STATISTICS IN MEDICINE, VOL. 16, 791–801 (1997)

CONFIDENCE INTERVALS FOR DIRECTLY STANDARDIZED RATES: A METHOD BASED ON THE GAMMA DISTRIBUTION

MICHAEL P. FAY AND ERIC J. FEUER

National Cancer Institute, Division of Cancer Prevention and Control, Executive Plaza North, Suite 344, 6130 Executive Blvd MSC 7354, Bethesda, Maryland 20892-7354, U.S.A.

SUMMARY

We offer an approximation to central confidence intervals for directly standardized rates, where we assume that the rates are distributed as a weighted sum of independent Poisson random variables. Like a recent method proposed by Dobson, Kuulasmaa, Eberle and Scherer, our method gives exact intervals whenever the standard population is proportional to the study population. In cases where the two populations differ non-proportionally, we show through simulation that our method is conservative while other methods (the Dobson *et al.* method and the approximate bootstrap confidence method) can be liberal. © 1997 by John Wiley & Sons, Ltd.

1. INTRODUCTION

In epidemiology it is common to compare incidence or mortality rates by directly standardized rates (DSRs) and to assume that one can model these DSRs as weighted sums of independent Poisson random variables where the weights are known.[1] In this paper we offer an approximation to the confidence intervals for DSRs under this assumption. We refer to these new confidence intervals as *gamma* intervals, since the approximation is based on the gamma distribution. The gamma intervals perform at least as well as existing methods in all situations studied here, but perform especially better than existing methods when the number of counts in any specific cell is small and there is large variability in the weights. Large variability in the weights occurs when comparing cancer rates across disparate populations using a single standard. For example, *Cancer Incidence in Five Continents*, *Vol IV*,[2] compares data from 163 different populations using the world population standard. For many cancers there are less than 10 cases across all age groups. Since we can write the gamma intervals as a simple function of the inverse chi-squared distribution, they are practical to use in any situation.

Dobson, Kuulasmaa, Eberle and Scherer[3] (hereafter DKES) introduced confidence limits for weighted sums of Poisson random variables that, unlike the traditional confidence limits based on the normal distribution (see Clayton and Hills[4]), do not require large cell counts. The DKES limits are *exact* for the case when all the weights are equal, a case for which the DSR reduces to a scaled Poisson random variable. The gamma intervals are also exact in this case as well as when some of the weights are equal and the rest are zero.

Recently, Swift applied the approximate bootstrap confidence (ABC) method of DiCiccio and Efron to this problem.[5,6] This method is not exact when the weights are equal but has fairly good

CCC 0277–6715/97/070791–11$17.50
© 1997 by John Wiley & Sons, Ltd.

Received December 1995
Revised May 1996

coverage properties. To test a variety of cases we perform a group of simulations on 500 different simulated data sets and one observed data set. We find that the gamma intervals remain conservative while the DKES intervals and the ABC intervals become anti-conservative as the sample variance of the weights increases.

2. BACKGROUND

2.1. Confidence Intervals

In this paper we examine only $100(1-\alpha)$ per cent *central* confidence intervals, that is, the intervals such that

$$\Pr[\mu \geqslant L(Y)|\mu;\theta] \geqslant 1-\frac{\alpha}{2} \quad \text{for all } \mu,\ \theta \tag{1}$$

and

$$\Pr[\mu \leqslant U(Y)|\mu;\theta] \geqslant 1-\frac{\alpha}{2} \quad \text{for all } \mu,\ \theta, \tag{2}$$

where $L(y)$ and $U(y)$ are the lower and upper confidence limits, respectively, for the observed y, Y is the random variable associated with y, μ is the parameter of interest, and θ is a vector of nuisance parameters. Some authors shrink the length of the intervals by using non-central confidence intervals, that is, intervals that do not meet equations (1) and (2) but still satisfy

$$\Pr[L(Y) \leqslant \mu \leqslant U(Y)] \geqslant 1-\alpha \quad \text{for all } \mu.$$

(See Crow and Gardner[7] or Casella and Robert.[8])

Exact confidence limits for discrete random variables derive from exact tests on the parameters.[9] In the usual one parameter discrete case with no nuisance parameters (for example, binomial or Poisson) the exact confidence interval is the solutions to

$$\Pr[Y \geqslant y|\mu = L(y)] = \frac{\alpha}{2} \tag{3}$$

and

$$\Pr[Y \leqslant y|\mu = U(y)] = \frac{\alpha}{2} \tag{4}$$

for $L(y)$ and $U(y)$ for each y in the sample space except $y=0$ where we define $L(0)=0$. For a single Poisson random variable we can write the solution to these equations in the form of the χ^2 distribution:[10,11]

$$L_{\mathrm{P}}(y) = \tfrac{1}{2}(\chi^2)^{-1}_{2y}(\alpha/2)$$

and (5)

$$U_{\mathrm{P}}(y) = \tfrac{1}{2}(\chi^2)^{-1}_{2(y+1)}(1-\alpha/2)$$

where $(\chi^2)^{-1}_n(p)$ is the pth quantiles of a χ^2 distribution with n degrees of freedom. Although in this paper we call confidence intervals that satisfy equations (3) and (4) exact, some authors (for

© 1997 by John Wiley & Sons, Ltd.

example, Johnson and Kotz[11]) use the term *approximate* with respect to these confidence intervals because $\Pr[\mu \geqslant L(Y)] = 1 - \alpha/2$ and $\Pr[\mu \leqslant U(Y)] = 1 - \alpha/2$ do not hold for all μ. In most cases one cannot solve the DSR in a direct manner using equations (3) and (4) because one cannot usually represent the DSR as a one parameter distribution.

