Practice of Epidemiology

Analyses of Injury Count Data: Some Do’s and Don’ts

Ian Shrier*, Russell J. Steele, James Hanley, and Benjamin Rich

* Correspondence to Dr. Ian Shrier, Centre for Clinical Epidemiology and Community Studies, SMBD-Jewish General Hospital, 3755 Cote Sainte Catherine Road, Montreal, Quebec H3T 1E2, Canada (e-mail: ian.shrier@mcgill.ca).

Initially submitted April 16, 2009; accepted for publication August 5, 2009.

The analysis of injury data requires different considerations from the analysis of other types of outcomes because an individual can experience the outcome many times. When describing injury patterns using numerator-only data (e.g., proportion of upper-extremity injuries vs. lower-extremity injuries), simple comparisons of proportions are inappropriate because 1) individuals are compared with themselves and 2) multiple testing increases the potential for incorrect inference. Bootstrapping (resampling) techniques can be used to determine confidence intervals and whether the frequencies significantly differ across categories. When describing injury rates, the authors suggest plotting the observed injury rate against the number of exposures to obtain a visual representation of the heterogeneity of risk across individuals. Because the distribution of injury rates is often skewed, some research questions may be best addressed by comparing the weighted median injury rates instead of the weighted mean injury rates (which are given by standard formulae). Again, resampling techniques can be used to obtain a null distribution for injury rates in order to determine whether there are subjects who have unexpectedly high injury rates. More advanced analyses are required to account for multiplicity.

Developing effective injury prevention programs requires an understanding of risk factors (1–7). Determining risk factors for injury is fundamentally different from that for many other conditions because injury represents multiple-event data, where a subject can have the “outcome” many times. Therefore, these data are best analyzed as “count data” (8, 9). Within some contexts, only injury counts are available (e.g., falls in the elderly); in other contexts, both counts and exposure data (e.g., person-time) are available.

Counts without exposure data still provide important information for prevention programs (e.g., high ski boots were introduced to prevent ankle fractures in alpine skiing). Where exposure data are available, one can calculate rate ratios and differences using standard formulae or regression (e.g., Poisson or negative binomial regression), or one can calculate them directly from the data distribution.

Each of the above methods can affect data interpretation. Our 2 objectives in this paper are to 1) provide an overview of the advantages/disadvantages of different analytical methods and 2) illustrate methods that appropriately describe the data and the heterogeneity within the data and lead to improved inferences. In our analyses, we have treated “injury” as the event, but a person can injure both the ankle and the shoulder during a single fall. For some research questions, it is more appropriate to treat the “fall” as the event rather than the injury. Our generic approach is applicable to whatever definition of “event” is used. Further, analysis of exacerbations/reinjuries requires careful consideration on many levels (e.g., healing has been defined as return to full participation (10) and date of last treatment (11)). Although we have used injury data in our example, most of the principles are applicable to other forms of count data (e.g., visits to a health-care professional).

METHODS

These analyses are based on work-related injury data from 2002–2006 obtained from Cirque du Soleil (Montreal, Quebec, Canada), and the methods are described in detail elsewhere (11). Below, we discuss issues related to numerator-only data (exposure data not available), which is common when data are obtained from recreational sports, hospital reports, or database studies. We first discuss frequently

1307 Am J Epidemiol 2009;170:1307–1315
used methods and then illustrate how appropriately accounting for multiple testing and repeated measures in the same individual affects interpretation. We then discuss rate data and compare the results from 1) standard population-based formulae, 2) regression formulae, and 3) the distribution of rates calculated per subject. In addition, we present different methods for exploring data heterogeneity.

We used the R software program (R Foundation for Statistical Computing; http://www.R-project.org) for all analyses.

**NUMERATOR-ONLY DATA**

*Are some injuries more common than others?*

Figure 1 demonstrates the types of injuries that occurred in our data according to anatomic location; this type of figure is common in injury studies (12). Within the data, 1 artist may contribute more than 1 injury to a single category and may contribute injuries to multiple categories. The distribution obtained is only a finite sample of a larger generic universe, and investigators may want to know whether the differences in frequencies are likely to be observed in other studies. We illustrate both the inappropriate and appropriate approaches.

