Statistical Methods in Medical Research 2006; **15**: 547–569

Efficient interval estimation for age-adjusted cancer rates

Ram C Tiwari National Cancer Institute, NIH, Bethesda, MD, USA, **Limin X Clegg** Office of Healthcare Inspections, OIG, Department of Veterans Affairs, Washington, DC, USA and **Zhaohui Zou** Information Management Services, Inc., Silver Spring, MD, USA

The age-adjusted cancer rates are defined as the weighted average of the age-specific cancer rates, where the weights are positive, known, and normalized so that their sum is 1. Fay and Feuer developed a confidence interval for a single age-adjusted rate based on the gamma approximation. Fay used the gamma approximations to construct an F interval for the ratio of two age-adjusted rates. Modifications of the gamma and F intervals are proposed and a simulation study is carried out to show that these modified gamma and modified F intervals are more efficient than the gamma and F intervals, respectively, in the sense that the proposed intervals have empirical coverage probabilities less than or equal to their counterparts, and that they also retain the nominal level. The normal and beta confidence intervals for a single age-adjusted rate are also provided, but they are shown to be slightly liberal. Finally, for comparing two correlated age-adjusted rates, the confidence intervals for the difference and for the ratio of the two age-adjusted rates are derived incorporating the correlation between the two rates. The proposed gamma and F intervals and the normal intervals for the correlated age-adjusted rates are recommended to be implemented in the Surveillance, Epidemiology and End Results Program of the National Cancer Institute.

1 Introduction

Despite rapid advances in medicine, cancer continues to be a major public health concern in the US and around the world. The total number of deaths due to cancer continues to rise, even though the age-adjusted mortality rates for many common cancer sites continue to decline.[1] Many public and private agencies dealing with cancer and related issues depend on these statistics for planning and resource allocation. Such figures have important social and economic ramifications, from deciding which programs get funded, to deciding how funds are allocated among various programs. Having reliable and accurate confidence intervals (CIs) for the means of the age-adjusted cancer mortality and incidence rates for recent years is very important for everyone concerned. The higher the coverage probabilities of the CIs, the more conservative the CIs are. Therefore, a desirable property of these CIs is that while retaining the nominal level, they have coverage probabilities as close to the nominal level as possible.

In the US, the data on cancer mortality are obtained from death certificates. Due to administrative and procedural delays, these data become fully available to the public

Address for correspondence: Limin Clegg, Office of Healthcare Inspections, VA OIG 801 1 Street, NW, Room 1013, Washington, DC 20001, USA. E-mail: lin_clegg@va.gov

 10.1177/0962280206070621

from the National Center for Health Statistics (NCHS) after approximately three years. The cancer incidence and mortality data are also available from the Surveillance, Epidemiology and End Results (SEER) Program of the National Cancer Institute (NCI). The SEER Program is an authoritative source for the cancer incidence and survival data in the US Population data are available from the US Census Bureau. The American Cancer Society (ACS) publishes reports on cancer trends in their widely circulated annual publication,[2] *Cancer Facts & Figures*, which is also available online: http://www.cancer.org/.

The state-level age-adjusted cancer (incidence or mortality) rates are given by

$$r_i = \sum_{j=1}^{J} w_j \frac{d_{ij}}{n_{ij}}, \quad i = 1, \ldots, I$$

where d_{ij} and n_{ij} are the number of cancer (incidence or mortality) counts and the count of mid-year population for the age-group j and the state i, respectively, and the w_j are the normalized proportion of mid-year population for the age-group j in the standard population, so that $\sum_{j=1}^{J} w_j = 1$. In the SEER Program, for each of the 51 regions (50 states and Washington D.C.) in the US, there are 19 standard age-groups consisting of $0-<1, 1-4, 5-9, \ldots, 85+$. The US-level age-adjusted cancer (incidence or mortality) rates are given by

$$r = \sum_{j=1}^{J} w_j \frac{d_j}{n_j}$$

with $d_j = \sum_{i=1}^{I} d_{ij}$ and $n_j = \sum_{i=1}^{I} n_{ij}$. The SEER Program contains age-adjusted mortality rates, based on the 2000 US standard population, for the US and for each of the 51 regions by cancer sites. The age-adjusted mortality rates for a selected number of cancer sites and a number of countries in the world are also reported in the *Cancer Facts & Figures* publication.[2] These age-adjusted rates are based on the World Health Organization's world standard population. Thus, the results of this paper, even though discussed in the context of the age-adjusted mortality rates for the US, apply to similar data sets for other countries.