2.2. Notation

Let $x_1, \ldots, x_n$ be the cell specific counts with associated random variables $X_1, \ldots, X_n$, and assume that $X_i \sim \text{Poisson}(\theta_i)$, $i = 1, \ldots, n$. Let the weight for the ith cell be

$$w_i = \frac{c_i}{n_i(\sum_{j=1}^{n} c_j)}$$

where n_i is the number of person years for the ith cell and c_i is the corresponding number of person years in the 'standard' population. Then the DSR is $y = \sum w_i x_i$, where here and throughout the paper we define all unmarked summations as going from 1 to n. An estimate of the variance of the DSR is $v = \sum w_i^2 x_i$. Let $Y = \sum w_i X_i$ and $E(Y) = \mu = \sum w_i \theta_i$. Our interest is in confidence intervals for μ, and the parameters $\theta_1, \ldots, \theta_n$ are nuisance parameters because we have no interest in them apart from their effect on μ.

2.3. ABC intervals

One can obtain a simple approximate confidence interval by assuming that Y is distributed normally with mean equal to y and variance equal to v. One can obtain another normal approximation by using the delta method on $\log(y)$.[4] A better approximation is the ABC method in which we assume that there exists some transformation on μ that gives a normal random variable, and we allow the variance to be a function of the transformed μ.[5, 6, 12] To find the ABC limits we do not need to know the transformation, instead we calculate the confidence limits from a few estimated additional parameters. For details see Efron and Tibshirani.[12] Swift[5] applied this method to DSRs to obtain

$$\begin{aligned} L(y) &= y + \frac{z_0 + \Phi^{-1}(\alpha/2)}{\{1 - a[z_0 + \Phi^{-1}(\alpha/2)]\}^2}\sqrt{v} \quad \text{for } y > 0 \\ U(y) &= y + \frac{z_0 + \Phi^{-1}(1 - \alpha/2)}{\{1 - a[z_0 + \Phi^{-1}(1 - \alpha/2)]\}^2}\sqrt{v} \quad \text{for } y > 0 \end{aligned} \tag{6}$$

where $\Phi^{-1}(p)$ is the pth percentile of the standard normal distribution, $a = z_0 = (\sum w_i^3 x_i)/(6v^{3/2})$. Following Swift, we let $L(0) = 0$ and $U(0) = (\sum w_i) U_{\mathrm{P}}(0)$, where $U_{\mathrm{P}}(x)$ is given in equations (5). The ABC intervals approach the standard normal interval as $\min(\theta_i) \to \infty$. We explore the ABC method for small θ_i through simulation in Section 4.

2.4. Dobson, Kuulasmaa, Eberle and Scherer Intervals

DKES offer a different approximation. DKES include both central and non-central intervals, but we focus on the central confidence intervals, labelled 'chi-squared' in that paper. DKES assume that Y is distributed approximately as T, where T is a scaled and shifted Poisson random variable such that $E(T) = y$ and $\text{var}(T) = v$. Specifically,

$$T = y + \frac{\sqrt{v}}{\sqrt{x}}(X - x)$$

© 1997 by John Wiley & Sons, Ltd.

where X is distributed Poisson with $E(X) = \mathrm{var}(X) = x = \sum x_i$. Thus, DKES obtain confidence intervals by scaling and shifting the exact intervals for x (see equation (5)) accordingly:

$$L(y) = y + \frac{\sqrt{v}}{\sqrt{x}}(L_{\mathrm{P}}(x) - x)$$

and (7)

$$U(y) = y + \frac{\sqrt{v}}{\sqrt{x}}(U_{\mathrm{P}}(x) - x).$$

The limits of equations (7) approach the usual normal approximation limits as x goes to infinity, since $(X - x)/\sqrt{x}$ approaches the standard normal distribution as x gets large.[11]

When all of the weights are equal, or when some of the weights are equal and the rest are zero, then the DSR becomes a scaled Poisson variate, and we can calculate exact intervals. In the former case, the DKES interval gives the exact interval, that is $(wL_{\mathrm{P}}(\sum x_i),\ wU_{\mathrm{P}}(\sum x_i))$. In the latter case, however, the exact interval is $(wL_{\mathrm{P}}(\sum_{i\in A} x_i),\ wU_{\mathrm{P}}(\sum_{i\in A} x_i))$ where $w_i = w$ for $i \in A$, while the DKES interval is

$$L(y) = w\sum_{i\in A} x_i + w\left(\frac{\sum_{i\in A} x_i}{\sum x_i}\right)^{1/2}(L_{\mathrm{P}}(\sum x_i) - \sum x_i)$$

with corresponding results for $U(y)$.