Given that Figure 1 is based on a $1 \times 5$ table (frequency of injuries by location), one might use the Pearson chi-squared ($\chi^2$) test (the comparison of proportions test is inappropriate because it requires at least 2 rows and 2 columns). The expected frequency for each cell is the total number of injuries divided by the number of categories. Our Cirque du Soleil data included 18,336 injuries, or 3,056 expected injuries per category ($\chi^2 = 8,138.9$, $P < 2.2 \times 10^{-16}$). The $P$ value is extremely low because of the large numbers of injuries. Table 1 provides examples of less extreme distributions for 100 injuries.

Although the $\chi^2$ test is simple, the underlying assumption of the $\chi^2$ test is that the events (injuries) are independent; this is inappropriate for our data, because 1 Cirque du Soleil subject could contribute multiple injuries to a particular category or could contribute information to multiple categories. Further, there are no simple methods for calculating the appropriate variances and confidence intervals. We therefore recommend bootstrapping as a generic method of determining the confidence interval. In brief, assume that we have data on 100 subjects with a total of 500 injuries. In the bootstrap approach, one first chooses one of the 100 subjects at random and records the injuries. The process is repeated as many times as there are subjects (i.e., 100 times in this example), and each time, the random choice is one of the full list of 100 subjects (i.e., a subject can be chosen several times). This list of 100 “subjects” (in which some subjects are duplicated) represents 1 sample of the bootstrap method. The process is then repeated to obtain many samples (typically $n = 1,000$) (13). Different methods can be used to calculate bootstrap confidence intervals (13). In this paper, we report the confidence intervals based on the percentiles of the distribution of percentages generated by the 1,000 data sets created (a nonparametric bootstrap). To obtain appropriate confidence intervals with injury data, it is important to bootstrap by the individual to appropriately account for recurrent injuries in the

---

**Table 1. Potential Injury Patterns in 6 Anatomic Locations for a Data Set That Includes a Total of 100 Injuries**

<table>
<thead>
<tr>
<th>Anatomic Location</th>
<th>No. of Injuries</th>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
<th>Example 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head and neck</td>
<td>20</td>
<td>20</td>
<td>25</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Trunk</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Spine</td>
<td>15</td>
<td>20</td>
<td>15</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Upper extremity</td>
<td>15</td>
<td>20</td>
<td>15</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Lower extremity</td>
<td>15</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>15</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Pearson $\chi^2$ test</td>
<td>$\chi^2$</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>$P$ value</td>
<td>0.85</td>
<td>0.16</td>
<td>0.16</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

---

$^a$ Analyses were based on work-related injury data for 2002–2006 obtained from Cirque du Soleil (Montreal, Quebec, Canada) (11).
Figure 2. Cumulative probability of observing a value less than or equal to the value shown on the x-axis using resampling methods based on the null distribution. The horizontal and vertical dotted lines represent the probabilities 0.025 and 0.975. The vertical dotted lines (and gray numbers on the x-axis) illustrate the 95% sampling interval for the percentage of injuries in any 1 category under the null hypothesis.

The same subject; bootstrapping by the injury would reproduce the error.

To determine whether all anatomic locations are equally likely to be injured, one can compare the observed pattern to a pattern created from a data set where all locations are equally likely, but is still based on the original data as a foundation. For example, subject 1 has 10 injuries. Under the null hypothesis, the expected numbers of injuries for all categories are identical (the number of injuries for that individual divided by the number of categories). We can therefore use the multinomial probability model to calculate how many injuries this subject is likely to have in each category, assuming that the observed total number of injuries will depend on the total number of injuries within each category. One then repeats this process for each individual to obtain a single complete data set. The pattern of injuries generated is based on multinomial probabilities, and one would expect the pattern to vary if the simulation were conducted a second time (similar to observed injury patterns for an individual over different time periods even if the “true” injury pattern remained fixed). To account for this “random variation,” one repeats the process many times (typically 1,000 times) and thus creates 1,000 data sets (“random” data sets) generated under the null hypothesis of a uniform distribution of injury rates.

