed as \( V(W_r) = r\lambda_0/t_1 + r^2\lambda_0/t_0 \).
Third, to construct a test statistic, we estimate the true rate \( \lambda_0 \) in the variances \( V_0(W_d) \) and \( V_0(W_r) \). For this purpose, the (unconstrained) maximum likelihood estimator (MLE) can be used. The MLE of \( \lambda_0 \) is given as \( \hat{\lambda}_0 = X_0/t_0 \) (Gu et al. 2008). Using this estimate, we construct the following test statistics, termed \( A \)-statistics. In the case of testing \( H_{0,d} \) versus \( H_{1,d} \), the \( A \)-statistic is

\[
A_{\text{MLE,d}} = \frac{X_1/t_1 - X_0/t_0 - d}{\sqrt{\hat{\lambda}_d \left[ \frac{1}{t_1} + \frac{1}{t_0} \right] + \frac{d}{t_1}}}
\]

Similarly, we construct the \( A \)-statistic that correspond to testing \( H_{0,r} \) against \( H_{1,r} \) as

\[
A_{\text{MLE,r}} = \frac{X_1/t_1 - rX_0/t_0}{r\sqrt{\hat{\lambda}_r \left[ \frac{1}{t_1} + \frac{1}{t_0} \right]}}
\]

A similar approach can be used to construct CMLE-based test statistics, that is, if the constrained (under the null hypothesis) maximum likelihood estimator (CMLE) for \( \lambda_0 \) is used to estimate the true rate \( \lambda_0 \) in the variance \( V_0(W_d) \) and \( V_0(W_r) \). In particular, the CMLE for \( \lambda_0 \) under \( H_{0,d} \) is given by

\[
\hat{\lambda}_d = \frac{-d}{2} + \frac{X_1 + X_0}{2(t_1 + t_0)} + \sqrt{\left( \frac{d}{2} - \frac{X_1 + X_0}{2(t_1 + t_0)} \right)^2 + \frac{X_{0,d}}{t_1 + t_0}}
\]

(Ng, Gu, and Tang 2007), which simplifies to \( \hat{\lambda}_d = \frac{(X_1 + X_0)}{(t_1 + t_0)} \) when \( d = 0 \). The CMLE for \( \lambda_0 \) under \( H_{0,r} \) is given by \( \hat{\lambda}_r = \frac{(X_1 + X_0)}{(rt_1 + t_0)} \) (Gu et al. 2008). Then the corresponding statistics are as follows.

In the case of testing \( H_{0,d} \) versus \( H_{1,d} \), the CMLE-based statistic is

\[
A_{\text{CMLE,d}} = \frac{X_1/t_1 - X_0/t_0 - d}{\sqrt{\hat{\lambda}_d \left[ \frac{1}{t_1} + \frac{1}{t_0} \right] + \frac{d}{t_1}}}
\]

In the special case of \( d = 0, A_{\text{CMLE,d}} \) simplifies to

\[
A_{\text{CMLE,d}=0} = \frac{X_1/t_1 - X_0/t_0}{\sqrt{\hat{\lambda}_d \left[ \frac{1}{t_1} + \frac{1}{t_0} \right]}}
\]

The CMLE-based statistic in the case of testing \( H_{0,r} \) against \( H_{1,r} \) is

\[
A_{\text{CMLE,r}} = \frac{X_1/t_1 - rX_0/t_0}{r\sqrt{\hat{\lambda}_r \left[ \frac{1}{t_1} + \frac{1}{t_0} \right]}}
\]

We note that, the above CMLE-based statistics for asymmetric comparisons reduce to the score statistics given by Ng, Gu, and Tang (2007) and Gu et al. (2008), as well as the corresponding tests.

Asymptotically, the null distribution of an \( A \)-statistic is standard normal. Let \( P(Z > z_\alpha) = \alpha \), where \( Z \sim \text{Normal}(0,1) \). Then the asymptotic asymmetric MLE-based test for the difference (AMD) and CMLE-based test for the difference (ACD) reject \( H_{0,d} \) in favor of \( H_{1,d} \) at level \( \alpha \) (provided that \( \lambda_i/t_i \) is sufficiently large for \( i = 0, 1 \)) if

\[
A_{\text{MLE,d}} > z_\alpha \quad \text{and} \quad A_{\text{CMLE,d}} > z_\alpha,
\]

Similarly, the asymptotic asymmetric MLE-based test for the ratio (AMR) and CMLE-based test for the ratio (ACR) reject \( H_{0,r} \) in favor of \( H_{1,r} \) at level \( \alpha \) if

\[
A_{\text{MLE,r}} > z_\alpha \quad \text{and} \quad A_{\text{CMLE,r}} > z_\alpha.
\]

Note that in the case of \( d = 0 \) and \( r = 1 \), the AMD and AMR are identical, as are the ACD and ACR. Similarly, asymptotic Wald-type normal tests based on the \( A \)-statistics can be constructed for asymmetric comparisons with other types of the alternative hypothesis.

In addition to these asymmetric tests, we consider the standard Wald MLE-based symmetric tests for the difference (NMD), discussed by Ng, Gu, and Tang (2007), and for the ratio (GMR), discussed by Gu et al. (2008). In the case of \( d = 0 \) and \( t_0 = t_1 \), the NMD and ACD are equivalent. Similarly, in the case of \( r = 1 \) and \( t_0 = t_1 \), the GMR and ACR are equivalent.