3. GAMMA CONFIDENCE INTERVALS

We motivate the gamma intervals by examining the derivation of the exact Poisson confidence limits given in equations (5). We begin with the well known relationship between the Poisson distribution and the gamma distribution, that is, if X is Poisson with mean μ then

$$\Pr[X \geqslant x|\mu] = \Pr[Z \leqslant \mu|E(Z) = x, \mathrm{var}(Z) = x] \tag{8}$$

where Z is a random variable distributed according to a gamma distribution with $E(Z) = x = \mathrm{var}(Z)$.[11] Symbolically, $Z \sim G(x, 1)$ where if $Z \sim G(a, b)$ then the density function for Z is

$$f(Z; a, b) = \frac{Z^{a-1}\exp\left(\dfrac{-Z}{b}\right)}{b^a\Gamma(a)}.$$

From equation (8) and the equations for exact confidence limits (equations (3) and (4)) we obtain $L_{\mathrm{P}}(x) = G^{-1}_{(x,1)}(\alpha/2)$ and $U_{\mathrm{P}}(x) = G^{-1}_{(x+1,1)}(1 - \alpha/2)$. These intervals are equivalent to equations (5) since the chi-squared distribution is a special case of the gamma distribution.

We form the gamma intervals by assuming that a relationship similar to equation (8) holds for the distribution of the directly standardized rate. We assume that if Y is the random variable that represents the DSR with mean μ then

$$\Pr[Y \geqslant y|\mu] \approx \Pr[Z \leqslant \mu|E(Z) = y, \mathrm{var}(Z) = v] \tag{9}$$

where $Z \sim G(\frac{y^2}{v}, \frac{v}{y})$ and y and v have been defined earlier.

© 1997 by John Wiley & Sons, Ltd.

We define the lower gamma confidence limit directly using equations (3) and (9), so that $L(y;v) = G^{-1}_{\left(\frac{y^2}{v}, \frac{v}{y}\right)}(\alpha/2)$. For the upper confidence limits we modify equation (4):

$$\begin{aligned} 1 - \frac{\alpha}{2} &= 1 - \Pr[Y \leqslant y | \mu = U(y)] \\ &= \Pr[Y > y | \mu = U(y)] \\ &\geqslant \Pr[Y \geqslant y + k | \mu = U(y)] \end{aligned} \tag{10}$$

where k is a positive constant representing a discrete increase in y. In the Poisson case, since all the non-zero weights are equal, say w, we know that $k = w$ and we reach equality in equation (10). When all the non-zero weights are not equal, it is unclear what value of k to use, and in addition, the distribution of y is not completely determined by μ. We hypothesize that we can obtain a conservative confidence limit if we let k be the maximum possible discrete increase in y, that is, $k = \max_{i \in \{1, \ldots, n\}}(w_i) \equiv w_{\mathrm{M}}$. Thus, for the upper gamma confidence limit, we let $k = w_{\mathrm{M}}$ and solve equation (9). In other words, we set $E(Z) = y + w_{\mathrm{M}}$ and $\mathrm{var}(Z) = v + w_{\mathrm{M}}^2$, so that

$$U(y;v;w_{\mathrm{M}}) = G^{-1}_{\left(\frac{(y+w_{\mathrm{M}})^2}{v+w_{\mathrm{M}}^2}, \frac{v+w_{\mathrm{M}}^2}{y+w_{\mathrm{M}}}\right)}(1 - \alpha/2).$$

We may equivalently write the gamma confidence limits in the form of the chi-squared distribution if we allow non-integer degrees of freedom:

$$L(y;v) = \frac{v}{2y}(\chi^2)^{-1}_{\frac{2y^2}{v}}(\alpha/2)$$

and (11)

$$U(y;v) = \frac{v + w_{\mathrm{M}}^2}{2(y + w_{\mathrm{M}})}(\chi^2)^{-1}_{\frac{2(y+w_{\mathrm{M}})^2}{v+w_{\mathrm{M}}^2}}(1 - \alpha/2).$$

The gamma intervals share two properties with the other intervals mentioned previously. First, if we let y go to infinity in such a way that v/y is constant, the gamma intervals approach the standard normal intervals, which asymptotically have the correct coverage. One can show this by noting that the gamma distribution approaches the normal distribution as the first parameter goes to infinity.[13] Second, the gamma intervals properly adjust for multiplying the weights by a constant, that is, if the limits for y are $[L(y), U(y)]$ then the limits for $y^* = \sum(cw_i)x_i = cy$ are $[cL(y), cU(y)]$, for any positive constant, c.