The principles for calculating the uncertainty due to the random variation (typically called sampling intervals) are similar to those for confidence intervals (but the two should not be confused (14)). Figure 2 is an illustration of the principle. In brief, the distribution of the proportions from the 1,000 data sets is plotted as a cumulative probability function. In our example, the 95% sampling interval for the proportion of injuries within each category was 16.1%–17.2%.

The resulting 95% sampling intervals for the proportion of injuries will depend on the total number of injuries within the data set. For example, subjects 1–20 had a total of 462 injuries, and the 95% sampling interval for the proportion of injuries at a particular anatomic location (all locations being equally likely under the null hypothesis) was 13.0%–20.1%. However, subjects 101–120 had only 210 injuries, and the 95% sampling interval for the proportion of injuries in a category was wider: 11.4%–21.9%. When we used 40 subjects with the same total number of injuries as the first 20 subjects, we obtained an interval of similar width (13.4%–20.1%) to that of subjects 1–20. These intervals correspond to what would be expected under a large-sample approximation to the distribution of the observed proportion (13.1%–20.2%). However, in situations with a large number of locations relative to the total number of injuries, sampling from the null distribution will perform better than the normal approximation.

Finally, the above analysis can be linked with the data in Table 1. We created simulated data by drawing a total number of injuries per subject from a Poisson distribution and then distributing the injuries to categories via a multinomial distribution. We simulated data for 50 subjects with an expectation of 2 injuries per subject (generated total number of injuries = 98) and 6 categories of possible injury. The 95% sampling interval for the proportion of injuries in a category was 9.2%–24.5%. In examples 1–3 of Table 1, the proportion in every category lies within or very close to the sampling interval and would not be expected to be statistically significant. However, the “other” category in example 4 is outside of the 95% sampling interval.

The above method essentially examines whether the different injuries are independent of each other. Therefore, deviations from the null distribution will occur if 1) an activity causes specific types of injuries, 2) injuries are correlated (e.g., head and spine injuries often occur together), or 3) reinjuries to the same location are more likely to occur than new injuries to other locations.

There is another informative analysis based on the square of the ratio of the 95% confidence interval range (upper limit of the 95% confidence interval of the observed data minus lower limit) to the 95% sampling interval range (upper limit of the 95% sampling interval of the null distribution data minus lower limit). In our example, the 95% confidence interval range for the “head and neck” category was 1.23%, and the 95% sampling interval range was 1.08%. The square of the ratio of these 2 numbers (1.30 in our
example) is a measure of the amount of dependence of the data (15). Therefore, after accounting for intrasubject correlations, a rough estimate of the “effective sample size” is 76.9% (1/1.3). Stated more clearly, the data contain approximately the same amount of independent information as one would have in a hypothetical sample that was 76.9% of the size of the original but where each subject had only 1 record.

We stress that in order to calculate the appropriate sampling intervals, one should create the simulated data using the number of injuries per subject observed in the actual data because the sampling interval depends on the total number of injuries. We also caution that as the number of categories considered increases, the actual probability of having at least 1 true value outside of the 95% sampling interval for one of the categories will rise above 5%. For example, with 2 injury categories and no correction for multiplicity, the simple 95% confidence interval (e.g., ±2%) underestimates the true 95% confidence interval (i.e., ±2.3%) by 13.0% (0.3/2.3); with 5 categories, the simple 95% confidence interval underestimates the true 95% confidence interval (±2.6%) by 23.1% (0.6/2.6).

Do injury patterns differ in different types of subjects?