### 2.2 Power and Sample Size Computing

#### 2.2.1 On Power Approximation

The approximate power function of the proposed tests for asymmetric comparisons is derived similarly to a method used to approximate the power of normal tests for the difference of binomial proportions (Falk and Koch 1998; Soulakova and Roy 2012). In the case of testing \( H_{0,d} \) against \( H_{1,d} \) via the tests given above, this method is as follows. First, consider \( \lambda_0 = \lambda_0^d \) and \( \lambda_1 = \lambda_1^d \) such that \( d^a = \lambda_1^d - \lambda_0^d > d \). Then if both means \( \lambda_0^d \) and \( \lambda_1^d \) are sufficiently large, the distribution of \( W_d/\sqrt{V_0(W_d)} \) can be approximated by the distribution of \( W_d/\sqrt{V_0(W_d)} \), where

\[
V_0(W_d) = \frac{\lambda_0^d + d}{t_0} + \frac{\lambda_1^d}{t_0}
\]

Since the distribution of \( W_d/\sqrt{V_0(W_d)} \) under the specific alternative is asymptotically standard normal, the power function of the AMD or ACD can be approximated using

\[
P(A_d > z_\alpha | \lambda_0^a, \lambda_1^a) \approx \Phi \left( -\frac{z_\alpha \sqrt{\frac{\lambda_0^a+d}{t_0} + \frac{\lambda_1^a}{t_0} + d-d^a}}{\sqrt{\frac{\lambda_0^a+d}{t_0} + \frac{\lambda_1^a}{t_0}}} \right),
\]

where \( \Phi(.) \) denotes the cumulative distribution function of the standard normal random variable. Similarly, for given values \( \lambda_0 = \lambda_0^a \) and \( \lambda_1 = \lambda_1^a \) such that \( r^a = \lambda_1^a/\lambda_0^a > r \), the approximate power function of the
AMR and ACR is given by

\[
P(A_r > z_\alpha | \lambda_0^a, \lambda_1^a) \approx \Phi \left( -\frac{z_\alpha}{\sqrt{\frac{1}{t_0^a} + \frac{r}{m}} + \lambda_1^a (r - r^a)} \right).
\]

(2)

Note that the power functions do not take into account the specific type of estimator, MLE or CMLE, used in the test statistic.

Table 1 presents the computed power values for some specific configurations. First, we note that when the total sample size, the margin, and the rates are fixed, balanced allocations result in the highest computed power when compared with designs with unbalanced allocations in all considered settings.

Next we assess the accuracy of the proposed approximate power. For this purpose, the power is computed via (1) or (2) for each given set of the rates (under a specific alternative hypothesis), fixed value of the margin \(d\) (or \(r\)), and group sample sizes \(t_0\) and \(t_1\). Additionally, for the same setting, the power is assessed via simulations (based on 100,000 simulation runs). The discrepancy between the approximate and simulated power is computed for each test as the absolute value of the difference between the approximate and simulated power. The accuracy is assessed using the maximum discrepancy and the average (computed across all configurations) discrepancy.

Table 1 presents the results. As is illustrated, the approximate power may overestimate or underestimate the (simulated) power for each test. For the tests for the difference, the largest discrepancy between the approximate and simulated power is 1.3% for the AMD and 4.3% for the ACD. On average, the approximate power provides a better estimate of the power of the AMD (the overall average discrepancy is about 0.6%) than of the power of the ACD (the overall average discrepancy is about 1.7%). When testing for the ratio, the largest observed discrepancy is 2.8% for the ACR and 2.2% for the AMR; and the approximate power appears to be reasonably close to the power of either test: the overall average discrepancy is 0.6% for the AMR and 1.1% for the ACR. Table 1 also illustrates that the approximate power is reasonably close to the simulated power in all settings with relatively large means, that is, in settings where the total sample size is 300. Thus, for relatively large values of the means, the power of any test can be reasonably approximated using (1) and (2).

2.2.2 On Sample Size Estimation

In this section, we discuss how the approximate power in (1) and (2) can be used to derive formulas suitable for estimating the minimum required sample size under some constraints. We consider the settings where \(t_1 = k t_0\), where \(k\) is the specified allocation constant (\(k > 0\)). For given values of the parameters \(\lambda_0^a\) and \(\lambda_1^a\), significance level \(\alpha\), and desired power \(100\% - \beta\), where \(\beta\) denotes the probability (in%) of a Type II error (under \(\lambda_0^a\) and \(\lambda_1^a\)), we solve the inequality \(P \geq 100\% - \beta\) for \(t_0\), where \(P\) is the power given in the right-hand side in (1) or (2). As a result, we estimate the minimum required reference group size for the AMD (ACD) by the smallest integer exceeding \(t_0\) such that

\[
t_0^{1/2} = \frac{z_\alpha \sqrt{\frac{\lambda_0^a (k+1)+d}{k} \lambda_1^a (r - r^a)}}{d^a - d}. \tag{3}
\]

Similarly, the minimum required reference group size corresponding to the AMR (ACR) is the smallest integer exceeding \(t_0\) such that

\[
t_0^{1/2} = \frac{z_\alpha \sqrt{\frac{\lambda_0^a (k+1)+d}{k} \lambda_1^a (r - r^a)}}{(r^a - r) \lambda_1^a}. \tag{4}
\]

Following this calculation, the experimental group sample size, \(t_{1,d}\) (or \(t_{1,r}\)), is computed using \(t_{1,d} = k t_0,d\) (or \(t_{1,r} = k t_0,r\)). A similar approach can be used to derive formulas for other design settings, for example, when the total sample size is fixed at \(n\) and \(t_1 = n - t_0\). Table 2 presents the computed sample sizes using the proposed formulas for different allocation strategies and parameter values.

Proximity of the proposed sample sizes is assessed via simulations for \(\alpha = 5\%\) and power of 80%, that is, \(\beta = 20\%). First, for each set of rates, fixed value of the margin \(d\) (or \(r\)), and a specified allocation constant \(k\), the group sample sizes are computed via the proposed formulas. Then these values are used in a simulation study to determine if the actual power of a test is close to the desirable 80% level.