For each i ($i = 1, \ldots, I$), let $d_{(-i)j} = d_j - d_{ij}$ and $n_{(-i)j} = n_j - n_{ij}$ and define

$$r_{(-i)} = \sum_{j=1}^{J} w_j \frac{d_{(-i)j}}{n_{(-i)j}}$$

to be the age-adjusted rate for the rest of the US after deleting the region i.

Let D_{ij}, D_j, $D_{(-i)j}$, R_i, $R_{(-i)}$ and R denote the random variables whose realizations are d_{ij}, d_j, $d_{(-i)j}$, r_i, $r_{(-i)}$ and r, respectively. We assume that D_{ij} are independent Poisson random variables[3] with parameters λ_{ij}, that is, $D_{ij} \stackrel{\text{ind}}{\sim} \text{Po}(n_{ij}\lambda_{ij})$. Note that by the moment generating function, $D_{(-i)j} \sim \text{Po}(\sum_{i' \neq i}^{I} n_{i'j}\lambda_{i'j})$ and $D_j \sim \text{Po}(\sum_{i=1}^{I} n_{ij}\lambda_{ij})$. Let $\xi_{ij} = n_{ij}/n_j$, $\xi_{(-i)j} = \sum_{i' \neq i}^{I} \xi_{i'j}$ and $\xi_j = n_j/n$, where $n = \sum_{j=1}^{J} n_j = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij}$. Let μ_i, $\mu_{(-i)}$, μ, $v_i = \sigma_i^2/n$, $v_{(-i)} = \sigma_{(-i)}^2/n$ and $v \equiv \sigma^2/n$ be the means and variances of R_i, $R_{(-i)}$ and R, respectively, and let and ρ_i/n be the $\text{Cov}(R_i, R)$, where their explicit expressions are derived in Appendix A. Let $w_{ij} = w_j/n_{ij}$ and define the estimates of μ_i, $\mu_{(-i)}$, μ, σ_i^2, $\sigma_{(-i)}^2$, σ^2 and ρ_i as

$$\hat{\mu}_i = r_i; \quad \hat{\mu}_{(-i)} = r_{(-i)}; \quad \hat{\mu} = r$$

$$\hat{\sigma}_i^2 = n \sum_{j=1}^{J} w_{ij}^2 d_{ij}; \quad \hat{\sigma}_{(-i)}^2 = n \sum_{j=1}^{J} w_j^2 \frac{d_{(-j)}}{n_{(-j)}^2}$$

$$\hat{\sigma}^2 = n \sum_{j=1}^{J} w_j^2 \frac{d_j}{n_j^2}; \quad \hat{\rho}_i = n \sum_{j=1}^{J} w_{ij} \frac{w_j}{n_j} d_{ij}$$

For a rare cancer site, as the observed total counts d_i are very small with $d_{ij} = 0$ plausibly for several j, the value of r_i is either close to 0 or equal to 0. As we will see subsequently, when $r_i = 0$, the gamma intervals of Fay and Feuer[4] is not defined. To avoid such situations, we introduce a correction factor, which amounts to distributing a count of 1 uniformly to all J categories, and hence adding $1/J$, the expected value under multinomial distribution with parameters 1 and cell probabilities $1/J$, to d_{ij}, $j = 1, \ldots, J$, in calculation of the estimates of μ_i, σ_i^2 and ρ_i. We redefine r_i as

$$\tilde{r}_i = \sum_{j=1}^{J} w_{ij} \left(d_{ij} + \frac{1}{J} \right) = r_i + \bar{w}_i$$