In Figure 3, injuries are plotted for males and females. To evaluate the evidence that the injury patterns between the 2 sexes occurred by chance, neither the comparison of proportions test nor the $\chi^2$ test (for a $2 \times K$ matrix) is appropriate because events remain dependent (each subject may contribute to more than 1 category). To calculate 95% confidence intervals for each sex for each category, bootstrapping is conducted on male and female data separately.

In this context, our approach is limited to inferences based on comparing point estimates and confidence intervals. One can obtain appropriate $P$ values for testing a null hypothesis, although there is some disagreement about the optimal approach (16). From a practical perspective, implementation of Monte Carlo algorithms to calculate appropriate $P$ values can be quite involved, particularly for nonstatisticians (16).

When the same subjects are being compared in 2 different contexts (e.g., injury patterns in athletes during training vs. games), the analysis is slightly more complicated, but bootstrapping techniques remain a simple solution for obtaining appropriate confidence intervals.

RATE DATA

Distribution of injury rates

When exposure data are available, the overall injury rate is often expressed per 1,000 units of person-time (8):

$$\frac{\sum \text{Cases}}{\sum \text{Person-Time}}$$

Equation 1 is equivalent to the weighted mean injury rate across subjects. The expanded formula for the weighted mean is shown below (equation 2), where the ratio on the
left represents a subject’s injury rate and the ratio on the right represents the weight for the subject; the “Person-Time,” terms in the numerator and denominator simply cancel each other out to yield equation 1.

\[
\sum \left( \frac{\text{Cases}_i}{\text{Person-Time}_i} \times \frac{\text{Person-Time}_i}{\sum \text{Person-Time}_i} \right). \tag{2}
\]

Using a weighted injury rate is important, because subjects with few exposures are expected to have a high variability in injury rates and therefore should not contribute as much information towards an overall group injury rate as subjects who have many exposures. The formula for the standard error (8) is

\[
\sqrt{\frac{\sum \text{Cases}}{\sum \text{Person-Time}}}. \tag{3}
\]

Is a weighted mean injury rate the best way to describe these data? Figure 4 shows a histogram of the injury rates in our data for injury rates of less than 30 injuries per 1,000 artist exposures. This histogram is highly skewed, and a median value would normally be preferred over a mean value to describe the center of a skewed distribution. One can calculate a weighted median (or any percentile) using the same principles as those used for the weighted mean; subjects with fewer exposures contribute less information to the overall weighted value.

The weighted mean and the weighted median both provide valuable information. The weighted mean injury rate is valuable because it allows one to calculate the consequences of injury in terms of workload or cost more easily (e.g., total injuries = weighted mean injury rate \times total number of subjects). However, because the data distribution is skewed, the individual injury rates are better summarized by the weighted median. For example, the weighted median in Table 2 shows that 50% of the artists had 7.8 injuries per 1,000 artist exposures or less.

Table 2 summarizes different ways of presenting injury rates along with the appropriate measures of variability.

### Table 2. Injury Rates (All Injuries Combined) for Cirque du Soleil Artists, Overall and by Sex, Calculated Using Several Different Methods

<table>
<thead>
<tr>
<th>Total</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate calculated by standard formula (8)</td>
<td>9.7 (9.4, 10.0)</td>
<td>10.2 (9.7, 10.7)</td>
</tr>
<tr>
<td>Weighted mean(^a) of individual injury rates (SD)</td>
<td>9.7 (8.8)</td>
<td>10.2 (13.4)</td>
</tr>
<tr>
<td>95% CI(^b) based on standard errors</td>
<td>9.5, 9.9</td>
<td>10.0, 10.3</td>
</tr>
<tr>
<td>95% CI(^d) based on bootstrap</td>
<td>9.1, 10.3</td>
<td>9.2, 11.2</td>
</tr>
<tr>
<td>Weighted median(^b) of distribution (IQR)</td>
<td>7.8 (11.5(^e))</td>
<td>8.8 (11.6)</td>
</tr>
<tr>
<td>95% CI(^d) based on bootstrap</td>
<td>7.0, 8.7</td>
<td>7.5, 10.1</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; IQR, interquartile range; SD, standard deviation.