Table 2 presents the results. As is expected from the power study discussed in the section above, balanced allocations result in the smallest total sample size (when the other parameters are fixed). Additionally, the proposed sample size formulas (3) and (4) are more suitable for estimating the sample size requirement for the AMD rather than ACD, while these sample size formulas are appropriate for both the AMR and ACR.

2.3 Type I Error Rate and Power of the Tests

A simulation study was conducted to investigate the control of the Type I error rate and power of the four asymmetric Wald-tests (AMD, ACD, AMR, and ACR) in comparison to the Wald MLE-based symmetric tests (NMD and GMR). The specific goals were (1) among the considered settings with relatively small, moderate, and
large means, compare the tests in terms of their control of Type I error rate and identify the settings that result in adequate performance of the tests; (2) investigate and compare the power of the tests; and (3) confirm that balanced designs result in higher power than unbalanced ones (for a fixed total size, margin, parameter values, and significance level).

For each simulation configuration specified in Tables 3 and 4, the data were generated from two independent Poisson distributions, Poisson \((\lambda_i, t_i), i = 0, 1.\) Then, the AMD, ACD, and NMD, or the AMR, ACR, and GMR were performed at a 5% significance level, and the corresponding decisions in terms of accepting or rejecting the null hypothesis were noted. The above steps were repeated 100,000 times. The percentage of replications where a true null hypothesis was rejected provided the estimated Type I error rate. We say that the simulations indicate inflation of Type I error rate if the

| \(\lambda_0^a\) | \(\lambda_1^a\) | \(k\) | \(d\) | \(t_0\) | \(t_1\) | Total size | Power (AMD ACD) | \(r\) | \(\lambda_0^a\) | \(\lambda_1^a\) | \(k\) | \(d\) | \(t_0\) | \(t_1\) | Total size | Power (AMR ACR) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 9 | 1/3 | -2 | 217 | 73 | 290 | 79.9 | 79.3 | 0.8 | 198 | 66 | 264 | 80.4 | 80.0 |
| 15 | 15 | 1/3 | -2 | 165 | 83 | 248 | 79.7 | 79.0 | 1.5 | 147 | 74 | 221 | 80.8 | 80.3 |
| 15 | 45 | 10 | -2 | 114 | 114 | 228 | 79.6 | 78.7 | 1.5 | 95 | 95 | 190 | 81.1 | 80.4 |
| 45 | 15 | 7 | 72.7 | 72.0 | 70.8 | 90.0 | 91.0 | 90.7 |
| 30 | 30 | 0 | 11 | 41.3 | 83 | 290 | 79.5 | 78.7 | 1 | 95 | 95 | 190 |
| 100 | 200 | 0 | 11 | 82.2 | 81.7 | 80.8 | 80.7 | 81.8 | 80.8 |
| 200 | 100 | 81.8 | 81.5 | 81.0 | 81.0 | 81.4 | 80.9 |
| 150 | 150 | 85.7 | 85.3 | 84.7 | 84.6 | 85.3 | 84.7 |
| 30 | 30 | 12 | 77.9 | 77.0 | 75.4 | 92.6 | 92.7 | 93.5 |
| 30 | 30 | 2 | 14 | 74.8 | 74.1 | 72.7 | 73.8 | 74.5 |
| 45 | 15 | 7 | 89.2 | 88.5 | 87.1 | 99.4 | 99.8 | 99.8 |
| 100 | 200 | 34.1 | 34.5 | 32.5 | 34.8 | 34.6 | 32.7 |
| 200 | 100 | 81.8 | 81.5 | 81.0 | 81.0 | 81.4 | 80.9 |
| 150 | 150 | 85.7 | 85.3 | 84.7 | 84.6 | 85.3 | 84.7 |
| 30 | 30 | 45 | 74.8 | 74.1 | 72.7 | 73.8 | 74.5 |
| 30 | 30 | 2 | 14 | 74.8 | 74.1 | 72.7 | 73.8 | 74.5 |
| 45 | 15 | 7 | 89.2 | 88.5 | 87.1 | 99.4 | 99.8 | 99.8 |
| 100 | 200 | 34.1 | 34.5 | 32.5 | 34.8 | 34.6 | 32.7 |
| 200 | 100 | 81.8 | 81.5 | 81.0 | 81.0 | 81.4 | 80.9 |
| 150 | 150 | 85.7 | 85.3 | 84.7 | 84.6 | 85.3 | 84.7 |
| 30 | 30 | 12 | 77.9 | 77.0 | 75.4 | 92.6 | 92.7 | 93.5 |
| 30 | 30 | 2 | 14 | 74.8 | 74.1 | 72.7 | 73.8 | 74.5 |
| 45 | 15 | 7 | 89.2 | 88.5 | 87.1 | 99.4 | 99.8 | 99.8 |
| 100 | 200 | 34.1 | 34.5 | 32.5 | 34.8 | 34.6 | 32.7 |
| 200 | 100 | 81.8 | 81.5 | 81.0 | 81.0 | 81.4 | 80.9 |
| 150 | 150 | 85.7 | 85.3 | 84.7 | 84.6 | 85.3 | 84.7 |
| 30 | 30 | 45 | 74.8 | 74.1 | 72.7 | 73.8 | 74.5 |
| 30 | 30 | 2 | 14 | 74.8 | 74.1 | 72.7 | 73.8 | 74.5 |
| 45 | 15 | 7 | 89.2 | 88.5 | 87.1 | 99.4 | 99.8 | 99.8 |
| 100 | 200 | 34.1 | 34.5 | 32.5 | 34.8 | 34.6 | 32.7 |
| 200 | 100 | 81.8 | 81.5 | 81.0 | 81.0 | 81.4 | 80.9 |
observed values are at least 5.14%, where 5.14% = 5% + 1.96√5%/100, 000. The proportion of replications where a false null hypothesis is rejected provides the estimated power.