where $\bar{w}_i = 1/J \sum_{j=1}^{J} w_{ij}$ and modify the estimates of μ_i, σ_i^2, $\sigma_{(-i)}^2$ and ρ_i accordingly by replacing d_{ij} by $(d_{ij} + 1/J)$. Thus,

$$\tilde{\mu}_i = \tilde{r}_i; \quad \tilde{\sigma}_i^2 = n \sum_{j=1}^{J} w_{ij}^2 \left(d_{ij} + \frac{1}{J} \right); \quad \tilde{\rho}_i = n \sum_{j=1}^{J} w_{ij} \frac{w_j}{n_j} \left(d_{ij} + \frac{1}{J} \right)$$

Note that $\tilde{r}_i \approx r_i$ for common cancer sites as $\bar{w}_i \approx 0$. Let

$$\hat{v}_i = \frac{\hat{\sigma}_i^2}{n}; \quad \tilde{v}_i = \frac{\tilde{\sigma}_i^2}{n}; \quad \hat{v} = \frac{\hat{\sigma}^2}{n}$$

The objectives of this paper include the construction of CIs for parameters such as i) the mean μ_i of the age-adjusted rate for the region i; ii) the mean μ of the age-adjusted rate for the US; iii) the ratio of the mean age-adjusted rates $\mu_i/\mu_{i'}$ for region i to region i'; iv) the ratio of the mean age-adjusted rates $\mu_i/\mu_{(-i)}$ for region i to the rest of the US; v) the ratio of the mean age-adjusted rates μ_i/μ for region i to the US; and vi) the difference of the mean age-adjusted rates $\mu_i - \mu$, between region i and the US. Fay and Feuer[4] derived a CI for μ_i (or μ) assuming that a mixture of Poisson distributions can be approximated by a gamma distribution and compared the performance of the gamma intervals with the approximate bootstrap confidence (ABC) intervals[5–7] and the 'chi-squared' intervals of Dobson *et al*.[8] through simulations. They observed that the gamma intervals retained at least the nominal coverage and were more conservative than the ABC intervals and chi-squared intervals. We propose a modification of the gamma interval for μ_i (or μ) developed by Fay and Feuer[4] and derive new CIs for μ_i (and μ) based on the beta and normal approximations of R_i (and R).

Fay[9] used the gamma approximation of Fay and Feuer[4] and developed a CI, based on an approximate F distribution, for the ratio of two age-adjusted rates that can be applied to $\mu_i/\mu_{i'}$ and $\mu_i/\mu_{(-i)}$, but not to μ_i/μ as the age-adjusted rate for the US involves the counts from the region i. We also propose a modification of the F interval of Fay[9]. We use the normal approximations of $R_i/R_{i'}$, $R_i/R_{(-i)}$, R_i/R and $R_i - R$, taking into account the correlation between R_i and R, and construct the CIs for $\mu_i/\mu_{i'}$, $\mu_i/\mu_{(-i)}$, μ_i/μ and $\mu_i - \mu$. It is important to mention that for comparing the state and US level age-adjusted rates, the current procedure[10] is to use the normal CI for $\mu_i - \mu$ based on $\rho_i = 0$. For simulations, we use the observed age-adjusted mortality rates for the 51 regions and the US for year 2002 from the SEER Program for a rare cancer site, the tongue cancer.

The rest of the paper is organized as follows. In Section 3, we briefly review the works of Fay and Feuer[4] and Fay[9] and present the modified gamma and F intervals. We also derive the CIs for the ratio of the means of the two age-adjusted rates namely the age-adjusted rates of any two regions, any region to the rest of the country, any region to the entire country and for the difference of the means of the age-adjusted rates between a region and the country. Simulations are carried out in Section 4, and we discuss our findings in Section 5. The conclusions are presented in Section 6.