One advantage of the gamma intervals over the DKES intervals and the ABC intervals is that in both cases where the DSR becomes a scaled Poisson variate, the gamma intervals yield the exact solutions. For these examples $w_{\mathrm{M}} = w$, $y = w\sum_{i \in A} x$ and $v = wy$, so by straightforward substitution we see that equations (11) give the exact solutions. To provide further justification for the gamma intervals we perform simulations in the next section.

4. SIMULATIONS AND EXAMPLE

In this section we perform two sets of simulations and calculate the intervals for a data example. For the first set we choose small cell means for the simulation, since as $y \to \infty$ (with v/y constant) all the approximation methods mentioned in this paper are asymptotically equivalent. For each simulation in this first set we choose the cell means as well as the weights from random samples,

© 1997 by John Wiley & Sons, Ltd.

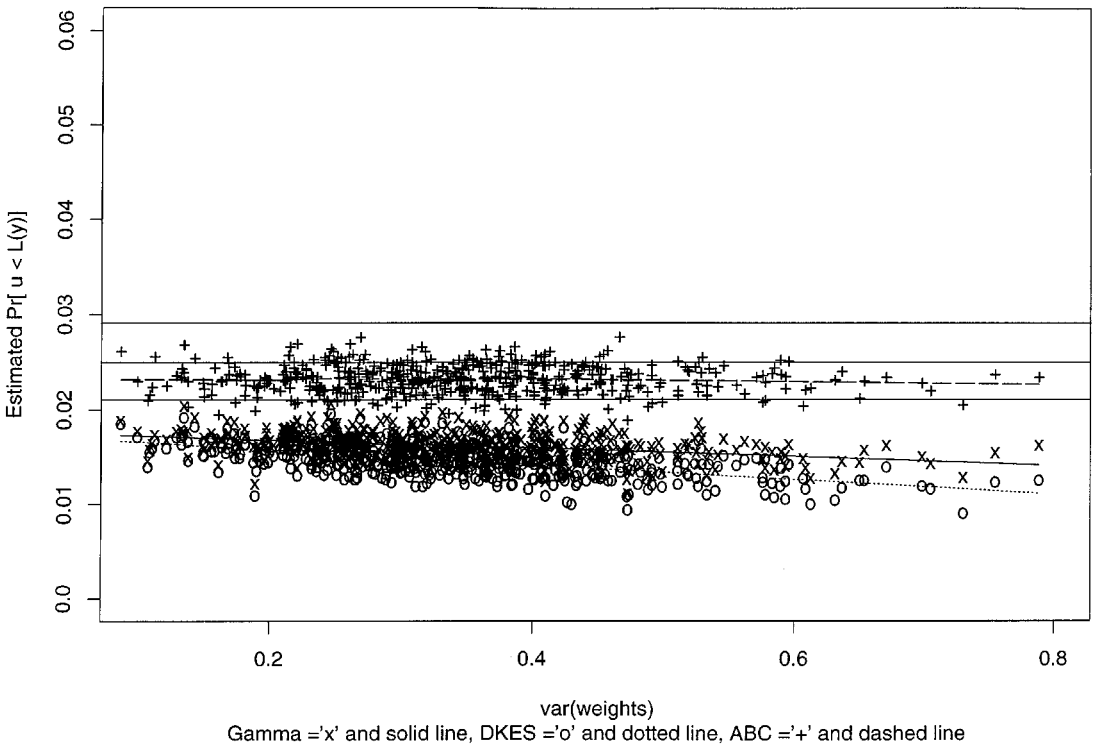

Figure 1. Simulated error on lower bounds, 500 simulations, 10,000 replications each

so that the set of simulations covers a wide variety of situations. The second set of simulations consists of one simulation where we take the weights and cell means from a real data example in which only two of the cell means have values less than 50. In all simulations we only compare the gamma method to the DKES and ABC methods since DKES and Swift[5] have previously shown the DKES and the ABC methods superior to the normal and log-normal approximations. All simulations were performed in S-plus.

The first set of simulations consists of 500 simulations of 10,000 replications each. For each simulation we let both the weights and the true means be independent random samples. For the true means, we drew a random sample of size $n = 18$ from a uniform distribution on $(0, 1)$ and standardized the true means such that $\sum \theta_i = 10$. We chose $n = 18$ to correspond to standard age groupings (0–4, 5–9, … , 85+) commonly used in demography. Similarly, for the weights we drew a separate random sample from a uniform distribution and standardized them such that $n^{-1} \sum w_i = 1$. For each replication we calculate the 95 per cent confidence limits by the DKES method (equations (7)), the ABC method (equations (6)), and by the gamma method (equations (11)). For each simulation we calculate the proportion of the lower limits greater than the true μ and the proportion of the upper limits less than the true μ which we call the simulated error.

We plot these results in Figures 1 and 2 by the sample variance of the weights. If we had an exact confidence interval for a *continuous* random variable, then for each tail 10,000 times the simulated error points would behave like binomial random variables with parameter equal to $p = 0{\cdot}025$. Using the binomial distribution, we expect 99 per cent of the simulated error points to be between 0·0211 and 0·0291. We plot these lines to aid in the interpretation of the graphs. To see

© 1997 by John Wiley & Sons, Ltd.