**Goal 1 Results.** Table 3 presents the results of the simulated Type I error rate. As is illustrated, all tests can potentially result in Type I error rate inflation. In addition, the CMLE-based tests (ACD and ACR) and Wald MLE-based symmetric tests (NMD and GMR) consistently result in a smaller Type I error rate than the corresponding MLE-based asymmetric tests (AMD and AMR). The CMLE-based tests and Wald symmetric MLE-based tests inflate the Type I error rate in a few settings. However, the Wald asymmetric MLE-based tests inflate the Type I error rate in several settings, even for relatively large sample sizes, and the largest Type

| (t₀, t₁) | d   | (λ₀, λ₁) | AMD  | ACD  | NMD  | r   | (λ₀, λ₁) | AMR  | ACR  | GMR  |
|----------|-----|----------|------|------|------|-----|----------|------|------|------|
| (10, 10) | −0.5| (2, 1.5) | 8.13 | 5.00 | 5.12 | 0.5 | (2, 1)   | 7.09 | 5.22 | 3.71 |
| (100, 100)| 5.87| 4.96     | 5.06 |      |      |     |          |      |      |      |
| (1000, 1000)| 5.21| 4.95     | 4.97 |      |      |     |          |      |      |      |
| (100, 100) | 5.58| 5.14     | 5.19 |      |      |     |          |      |      |      |
| (300, 300) | 5.21| 4.95     | 4.99 |      |      |     |          |      |      |      |
| (1000, 1000) | 5.10| 4.97     | 4.99 |      |      |     |          |      |      |      |
| (500, 1500)| 5.15| 4.94     | 5.08 |      |      |     |          |      |      |      |
| (1500, 500) | 5.09| 5.01     | 4.88 |      |      |     |          |      |      |      |
| (10, 10) | 0   | (2, 2)   | 7.41 | 5.04 | 5.04 | 1   | (3, 3)   | 7.09 | 4.95 | 4.95 |
| (100, 100) | 5.20| 4.98     | 4.98 |      |      |     |          |      |      |      |
| (100, 100) | 5.13| 5.02     | 5.02 |      |      |     |          |      |      |      |
| (10, 10) | 2   | (80, 82)| 5.05| 4.94 | 4.94 | 1.25| (80, 100)| 5.09 | 4.96 | 4.96 |
| (10, 10) | 4   | (2, 6)  | 6.04| 5.17 | 4.50 | 1.5| (80, 100)| 7.79 | 4.72 | 4.50 |
| (300, 300) | 5.26| 5.17     | 4.89 |      |      |     |          |      |      |      |
| (200, 400) | 5.11| 4.87     | 4.90 |      |      |     |          |      |      |      |
| (400, 200) | 5.26| 5.17     | 4.89 |      |      |     |          |      |      |      |
| (1000, 1000)| 5.08| 5.01     | 4.93 |      |      |     |          |      |      |      |
| (100, 100) | 5.10| 5.01     | 5.00 |      |      |     |          |      |      |      |
| (50, 150)  | 5.12| 4.93     | 5.06 |      |      |     |          |      |      |      |
| (150, 50)  | 5.07| 5.01     | 4.99 |      |      |     |          |      |      |      |

**Statistics in Biopharmaceutical Research: February 2015, Vol. 7, No. 1**

**Table 3.** Simulation results: Type I error rate (%) of Wald-tests (α = 5%)

**Table 4.** Simulation results: Power (%) of Wald-tests
I error rate is observed for settings with small Poisson rates.

Goal 2 Results. Table 4 presents the power results. As is expected, the power of any test highly depends on $\lambda_i t_i$, $i = 0, 1$, given that the other simulation parameters are fixed; for example, when $d = -0.5$, $\lambda_0 = 2$, and $\lambda_i = 2$, the power of any of the three tests for the difference ranges from 19% to 82% depending on the values of the group sample sizes. The power of the tests for the difference (ratio) also depends on the relative difference $\lambda_i^a - \lambda_0^a$ (ratio $\lambda_i^a/\lambda_0^a$); for example, when $r = 1.5$ and $t_0 = t_1 = 10$, the power difference can be as large as 19% when settings with $\lambda_i^a/\lambda_0^a = 6/3$ and settings with $\lambda_i^a/\lambda_0^a = 125/80$ are compared.

When the three tests for the difference (ratio) are compared in terms of the power, it can be concluded that the MLE-based asymmetric tests outperform the CMLE-based tests and Wald MLE-based symmetric tests in all settings, where the maximum discrepancy in power is 6.8% when tests for the difference are compared (and 9.2% when tests for the ratio are compared). Whether the CMLE-based tests or the corresponding Wald MLE-based symmetric tests have higher power depends on the parameter settings.

Goal 3 Results. Among different sample size allocations for the designs with the same total sample size (when the other parameters are fixed), the highest power is achieved in the balanced design cases, which agrees with our conclusions stated in Section 2.2.

2.4 Example

Boice et al. (1991) investigated the relationship between repeated exposure to X-ray fluoroscopy as part of a treatment regimen and incidence of breast cancer in women. The study involved examining the hospital records of female patients who have been treated for pulmonary tuberculosis between 1925 and 1954 in Massachusetts and determining who received lung collapse treatment using repeated X-ray fluoroscopy (exposed group) and who were unexposed to this source of radiation (unexposed group). In the sample, there are 87 cases of breast cancer over a total of 48,919 person-years observed in the unexposed group and 147 cases of breast cancer observed over a total of 56,965 person-years in the exposed group. Thus, the breast cancer incidence per year is computed as 0.00178 and 0.00258 for the unexposed and exposed groups, respectively. The authors concluded that exposure to X-ray fluoroscopy (as part of a treatment regimen) results in excess risk of developing breast cancer. This example has been widely used to illustrate inference methods for the Poisson rates (Graham, Mengersen, and Morton 2003; Ng and Tang 2005; Gu et al. 2008; Krishnamoorthy and Lee 2010; Li et al. 2011; Li, Tang, and Wong 2014).