2 Confidence intervals for age-adjusted rates

2.1 Gamma and *F* approximations

Note that if $X \sim \text{Po}(\theta)$, then for an integer $x \geq 0$,

$$P(X \geq x|\theta) = \int_0^{\theta} f_Z(z|x, 1)\,\mathrm{d}z$$

where $Z \sim G(x, 1) =^d 1/2\chi^2_{2x}$ and, in general, $G(\alpha, \beta) =^d \beta/2\chi^2_{2\alpha}$ (allowing non-integer degrees of freedom) has density

$$f_Z(x|\alpha, \beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \exp\left(-\frac{x}{\beta}\right) x^{\alpha-1}, \quad x > 0$$

with mean $E(Z) = \alpha\beta$ and variance $\text{Var}(Z) = \alpha\beta^2$. Let x be the observed value of X and let $(L(x;\alpha), U(x,\alpha))$ denote the $100(1-\alpha)\%$ CI for θ, where $L(x;\alpha)$ is obtained by solving the equation

$$P(X \geq x|\theta = L(x;\alpha)) = \frac{\alpha}{2}$$

and $U(x,\alpha)$ is obtained by solving

$$P(X \leq x|\theta = U(x;\alpha)) = \frac{\alpha}{2}$$

or equivalently by solving

$$P(X > x|\theta = U(x;\alpha)) = P(X \geq x+1|\theta = U(x;\alpha)) = 1 - \frac{\alpha}{2}$$

Thus, $L(x;\alpha) = 1/2(\chi^2_{2x})^{-1}(\alpha/2)$ and $U(x;\alpha) = 1/2(\chi^2_{2(x+1)})^{-1}(1-\alpha/2)$. Fay and Feuer[4] called the interval $(L(x;\alpha), U(x;\alpha))$ 'exact' while others, for example, Johnson and Kotz,[11] use the term 'approximate' interval.

Let $w_{i(1)} \leq \cdots \leq w_{i(J)}$ be the ordered values of w_{ij}, $j = 1,\ldots,J$. Fay and Feuer[4] assumed that a mixture of Poisson distributions is approximately distributed as a gamma distribution; that is,

$$P\left(\sum_{j=1}^{J} w_{ij}D_{ij} \geq y|\mu_i,\nu_i\right) \approx \int_0^{\mu_i} f_{Z_i}\left(z\left|\frac{y^2}{\nu_i},\frac{\nu_i}{y}\right.\right)\,\mathrm{d}z$$

where $Z_i \sim G(y^2/\nu_i, \nu_i/y)$. This assumption essentially means that the distribution of a linear combination of independent Poisson random variables is approximately distributed as a gamma random variable with the mean and variance of the gamma distribution equal to the mean and variance of the linear combination, respectively. Fay and Feuer[4] used this approximation to construct approximate $100(1-\alpha)\%$ CIs for the true age-adjusted rates μ_i.

The lower confidence limit $L(r_i;\alpha)$ was obtained by solving the equation

$$\frac{\alpha}{2} = P\left(\sum_{j=1}^{J} w_{ij}D_{ij} \geq r_i|\mu_i,\nu_i\right) \approx \int_0^{L(\mu_i;\nu_i)} f_{Z_i}\left(z\left|\frac{r_i^2}{\nu_i},\frac{\nu_i}{r_i}\right.\right)\,\mathrm{d}z$$

This yields

$$L(r_i;\hat{\nu}_i;\alpha) = G^{-1}\left(\frac{\alpha}{2};\frac{r_i^2}{\hat{\nu}_i},\frac{\hat{\nu}_i}{r_i}\right) = \frac{\hat{\nu}_i}{2r_i}\left(\chi^2_{2r_i^2/\hat{\nu}_i}\right)^{-1}\left(\frac{\alpha}{2}\right)$$

where G^{-1} is the inverse function of the gamma distribution function and $(\chi^2_l)^{-1}(\alpha)$ denotes the 100αth percentile of the chi-squared distribution with l degrees of freedom.