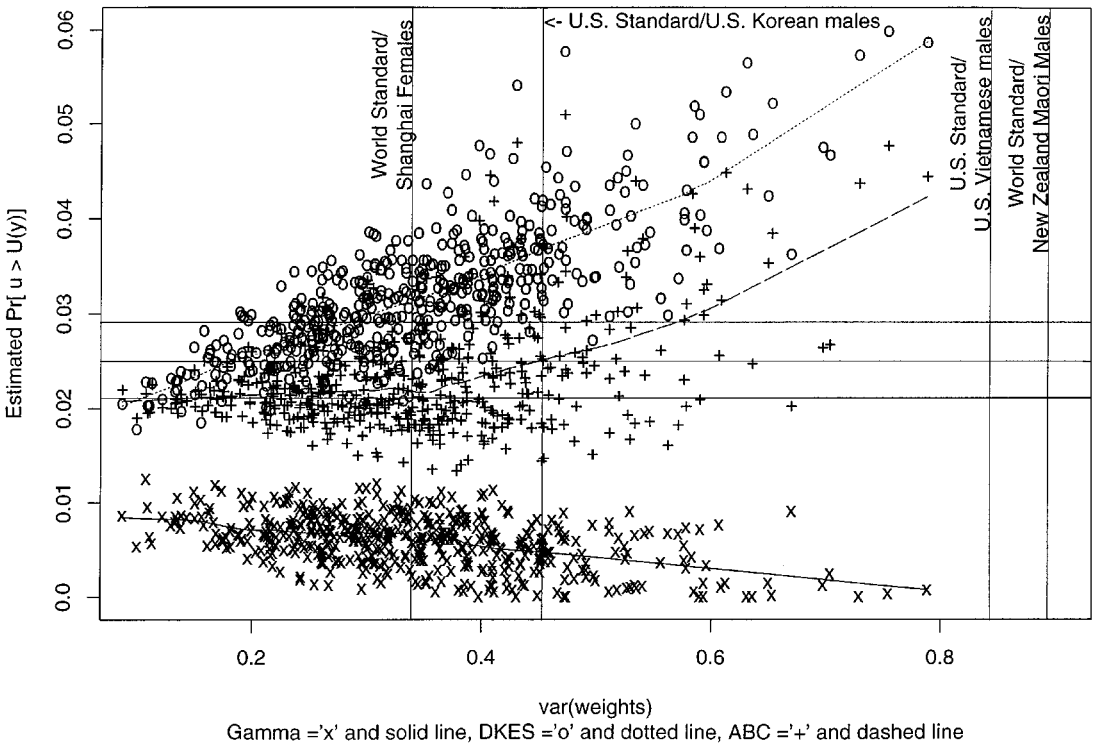

Figure 2. Simulated error on upper bounds, 500 simulations, 10,000 replications each

the relationship between the methods we plot non-parametric smoothed regression lines (using the *supsmu* function of S-plus[14]) of the simulated error regressed onto the sample variance of the weights. From Figure 1 we see that both the DKES method and the gamma method are conservative while the ABC method appears mostly within the limits of sampling error but slightly conservative. From Figure 2 we see that the DKES method and the ABC method become liberal as the sample variance of weights, $\widehat{\mathrm{var}}(w)$, becomes large, while conversely the gamma method becomes more conservative. For small $\widehat{\mathrm{var}}(w)$ the ABC limits look good in Figure 2, however, there are some cases where the ABC limits are liberal even at $\widehat{\mathrm{var}}(w) = 0$ when the problem reduces to a Poisson interval. Figure 1(b) of Swift[5] plots the ABC method for the Poisson interval. In this case, similar to the mid-P confidence intervals,[15] the ABC method may have slightly conservative or slightly liberal coverage depending on the true value of the parameter. Recall that the gamma intervals are exact in this case and exact intervals for discrete random variables may be conservative in many situations to ensure that they are not liberal in any situation.

To relate Figure 2 to practical examples, we consider weights associated with four different populations with two different standards. For each population we calculate the sample variance of the weights calculated from the ratio of the standard population to that population, where we calculated the weights from 18 standard age groups and standardized them so that $n^{-1} \sum w_i = 1$. We plot these sample variances for the four populations as vertical lines in Figure 2. Figure 3 displays graphically the four populations and two standards. From *Cancer Incidence in Five Continents, Vol IV*,[2] we use the world standard in calculating weights for the average annual

© 1997 by John Wiley & Sons, Ltd.