To illustrate the application of the proposed Wald-tests, we assume that it is reasonable to treat the treatment group without X-ray fluoroscopy as the reference one. In this case, $\lambda_1$ and $\lambda_0$ denote the rates of breast cancer corresponding to the treatment with and without repeated X-ray fluoroscopy, respectively. The significance level is 5%, so we use the critical value $z_{0.05} = 1.645$.

To demonstrate the application of the methods for the asymmetric testing of the difference, we suppose that the objective of the study is to demonstrate that treatment with repeated X-ray fluoroscopy is inferior (i.e., results in a higher rate of breast cancer) to treatment without X-ray fluoroscopy, that is, $H_{0,d} : \lambda_1 - \lambda_0 = 0$ and $H_{1,d} : \lambda_1 - \lambda_0 > 0$. We compute $A_{\text{MLE},d=0} = 3.086 > 1.645$, therefore, the AMD rejects the null hypothesis. The ACD leads to the same conclusion ($A_{\text{CMLE},d=0} = 2.768$).

To demonstrate the application of the methods for the asymmetric testing of the ratio, we suppose that the objective of the study is to demonstrate that the breast cancer rate associated with repeated X-ray fluoroscopy exceeds 15% of the rate of the reference treatment, that is, $\lambda_1 > 1.15\lambda_0$. Thus, we test $H_{0,r} : \lambda_1 = 1.15\lambda_0$ versus $H_{1,r} : \lambda_1 > 1.15\lambda_0$. Since $A_{\text{MLE},r=1.15} = 1.847 > 1.645$, the AMR rejects the null hypothesis: the repeated X-ray fluoroscopy (as a part of the treatment) group breast cancer incidence exceeds 115% of the incidence of the treatment without the X-ray exposure. If we use the ACR, $A_{\text{CMLE},r=1.15} = 1.723$, we reach the same conclusion.

3. Wald-Type Confidence Intervals for the Difference and Ratio in Asymmetric Settings

The proposed tests can be used to derive the corresponding asymptotic $100(1 - \alpha)%$ one-sided confidence intervals for the parameters of interest, that is, the difference and ratio of the rates. The two-sided confidence intervals can be derived similarly. In this section, we provide the one-sided lower confidence intervals equivalent to the AMD, AMR, and ACR. The approach applied with respect to the ACD results in a very complex expression for the lower bound, and thus, it is not presented here. The Appendix illustrates the approach for constructing the asymptotic confidence interval equivalent to the AMD. Let $x_0$ and $x_1$ denote the observed values of $X_0$ and $X_1$, respectively. The resulting asymptotic $100(1 - \alpha)%$ one-sided lower interval for the difference $\lambda_1 - \lambda_0$ is given by
| $\lambda_0$ | $\lambda_1$ | $(t_0, t_1)$ | MLE-based confidence interval for the difference | MLE-based confidence interval for the ratio | CMLE-based confidence interval for the ratio |
|------------|------------|-------------|----------------------------------------|-------------------------------|----------------------------------|
| 10 | 7 | (100, 100) | 94.64 | 94.68 | 94.96 |
| | | (500, 500) | 94.88 | 94.90 | 95.02 |
| | | (3000, 3000) | **94.84** | **94.88** | **94.93** |
| 10 | 10 | (100, 100) | 94.69 | 94.69 | 95.00 |
| | | (500, 500) | 94.82 | 94.82 | 94.96 |
| | | (3000, 3000) | 95.03 | 95.03 | 95.10 |
| 13 | 10 | (100, 100) | 94.77 | 94.73 | 95.06 |
| | | (500, 500) | 94.96 | 94.93 | 95.08 |
| | | (3000, 3000) | 94.93 | 94.90 | 94.97 |
| 8 | 5 | (100, 100) | 94.42 | **94.54** | **94.87** |
| | | (500, 500) | **94.75** | **94.81** | 94.96 |
| | | (3000, 3000) | 94.92 | **94.85** | 94.91 |
| 8 | 10 | (100, 100) | **94.52** | **94.52** | 94.89 |
| | | (500, 500) | 94.88 | 94.88 | 95.03 |
| | | (3000, 3000) | **94.86** | **94.86** | 94.92 |
| 11 | 10 | (100, 100) | **94.80** | **94.66** | 95.02 |
| | | (500, 500) | 94.92 | 94.93 | 95.10 |
| | | (3000, 3000) | 95.02 | 95.02 | 95.09 |

NOTE: Values at or below 94.87% are in bold.