\(^a\) Analyses were based on work-related injury data for 2002–2006 obtained from Cirque du Soleil (Montreal, Quebec, Canada) (11).


\(^d\) Calculated using the BSDA or PASWR package in R.

\(^e\) 75th percentile minus 25th percentile.

---

*Am J Epidemiol* 2009;170:1307–1315
Figure 5. Injury rates for individual Cirque du Soleil artists (open circles) according to the number of exposures for that artist, 2002–2006. For clarity, only data for ≤60 injuries per 1,000 artist exposures are shown. There were 3 males (out of 618) and 1 female (out of 348) with injury rates greater than 60 injuries per 1,000 artist exposures (maximum injury rate = 200 injuries per 1,000 artist exposures). The solid line represents the weighted median injury rate, and the dotted line represents the weighted mean injury rate. The apparently curved pattern of the data points occurs because a particular injury rate is not possible in the Poisson distribution at certain exposure values, and it then becomes possible at particular discrete values. Making inferences about heterogeneity requires the addition of sampling intervals (see Figure 6).

Exploring heterogeneity

The histogram in Figure 4 shows the distribution in rates overall and by sex. Note the discrepancies between 95% confidence intervals calculated by different methods. In brief, the standard method in equation 3 assumes a Poisson distribution (mean of the distribution = variance of the distribution). When there is increased heterogeneity (variance > mean, known as overdispersion), the formula underestimates the confidence interval. When overdispersion exists, negative binomial regression is often used. However, negative binomial regression makes an assumption of a particular distribution for the heterogeneity of the individual-specific rates, which must also be assessed because over-/underdispersion could still remain a problem. The bootstrap method does not make assumptions about the underlying distribution of the population individual-specific rates or their heterogeneity. For modeling purposes, one could use 1) the bootstrap in conjunction with a negative binomial distribution, if appropriate, or 2) the quasi-Poisson regression, which does not assume an underlying distribution. Although both of these alternatives are complicated and require statistical expertise, they are nonetheless necessary for appropriate inferences when the data are over-/underdispersed.

Figure 5 is a plot of the injury rate against the number of exposures for injury rates between 0 and 60 injuries per 1,000 artist exposures. Even here, risk assessment for individual subjects must be interpreted cautiously. For example, random chance dictates that some subjects may fall into the highest tertile even if they do not have an underlying increased risk of injury. In addition, some readers might believe that the data in Figure 5 strongly suggest that subjects with less experience (fewer exposures) have a higher risk of injury, but below we show this to be incorrect. For data focusing on only the time to first injury for each artist, one could perform a more sophisticated analysis which attempts to estimate both the population mean injury rate and the heterogeneity of the distribution of rates simultaneously by using a beta-geometric model (17), but that approach is beyond the scope of this paper.

A simpler but approximately correct analysis is to superimpose expected sampling intervals for injury rates across all levels of exposure using an estimate of the population probability distribution (Figure 6). The distribution is based on the number of exposures and the injury heterogeneity of the observed data (similar to what we suggested for patterns of injury) so that it reflects the correct uncertainty in observed rates. There are several steps in this process.

Step 1. According to the hypothesized model that generated the data, the overall weighted mean injury rate should be the true weighted mean injury rate for all subjects. Therefore, the expected number of injuries for each subject is the overall weighted mean injury rate multiplied by that subject’s number of exposures.

Step 2. The probability for the lowest and highest 2.5% of the Poisson distribution represents the approximate 95% sampling interval for injury rates, assuming that all variation
in the rates occurred simply by chance. These values can be superimposed on the previous graph to determine which of the observed points fall outside of the 95% sampling intervals. Although subjects with injury rates outside of these boundaries have extreme values relative to what one would expect under a hypothesis of no variation in rates, the boundaries do not quite represent 95% regions, as explained in step 3.

**Step 3.** There are 3 reasons why the 95% sampling intervals calculated above do not represent true 95% sampling regions.