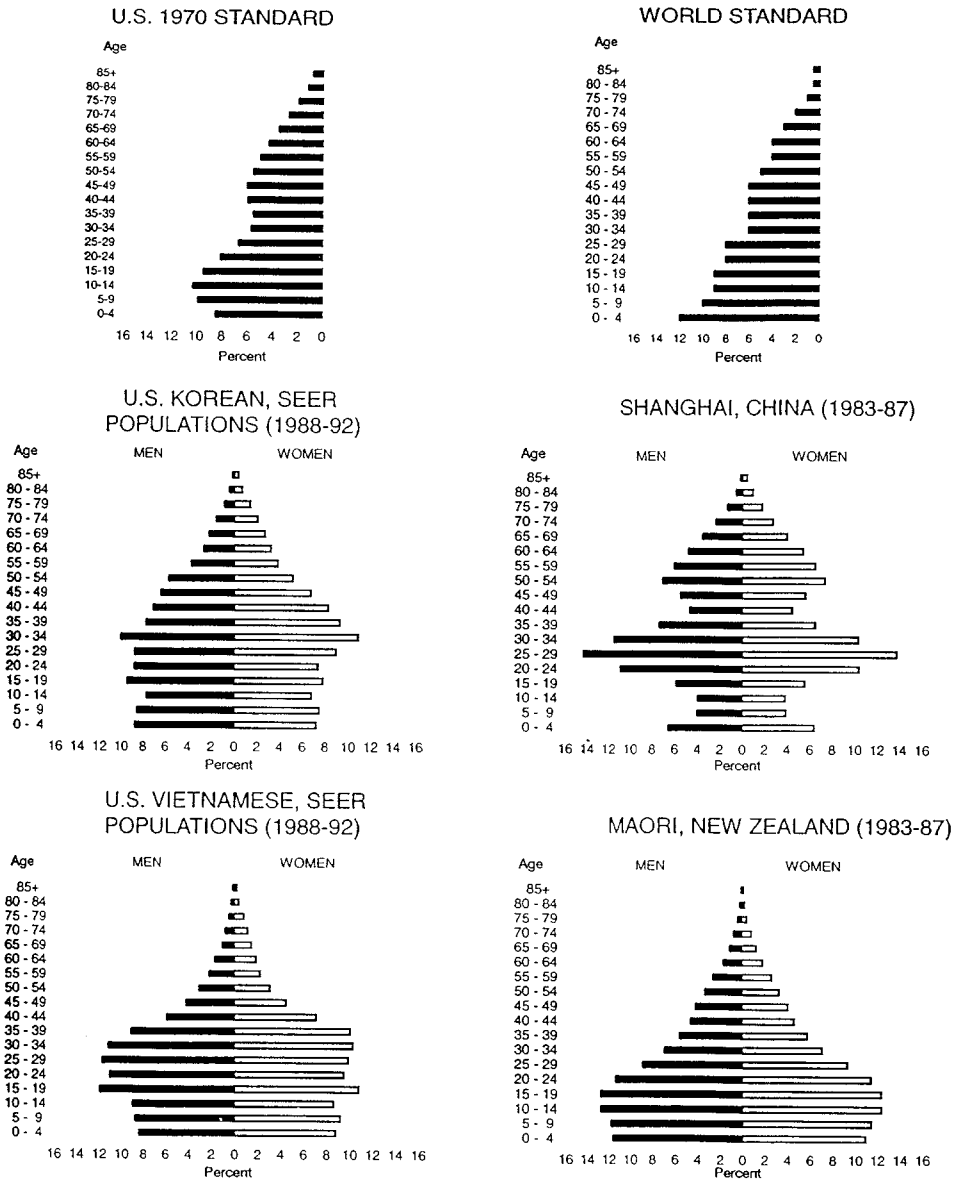

Figure 3. Comparative age distributions, United States and World. 5-year age groups by race/ethnicity and sex

populations from 1983–1987 for (i) females in Shanghai, China $\widehat{\mathrm{var}}(w) = 0{\cdot}339$, and (ii) males of Maori ethnicity in New Zealand, $\widehat{\mathrm{var}}(w) = 0{\cdot}896$. As another example, consider a monograph under preparation by the National Cancer Institute on minority populations from the Surveillance, Epidemiology, and End Results (SEER) programme. The SEER programme is a collection of nine population-based registries that constitute about 10 per cent of the U.S. population. We use the 1970 U.S. standard in calculating weights for the average annual populations from 1988–1992 for (iii) Korean males, $\widehat{\mathrm{var}}(w) = 0{\cdot}453$, and (iv) Vietnamese males, $\widehat{\mathrm{var}}(w) = 0{\cdot}844$.[16]

© 1997 by John Wiley & Sons, Ltd.

Table I. 95 per cent central confidence intervals for rate of mongoloids per 100,000 live births (data listed in Fleiss,[18] p. 249)

Birth order	Age-adjusted rate	ABC method (equations (6))	Dobson *et al.*[3] (equations (7))	Gamma method (equations (11))
1	92·3	(80·8, 105·4)	(80·3, 105·2)	(80·4, 105·8)
2	91·2	(82·5, 100·6)	(82·3, 100·6)	(82·4, 100·9)
3	85·1	(77·3, 93·5)	(77·2, 93·5)	(77·2, 94·2)
4	92·7	(82·8, 111·2)	(79·9, 106·6)	(80·0, 114·7)
5 +	75·5	(68·4, 84·6)	(67·6, 83·9)	(67·7, 188·3)

Although we chose these four populations because of their unusually large values for $\widehat{\mathrm{var}}(w)$, there are larger values; for non-Kuwaiti males living in Kuwait, using the world standard,[2] $\widehat{\mathrm{var}}(w) = 1{\cdot}812$. From Figure 2 we expect that the gamma method will give a noticeable improvement over the DKES method and the ABC method when comparing rare cancers in populations that differ considerably from the standard population.