| $\lambda_0$ | $\lambda_1$ | $(t_0, t_1)$ | MLE-based confidence interval for the difference | MLE-based confidence interval for the ratio | CMLE-based confidence interval for the ratio |
|------------|------------|-------------|----------------------------------------|-------------------------------|----------------------------------|
| 10 | 7 | (50, 150) | 94.36 | 94.45 | 94.99 |
| | | (150, 50) | **94.84** | **94.92** | **95.09** |
| | | (250, 750) | **94.72** | **94.72** | **94.95** |
| | | (750, 250) | **94.96** | **94.96** | **95.02** |
| | | (1500, 4500) | **94.85** | **94.90** | **94.99** |
| | | (4500, 1500) | 95.03 | 95.06 | 95.09 |
| 10 | 10 | (50, 150) | **94.56** | **94.56** | 95.08 |
| | | (150, 50) | **94.80** | **94.80** | **94.98** |
| | | (250, 750) | 94.90 | 94.90 | 95.13 |
| | | (750, 250) | **94.83** | **94.83** | **94.91** |
| | | (1500, 4500) | 94.90 | 94.90 | 94.93 |
| | | (4500, 1500) | 94.90 | 94.90 | 94.94 |
| 13 | 10 | (50, 150) | 94.60 | **94.58** | **94.97** |
| | | (150, 50) | **94.78** | **94.74** | **94.92** |
| | | (250, 750) | 94.73 | 94.73 | 94.98 |
| | | (750, 250) | 94.90 | 94.88 | 94.96 |
| | | (1500, 4500) | 94.90 | 94.94 | 95.04 |
| | | (4500, 1500) | **94.86** | **94.87** | **94.91** |
| 8 | 5 | (50, 150) | **94.32** | 94.42 | 94.98 |
| | | (150, 50) | **94.73** | 94.82 | 95.00 |
| | | (250, 750) | **94.85** | **94.87** | **95.12** |
| | | (750, 250) | **94.75** | **94.82** | **94.90** |
| | | (1500, 4500) | **94.82** | **94.83** | **94.96** |
| | | (4500, 1500) | **94.86** | **94.90** | **94.93** |
| 8 | 11 | (50, 150) | **94.38** | **94.38** | 95.04 |
| | | (150, 50) | **94.87** | **94.87** | **95.06** |
| | | (250, 750) | **94.86** | **94.86** | **95.11** |
| | | (750, 250) | 94.89 | 94.89 | 94.98 |
| | | (1500, 4500) | **94.87** | **94.87** | **94.99** |
| | | (4500, 1500) | 95.04 | 95.04 | 95.09 |
| 11 | 5 | (50, 150) | **94.61** | **94.49** | **95.16** |
| | | (150, 50) | **94.74** | **94.64** | **94.90** |
| | | (250, 750) | **94.74** | **94.73** | **95.00** |
| | | (750, 250) | 94.93 | 94.92 | 95.02 |
| | | (1500, 4500) | 94.93 | 94.88 | 95.00 |
| | | (4500, 1500) | 94.97 | 94.92 | 94.96 |

NOTE: Values at or below 94.87% are in bold.
Wald-Type Testing and Estimation Methods for Asymmetric Comparisons of Poisson Rates

\[(L_{\text{MLE,d}}, \infty), \text{where} \]
\[
L_{\text{MLE,d}} = -\frac{2t_0x_1 + 2t_0x_1 + t_0z^2_a}{2t_0t_1} - \frac{1}{2} \sqrt{\frac{4t^2_1x_0z^2_a + 4t^2_0x_1z^2_a + t^2_0z^4_a}{t^2_0t_1}}.
\]

The asymptotic 100(1 - \(\alpha\))% one-sided lower intervals for the ratio \(\lambda_1/\lambda_0\) are given by \((L_{\text{MLE,r}}, \infty)\) and \((L_{\text{CMLE,r}}, \infty)\), where
\[
L_{\text{MLE,r}} = \frac{2t_0x_1 + t_0z^2_a}{2t_1(x_0 - z^2_a)} - \frac{1}{2} \sqrt{\frac{4t^2_0x_0x_1z^2_a + 4t^2_0x_1z^2_a + t^2_0z^4_a}{t^2_0x_0(x_0 - z^2_a)^2}}
\]

provided that \(x_0 > z^2\), and
\[
L_{\text{CMLE,r}} = \frac{2t_0x_0x_1 + t_0x_0z^2_a + t_0x_1z^2_a}{2t_1x^2_0} - \frac{1}{2} \sqrt{\frac{4t^2_0x_0x_1z^2_a + 4t^2_0x_0z^2_a + t^2_0x^2_0 + 2t^2_0x_0x_1z^4_a + t^2_0x^2_1z^4_a}{t^2_0x^4_0}}.
\]

To check the proposed formulas, we also used Mathematica 9.0 (Wolfram Research Inc. 2012). To illustrate the proposed confidence intervals, we use the set-up of the example presented in Section 2.4, and construct the asymptotic one-sided 95% confidence intervals. The lower confidence bound for the difference is computed as \(L_{\text{MLE,d}} = 0.0004\), which is greater than zero. Thus, based on the confidence interval we reject the null hypothesis. The MLE-based lower confidence bound for the ratio is \(L_{\text{MLE,r}} = 1.178\) and the CMLE-based bound is \(L_{\text{CMLE,r}} = 1.162\), which both exceed 1.15. Therefore, each inference is consistent with the one based on the corresponding test.

A simulation study was conducted to evaluate the coverage probability of the asymmetric confidence intervals for the difference and ratio of two Poisson rates. These simulations were performed for several rate settings, which include settings with relatively small rates (\(\lambda_0 \leq \lambda_1\) and \(\lambda_0 > \lambda_1\)) and various group sample sizes, including balanced and unbalanced settings. The simulation configurations are specified in Tables 5 and 6. For each configuration, random data were generated similarly to the study discussed in Section 2.3. Then the proposed 95% lower confidence intervals were constructed. An estimate of the true coverage probability for each approach was calculated as the percentage of replications where the confidence interval contained the true parameter value (difference of rates, \(\lambda_1 - \lambda_0\), or ratio of rates, \(\lambda_1/\lambda_0\)). Taking into account the simulation error, an interval-estimating method was said to result in under-coverage if the observed probability coverage was 94.87% or below.

Tables 5 and 6 present the coverage probability results for balanced and unbalanced designs, respectively. The simulations do not detect under-coverage for the CMLE-based interval for the ratio in almost all settings. The MLE-based intervals, however, result in under-coverage in settings where at least one group sample size is not sufficiently large, for example, less than 500. The extent of under-coverage does not seem to be more pronounced in balanced settings when compared to the unbalanced ones (or vice versa) with the same total sample size (and other simulation parameters). All methods tend to perform satisfactorily when the sample sizes and means are sufficiently large.