1. In step 1, we assumed that the true weighted mean injury rate for all individuals was fixed and that there was no variation. However, groups of individuals will have weaknesses or strengths that would make their true injury rate different from that of other groups of subjects. Therefore, it is likely that there is more than 1 “true injury rate” for the group, and one could account for this uncertainty as well (in statistical terms, this means that the injury rate follows a random-effects model). Because our data included over 6,700 injuries, the standard error of the weighted mean injury rate (i.e., incorporating the uncertainty of the true injury rate) was only 3%. Therefore, the results derived from the random-effects and fixed-effects models were very similar to each other. However, the results from the 2 models would be expected to differ if the total number of subjects were smaller, the weighted mean injury rate were lower, and/or there were more heterogeneity between subjects.

2. The 95% sampling intervals reflect probabilities at each level of exposure, and this must be incorporated when making inferences. For example, if there were 100 points at each exposure level, one would expect 5 of the points to lie outside of the interval at that exposure level. Although it might appear that one could simply expect 5% of all points to lie outside of the 95% sampling bands, this is not actually exactly true. Such an interpretation would require a more sophisticated analysis to create 95% probability regions over the range of exposure values (for a recent review, see Uusipaikka (18)).

3. An alternative to the 95% probability regions is to adjust the perceived significance of the values outside of the pointwise intervals via some false discovery rate procedure. These types of procedures range from the relatively computationally simple (19) to the more complex (20, 21). See the article by Pounds (22) for a nontechnical review of these techniques.

### Comparing rates

Rate ratios and rate differences are often used to compare different subgroups (e.g., males vs. females). For rate ratios, one can use the standard formula (8), weighted means of individual artist rates, weighted medians of individual artist rates, or Poisson regression (or some related model such as negative binomial regression, depending on the data distribution). Poisson regression also allows one to adjust for other potentially confounding variables. For rate differences, one can use the standard formula, the weighted mean, or weighted medians. In Table 3, we summarize the results derived from the use of these different methods.

For rate ratios, the point estimates for the standard formula, the Poisson regression, and the individual weighted-mean artist rates are the same, but the 95% confidence
intervals are too narrow because they incorrectly assume no standard formula and the Poisson regression confidence intervals are too narrow because they incorrectly assume no overdispersion. Note that this is not due simply to approximation error of the standard formulae, because our sample was sufficiently large to avoid that problem (23). The bootstrap intervals are nonparametric and provide the appropriate 95% confidence interval (a quasi-Poisson regression model also provided the same intervals as the bootstrap method, and the dispersion parameter was taken to be 3.899). We also tried negative binomial regression. First, a few artists had very high injury rates, which were larger than expected for a negative binomial distribution (gamma distribution). With the use of more advanced statistical techniques (we assumed that the observed rates were true rates, generated gamma distributions for males and females separately, recategorized the gamma distributions into cells, added 0.5 to cells with 0 values, and then compared the pattern with the observed pattern), the negative binomial distribution could not be rejected (P = 0.68 for males and P = 0.30 for females). We also fitted a quasi-likelihood version of the negative binomial regression model and found little evidence of further overdispersion (0.96 instead of 1). Although it is not possible to determine whether the negative binomial model is more or less appropriate than the quasi-Poisson model (24), both of these analyses suggest no significant effect of sex, which is qualitatively different from the Poisson regression. These additional analyses clearly require significant statistical expertise and would often require collaboration with statisticians. Finally, although one should test for overdispersion to avoid inappropriate inferences when conducting Poisson regression, this cannot be done for analyses based on the standard formula, where only total numbers of events and exposures are available.

It is also possible to compare the weighted medians of the injury rates in both groups. The point estimate from the ratio of weighted medians in our data suggests a larger effect. We stress that this is not a rate ratio in the traditional epidemiologic sense and cannot be interpreted in the same way as the ratio of means. However, it does provide an assessment of the probability of injury for individual subjects. For example, our data suggest that other studies would be expected to find that 50% of the men have 0.86 times the probability of 50% of the women to have an injury rate of 8.6 injuries per 1,000 artist exposures or less. This suggests a greater decrease in the risk of injuries for a particular subject than is provided by the ratio of weighted means.