Consider a more striking example of extreme variability of the weights from Stark and Mantel[17] presented in Fleiss.[18] In this case the variability of the weights is so great that coverage of the standard methods appears to be liberal even when there is a large total number of counts (although each cell does not have a large number of counts). The age-adjusted rates for mongolism per 100,000 live births in Michigan (1950–1964) are calculated according to the birth order of the child (that is, oldest, second oldest, etc.) using the distribution of the mother's age for all births as a standard. For the largest birth order group (5+) the weights for the first two maternal age groups (34·63 and 1·07) are very large compared to those of the older age groups (0·23, 0·12, 0·08, 0·06, respectively). Note also that the number of counts is large; there are over 400 mongoloid births in each of the birth order groups. We list the confidence intervals calculated by the ABC method, the DKES method and the gamma method in Table I.

In the first three birth orders all the methods are similar, but in the last two birth orders there are differences. In 4 and 5+ birth order groups the cells that represent younger mothers have very large weights since there are very few mothers who have their fourth or fifth child before they are 20 years old. Thus, if there were one more mongoloid case in the younger of the last two groups, then the age-adjusted rates would be considerably larger, 97·7 (for birth order =4) and 110·2 (for birth order = 5+). Notice that, particularly for birth order 5+ , the gamma confidence intervals reflect this instability while the DKES intervals and the ABC intervals do not.

We perform a simulation to test the coverage for the group with birth order =5+ . The data appear in Table II. For the simulation, we define the true rates as the observed rates except for mothers under 20 years old, where we set the true rate equal to the rate from all mothers under 20 years old (that is, 42·5 per 100,000). We make this adjustment since there are no observed mongoloid births of the 327 live births that are the fifth oldest or greater for mothers under 20 years old. A rate of 42·5 per 100,000 is not inconsistent with the observed data since the probability of observing 0 events out of 327 is 87 per cent with the rate of 42·5 per 100,000. We calculate 95 per cent confidence limits for each of 10,000 replications using the observed weights. The lower confidence limits are conservative for all three methods; the upper confidence limits, however, are conservative for only the gamma method. For the gamma method the upper confidence limit was always greater than the true DSR, while the ABC method (equations (6)) and the DKES method had, respectively, 16·15 per cent and 19·80 per cent upper confidence limits less

© 1997 by John Wiley & Sons, Ltd.

Table II. Birth data from Michigan, 1950–1964 (see Fleiss,[18] p. 240)

Maternal age	Birth order = 5 +			All birth orders		
	Number of mongoloids	Number of live births	Rate per 100,000	Number of mongoloids	Number of live births	Rate per 100,000
Under 20	0	327	0	136	319,933	42·5
20–24	8	30,666	26·1	396	931,318	42·5
25–29	63	123,419	51·0	411	786,511	52·3
30–34	112	149,919	74·7	428	488,235	87·7
35–39	262	104,088	251·7	628	237,863	264·0
40 and over	295	34,392	857·8	530	61,313	864·4

than the true DSR (the true value should be less than 2·5 per cent). Thus in this extreme situation the gamma method is very conservative, but the other methods are very liberal.

5. DISCUSSION

We have offered a method for calculating confidence intervals for weighted sums of Poisson variables. This new confidence interval (gamma interval) gives improved coverage compared to existing methods and is easy to calculate. Many statistical packages (for example, SAS, S-plus) have a function to calculate the inverse chi-square distribution. Thus, we recommend it for standard use over existing methods.

The gamma interval is given by the simple form of equations (11) where we have allowed for continuous degrees of freedom for the chi-square distribution. In terms of an age standardized rate, $y = \sum w_i x_i$ is the age standardized rate, w_i is the ratio of the proportion of the standard population in the ith age group to the observed population at that age group, w_{M} is the largest of these ratios, x_i is the count at the ith age group, and $v = \sum w_i^2 x_i$ is the estimated variance of the age standardized rate. If continuous degrees of freedom are unavailable, one may calculate the confidence intervals from equations (11) with the degrees of freedom rounded to the nearest integer. Simulations on these 'rounded' confidence intervals give similar results as the non-rounded ones.

The gamma confidence intervals appear overly conservative in both sets of simulations, so one may be tempted to try smaller values for k. For example, one could use the minimum or the weighted average of the weights rather than the maximum weight for k. We simulated the latter case (that is, let $w_{\mathrm{M}} = \sum x_i w_i / \sum x_i$ in equations (11)) and found that the intervals became anti-conservative as the sample variance of the weights increased. We note that from the 5,000,000 replications from the first set of simulations the average length of the gamma intervals was only 9.6 per cent larger than the average length of the DKES intervals and 11.6 per cent larger than the average length of the ABC intervals. So the added cost of conservatism does not appear too great.