4. Discussion

This article, proposes several asymptotic Wald-type tests and interval-estimation methods for asymmetric comparisons, that is, settings where one treatment (reference one) serves as the standard treatment. The asymmetric tests reflect this: the variance of the point estimator for the difference or ratio of two Poisson rates (under the null hypothesis) incorporates only the rate corresponding to the reference treatment. The standard error is then computed using one of two commonly used estimators for the Poisson, that is, the MLE or CMLE. The CMLE-based asymmetric tests, the ACD and ACR, are shown to be equivalent to the standard (i.e., symmetric) tests proposed by Ng, Gu and Tang (2007) and Gu et al. (2008). The asymmetric tests, that is, the AMD, ACD, AMR, and ACR, are assessed in terms of Type I error rate and power. The advantages of these tests over the standard MLE-based tests (Ng, Gu, and Tang 2007; Gu et al. 2008) are discussed. The formulas for estimating the approximate power and sample size are provided. In addition, the interval-estimating methods for the difference and ratio of the rates are proposed.

The proposed tests for asymmetric and symmetric comparisons are large-sample tests and are only expected to perform well when sample sizes are sufficiently large. In fact, we have detected several settings with relatively small values of the rates and/or sample sizes where the proposed MLE-based tests inflated the Type I error rate, for example, when testing for the difference with \(d = -2\) and \(t_0 = t_1 = 300\), settings with \(\lambda_0 = 8\) and \(\lambda_1 = 6\) resulted in an inflated Type I error rate, while similar settings with \(\lambda_0 = 80\) and \(\lambda_1 = 78\) did not result in any inflation. However, for sufficiently large values of the rates, all large-sample tests satisfactorily control the Type I error rate. In these cases, the proposed MLE-based tests are more powerful than the corresponding CMLE-based tests and the standard MLE-based tests. The proposed
formulas for estimating the power and minimum sample size can be used in practice in the design stage to obtain additional information regarding the anticipated test performance. For a fixed total sample size, the tests are expected to perform with the highest power in balanced settings, although there could be some exceptions.

We also presented Wald-type confidence interval methods that can be used for asymmetric comparisons. Similar to the conclusions stated in terms of the Type I error rate control, the MLE-based confidence interval methods may result in under-coverage when estimating the difference or ratio of the rates. The CMLE-based confidence interval method performs reasonably well.

**Supplementary Material**

*Appendix: Constructing the Asymptotic Confidence Interval Equivalent to the Wald-MLE Asymmetric Test for the Difference of Two Poisson Rates*

The Appendix illustrates the approach for using the proposed Wald-MLE asymmetric test to derive the corresponding asymptotic 100(1 − α)% one-sided confidence intervals for the difference of two rates.

[Received May 2014. Revised September 2014]

**References**

Anbar, D. (1983), “On Estimating the Difference Between two Probabilities, with Special Reference to Clinical Trials,” *Biometrics*, 39, 257–262. [2]

Barker, L. (2002), “A Comparison of Nine Confidence Intervals for a Poisson Parameter when the Expected Number of Events is ≤ 5,” *The American Statistician*, 56, 85–89. [1]

Barker, L., and Cadwell, B. L. (2008), “An Analysis of Eight 95 Per Cent Confidence Intervals for a Ratio of Poisson Parameters when Events are Rare,” *Statistics in Medicine*, 27, 4030–4037. [1]

Boice, J. D., Preston, D., Davis, F. G., and Monson, R. R. (1991), “Frequent Chest X-ray Fluoroscopy and Breast Cancer Incidence Among Tuberculosis Patients in Massachusetts,” *Radiation Research*, 125, 215–222. [7]

Bright, B. C., and Soulakova, J. N. (in press), “Evaluating the Performance of Simultaneous Stepwise Confidence Intervals for the Difference Between two Poisson Rates,” *Communications in Statistics—Simulation and Computation*. [1]

Brown, L. D., Cai, T., and Dasgupta, A. (2001), “Interval Estimation for a Binomial Proportion” (with discussion), *Statistical Science*, 16, 101–133. [1]

Byrne, J., and Kabaila, P. (2005), “Comparison of Poisson Confidence Intervals,” *Communications in Statistics—Theory and Methods*, 34, 545–556. [1]

Colditz, G. A., Stampfer, M. J., Willett, W. C., Hennekens, C. H., Rosner, B., and Speizer, F. E. (1990), “Prospective Study of Estrogen Replacement Therapy and Risk of Breast Cancer in Post-menopausal Women,” *Journal of the American Medical Association*, 264, 2648–2653. [2]

Falk, R. W., and Koch, G. G. (1998), “Testing a Specified Difference Between Proportions,” *Biometrics*, 54, 1602–1614. [2,3]

Graham, P. L., Mengersen, K., and Morton, A. P. (2003), “Confidence Limits for the Ratio of two Rates Based on Likelihood Scores: Non-iterative Method,” *Statistics in Medicine*, 22, 2071–2083. [2,7]

Gu, K., Ng, H. K. T., Tang, M. L., and Schucany, W. R. (2008), “Testing the Ratio of Two Poisson Rates,” *Biometrical Journal*, 50, 283–298. [1,2,3,7,9]

Hall, P. (1982), “Improving the Normal Approximation when Constructing One-Sided Confidence Intervals for Binomial or Poisson Parameters,” *Biometrika*, 69, 647–652. [1]

Hargarten, S. W., and Baker, S. P. (1985), “Fatalities in the Peace Corps: A Retrospective Study: 1962 Through 1983,” *Journal of the American Medical Association*, 254, 1326–1329. [2]

Johnson, N., Kemp, A., and Kotz, S. (2005), *Univariate Discrete Distributions* (3rd ed.), Hoboken, NJ: Wiley-Interscience. [1]

Khamkong, M. (2012), “Approximate Confidence Interval for the Mean of Poisson Distribution,” *Open Journal of Statistics*, 2, 204–207. [1]

Krishnamoorthy, K., and Lee, M. (2010), “Inference for Functions of Parameters in Discrete Distributions Based on Fiducial Approach: Binomial and Poisson Cases,” *Journal of Statistical Planning and Inference*, 140, 1182–1192. [1,7]