For rate differences (23), the population-based formula again gave 95% confidence intervals that were narrower than those obtained from the nonparametric bootstrap method and might lead to inappropriate inferences.

In summary, researchers should calculate overall injury rates using individual data where possible, because the underlying assumptions of data distributions for standard formulas may not be valid. The advantages and disadvantages of weighted means versus weighted medians depend on the particular research question. Bootstrap methods provide a nonparametric method for comparing rates, but Poisson or quasi-Poisson or negative binomial regression is appropriate if the underlying assumptions about the distribution are confirmed.

**CONCLUSION**

Using examples from the Cirque du Soleil injury database, we have demonstrated why some commonly used statistical techniques are not appropriate for many of the data

### Table 3. Rate Ratios and Rate Differences for Risk of Injury Among Cirque du Soleil Artists, Calculated Using Several Different Methods

<table>
<thead>
<tr>
<th></th>
<th>Rate Ratio</th>
<th>95% CI</th>
<th>Rate Difference</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard formula</td>
<td>0.93</td>
<td>0.87, 0.98</td>
<td></td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Poisson regression</td>
<td>0.93</td>
<td>0.87, 0.98</td>
<td></td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Individual weighted means</td>
<td>0.93</td>
<td>0.82, 1.05</td>
<td></td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Quasi-Poisson regression</td>
<td>0.93</td>
<td>0.83, 1.04</td>
<td></td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>Negative binomial regression</td>
<td>0.97</td>
<td>0.85, 1.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual weighted medians</td>
<td>0.84</td>
<td>0.66, 1.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population-based formula</td>
<td></td>
<td>–0.74</td>
<td>–1.33, –0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual weighted means</td>
<td>–0.74</td>
<td></td>
<td>–1.97, 0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual weighted medians</td>
<td>–1.43</td>
<td></td>
<td>–3.30, 0.24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.

a Analyses were based on work-related injury data for 2002–2006 obtained from Cirque du Soleil (Montreal, Quebec, Canada) (11).

b The wider 95% CIs obtained with bootstrapping methods or quasi-Poisson regression are more accurate because the underlying data were overdispersed (the variance was greater than the mean).

c Calculated from 1,000 bootstrap samples in R (R Foundation for Statistical Computing; http://www.R-project.org). The 95% CIs were calculated using the percentile method.
generated during injury research. If the data are not independent, the easiest, most general solution is to use bootstrapping techniques to obtain appropriate confidence intervals. Plotting injury rates against exposures provides an overview of the data that allows for easy exploration of data heterogeneity, accounting for the associated uncertainty when subjects have few exposures. Researchers need to verify underlying assumptions about the data distribution when comparing rates among different groups, or inappropriate inferences may result.

ACKNOWLEDGMENTS

Author affiliations: Centre for Clinical Epidemiology and Community Studies, Lady Davis Institute for Medical Research, SMBD-Jewish General Hospital, McGill University, Montreal, Quebec, Canada (Ian Shrier, Russell J. Steele); Department of Mathematics and Statistics, McGill University, Montreal, Quebec, Canada (Russell J. Steele); and Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec, Canada (James Hanley, Ben Rich).

Dr. Ian Shrier was supported by the Senior Clinician Scientist award from the Fonds de la Recherche en Santé du Quebec.

The authors acknowledge Cirque du Soleil for providing the data. The analyses could not have been completed without major contributions from Cirque du Soleil’s Human Resources Department (Stéphanie Tremblay), Performance Medicine Department (Janet Pundick, Eric Lamme, Jay Mellette), Coaching Staff (Marie-Hélène Martineau, Boris Verkhovsky) and Vice Presidency of Casting and Performance (Sylvie Geneau, Bernard Petiot).

Conflict of interest: none declared.

REFERENCES