Note that DiCiccio and Efron[6] in presenting the ABC method warned against possible boundary problems when any observed cell count is zero (that is, $x_i = 0$ for any i). They recommend adding a small value, like 0·5, to each cell in which the observed count is zero. This is not the method that Swift[5] uses and appears to give more conservative estimates. We ran the first set of simulations 100 times adding 0·5 to each cell with a zero count in calculating the ABC upper confidence interval, and ran the second simulation on the data in Table II using this ABC upper confidence interval. In both simulations this ABC method had larger intervals and was much

© 1997 by John Wiley & Sons, Ltd.

more conservative than the gamma method. So the gamma method appears better than *this* ABC method as well.

In this paper we have assumed that the Poisson assumption was valid, however, sometimes there exists extra-Poisson variation in the counts. One can modify the gamma method to account for this extra-Poisson variation. Let $\widehat{\mathrm{var}}(x_i)$, $i = 1, \ldots, n$, be estimates for the variance of each X_i (see for example McCullagh and Nelder,[19] p. 198–200). Replace the v of equations (11) with $v_{\mathrm{E}} = \widehat{\mathrm{var}}(Y) = \sum w_i^2 \widehat{\mathrm{var}}(x_i)$ to obtain confidence intervals that accommodate the extra-Poisson variation. We are currently studying methods of estimating overdispersion for DSRs.

ACKNOWLEDGEMENTS

The authors thank Jian-Lun Xu and Charles Brown for discussions on this topic.

REFERENCES

1. Breslow, N. and Day, N. E. *Statistical Methods in Cancer Research: Volume 2 – The Design and Analysis of Cohort Studies*, Oxford University Press, London, 1987.
2. Parkin, D. M., Muir, C. S., Whelan, S. L., Gao, Y. T., Ferlay, J. and Powell, J. (eds), *Cancer Incidence in Five Continents*, International Agency for Research on Cancer, Lyon, 1992.
3. Dobson, A. J., Kuulasmaa, K., Eberle, E. and Scherer, J. 'Confidence intervals for weighted sums of Poisson parameters', *Statistics in Medicine*, **10**, 457–462 (1991).
4. Clayton, D. and Hills, M. *Statistical Models in Epidemiology*, Oxford University Press, New York, 1993.
5. Swift, M. B. 'Simple confidence intervals for standardized rates based on the approximate bootstrap method', *Statistics in Medicine*, **14**, 1875–1888 (1995).
6. DiCiccio, T. and Efron, B. 'More accurate confidence intervals in exponential families', *Biometrika*, **79**, 231–245 (1992).
7. Crow, E. L. and Gardner, R. S. 'Confidence intervals for the expectation of a Poisson variable', *Biometrika*, **46**, 441–453 (1959).
8. Casella, G. and Robert, C. 'Refining Poisson confidence intervals', *Canadian Journal of Statistics*, **17**, 45–57 (1989).
9. Breslow, N. and Day, N. E. *Statistical Methods in Cancer Research: Volume 1 – The Analysis of Case-Control Studies*, Oxford University Press, London, 1980.
10. Garwood, F. 'Fiducial limits for the Poisson distribution', *Biometrika*, **28**, 437–442 (1936).
11. Johnson, N. L. and Kotz, S. *Discrete Distributions*, Wiley, New York, 1969.
12. Efron, B. and Tibshirani, R. J. *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.
13. Johnson, N. L. and Kotz, S. *Continuous Univariate Distributions-1*, Wiley, New York, 1970.
14. Statistical Sciences, *S-PLUS Guide to Statistical and Mathematical Analysis, Version* 3.2, StatSci, a division of MathSoft Inc., Seattle, 1993.
15. Cohen, G. R. and Yang, S-Y. 'Mid-P confidence intervals for the Poisson expectation', *Statistics in Medicine*, **13**, 2189–2203 (1994).
16. Miller, B. A., Kolonel, L. N., Bernstein, L., Young, Jr. J. L., Swanson, G. M., West, D., Key, C. R., Liff, J. M., Glover, C. S., Alexander, G. A. *et al.* (eds), *Racial/Ethnic Patterns of Cancer in the United States 1988–1992*, National Cancer Institute, NIH Pub., Bethesda, MD, 1996.
17. Stark, C.R. and Mantel, N. 'Effects of maternal age and birth order on the risk of mongolisma and leukemia', *Journal of the National Cancer Institute*, **37**, 687–698 (1966).
18. Fleiss, J. L. *Statistical Methods for Rates and Proportions*, Wiley, New York, 1981.
19. McCullagh, P. and Nelder, J. A. *Generalized Linear Models*, 2nd edn, Chapman and Hall, New York, 1989.

© 1997 by John Wiley & Sons, Ltd.