——— (2013), “New Approximate Confidence Intervals for the Difference between two Poisson Means and Comparison,” *Journal of Statistical Computation and Simulation*, 83, 2232–2243. [1]

Li, H. Q., Tang, M. L., Poon, W. Y., and Tang, N. S. (2011), “Confidence Intervals for Difference Between two Poisson Rates,” *Communications in Statistics—Simulation and Computation*, 40, 1478–1493. [1,7]

Li, H. Q., Tang, M. L., and Wong, W. K. (2014), “Confidence Intervals for Ratio of two Poisson Rates using the Method of Variance Estimates Recovery,” *Computational Statistics*, 29, 869–889. [1,7]

Liu, G. F., Wang, J., Liu, K., and Snayele, D. B. (2006), “Confidence Intervals for an Exposure Adjusted Incidence Rate Difference with Applications to Clinical Trials,” *Statistics in Medicine*, 25, 1275–1286. [1]

Liu, K. J. (2004), “Eight Interval Estimators of a Common Rate Ratio under Stratified Poisson Sampling,” *Statistics in Medicine*, 23, 1283–1296. [1]

Mee, R. W. (1984), “Confidence Bounds for the Difference Between two Proportions,” *Biometrics*, 40, 1175–1176. [2]

Miettinen, O., and Nurminen, M. (1985), “Comparative Analysis of two Proportions,” *Statistics in Medicine*, 4, 213–226. [1]

Nelson, W. (1970), “Confidence Intervals for the Ratio of two Poisson Means and Poisson Predictor Intervals,” *IEEE Transactions on Reliability*, R-10, 42–49. [1]

Ng, H. K. T., Gu, K., and Tang, M. L. (2007), “A Comparative Study of Tests for the Difference of two Poisson Means,” *Computational Statistics and Data Analysis*, 51, 3085–3099. [1,2,3,9]

Ng, H. K. T., and Tang, M. L. (2005), “Testing the Equality of two Poisson Means using the Rate Ratio,” *Statistics in Medicine*, 24, 955–965. [1,7]

Patil, V. V., and Kulkarni, H. V. (2012), “Comparison of Confidence Intervals for the Poisson Mean: Some New Aspects,” *REVSTAT—Statistical Journal*, 10, 211–227. [1]
Pearson, E. S., and Hartley, H. O. (eds.) (1966), *Biometrika Tables for Statisticians* (3rd ed., Vol. 1), Cambridge: Cambridge University Press. [1]

Pitt, B., Segal, R., Martinez, F. A., Meurers, G., Cowley, A. J., Thomas, I., Deedwania, P. C., Ney, D. E., Snaively, D. B., and Chang, P. I. (1997), “On Behalf of ELITE Study Investigators. Randomised Trial of Losartan versus Captopril in Patients over 65 with Heart Failure (Evaluation of Losartan in the Elderly Study, ELITE),” *Lancet*, 349, 747–752. [2]

Price, R. M., and Bonett, D. G. (2000), “Estimating the Ratio of two Poisson Rates,” *Computational Statistics and Data Analysis*, 34, 345–356. [1]

Rochon, J., Aprile, M. A., and Cardella, C. J. (1992), “Analyzing the Number of ‘Rejection Episodes’ in Renal Transplant Studies,” *Transplantation*, 53, 1236–1242. [1]

Sahai, H., and Khurshid, A. (1993), Confidence Intervals for the Ratio of Two Poisson Means,” *The Mathematical Scientist*, 18, 43–50. [1]

Schwertman, N., and Martinez, R. (1994a), “Approximate Confidence Intervals for the Difference in two Poisson Parameters,” *Journal of Statistical Computation and Simulation*, 50, 235–247. [1]

——— (1994b), “Approximate Poisson Confidence Limits,” *Communications in Statistics—Theory and Methods*, 23, 1507–1529. [1]

Soulakova, J. N., and Bright, B. C. (2013), “Applications of Asymptotic Confidence Intervals with Continuity Corrections for Asymmetric Comparisons in Non-inferiority Trials,” *Pharmaceutical Statistics*, 12, 147–155. [2]

Soulakova, J. N., and Roy, A. (2011), “Application of Anbar’s Approach to Hypothesis Testing to Detect the Difference Between Two Proportions,” *Communications in Statistics—Theory and Methods*, 40, 1866–1878. [2]

——— (2012), “On Power Approximations and Comparison of Several Asymptotic Tests to Detect a Specified Difference Between Two Proportions,” *Communications in Statistics—Simulation and Computation*, 41, 376–390. [3]

Stampfer, M. J., and Willett, W. C. (1985), “A Prospective Study of Postmenopausal Estrogen Therapy and Coronary Heart Disease,” *New England Journal of Medicine*, 313, 1044–1049. [2]

Swift, M. B. (2009), “Comparison of Confidence Intervals for a Poisson Mean—Further Considerations,” *Communications in Statistics—Theory and Methods*, 38, 748–759. [1]

Tang, M. L., and Ng, H. K. T. (2004), Comment on “Confidence Limits for the Ratio of two Rates based on Likelihood Scores: Non-Iterative Method,” *Statistics in Medicine*, 23, 685–693. [1]

Weissfeld, L., St. Laurent, R., and Moulto, L. (1991), “Confidence Intervals for the Comparison of Two Poisson Distributions,” *Communications in Statistics—Theory and Methods*, 20, 3071–3081. [1]

Wolfram Research, Inc. (2012), *Mathematica, Version 9.0*, Champaign, IL: Wolfram Research, Inc. [9]

**About the Authors**

Brianna C. Bright (E-mail: Brianna.Bright@gmail.com) is Graduate Student and Julia N. Soulakova (E-mail: jsoulakova2@unl.edu) is Associate Professor, Department of Statistics, University of Nebraska-Lincoln, 340 Hardin Hall-North, Lincoln, NE 68583-0963.