Note that when $r_i = 0$, $L(r_i; \hat{\nu}_i; \alpha)$ is not defined. For the upper confidence limit $U(r_i; \alpha)$, Fay and Feuer[4] solved the equation

$$1 - \frac{\alpha}{2} = P\left(\sum_{j=1}^{J} w_{ij} D_{ij} > r_i | \mu_i, \nu_i\right) \geq P\left(\sum_{j=1}^{J} w_{ij} D_{ij} \geq r_i + w_{i(J)} | \mu_i, \nu_i\right)$$

$$\approx \int_0^{U(\mu_i;\nu_i)} f_{Z_i}\left(z \left| \frac{(r_i + w_{i(J)})^2}{\nu_i + w_{i(J)}^2}, \frac{\nu_i + w_{i(J)}^2}{r_i + w_{i(J)}}\right.\right) \mathrm{d}z$$

resulting in

$$U(r_i; \hat{\nu}_i, w_{i(J)}; \alpha) = G^{-1}\left(1 - \frac{\alpha}{2}; \frac{r_i^2}{\hat{\nu}_i}, \frac{\hat{\nu}_i}{r_i}, w_{i(J)}\right)$$

$$= \frac{\hat{\nu}_i + w_{i(J)}^2}{2(r_i + w_{i(J)})} \left(\chi^2_{(2(r_i + w_{i(J)})^2/\hat{\nu}_i + w_{i(J)}^2)}\right)^{-1} \left(1 - \frac{\alpha}{2}\right)$$

Fay and Feuer[4] performed simulations to study the performance of their gamma CIs ($L(r_i; \hat{\nu}_i; \alpha)$, $U(r_i; \hat{\nu}_i; w_{i(J)}; \alpha)$) and found that the upper confidence limits were more conservative than those based on the ABC intervals[5–7] and the chi-squared intervals of Dobson *et al.*,[8] henceforth referred to as DKES intervals. For completeness the ABC and DKES intervals are given in Appendix B.

Fay and Feuer[4] have mentioned that when the weights w_{ij} for all j are equal to a constant, $c_i > 0$, say, the CI for $\mu_i = E(\sum_{j=1}^{J} w_{ij} D_{ij}) = c_i E(D_i)$ is $(c_i L(d_i; \alpha), c_i U(d_i : \alpha))$ exact with $D_i \sim \mathrm{Po}(\sum_{j=1}^{J} n_{ij} \lambda_{ij})$. However, note that since $w_{ij} = w_j / n_{ij}$ depend on both the standards w_j and on the mid-year populations n_{ij}, the condition that w_{ij} are equal to a constant for all j is not easily satisfied. For example, a sufficient condition for this condition to hold is that w_j are all equal and n_{ij} are all equal. Another sufficient condition for w_{ij} to be equal to a constant for all j is to assume n_{ij} is proportional to w_j, independent of i, for all j. If a populous state like California or New York has the age-group distribution of its population similar to that of the entire US, then for that state, one may expect n_{ij} to be proportional to w_j and hence the CI for μ_i to be exact.

Since $w_{i(l)} \leq w_{i(l+1)}$, we have

$$P\left(\sum_{j=1}^{J} w_{ij} D_{ij} > r_i | \mu_i, \nu_i\right) \geq P\left(\sum_{j=1}^{J} w_{ij} D_{ij} \geq r_i + w_{i(l)} | \mu_i, \nu_i\right)$$

$$\geq P\left(\sum_{j=1}^{J} w_{ij} D_{ij} \geq r_i + w_{i(l+1)} | \mu_i, \nu_i\right), \quad l = 1, \ldots, J-1$$

Thus proceeding as above, one can construct the upper confidence limits $U(r_i; \hat{\nu}_i; w_{i(1)}; \alpha)$, $U(r_i; \hat{\nu}_i; w_{i(2)}; \alpha), \ldots, U(r_i; \hat{\nu}_i, w_{i(J)}; \alpha)$ varying from being the most liberal upper limit to the most conservative upper limit. In fact, there are an infinite number of choices for such an upper confidence limit.

As a compromise, we propose an upper limit that is based on the mean $\bar{w}_i = 1/J \sum_{j=1}^{J} w_{ij}$ and that depends on all $w_{i(l)}$, $l = 1, \ldots, J$. As mentioned earlier, this assumes distributing a count of 1 uniformly to all J age-groups. Thus,

$$1 - \frac{\alpha}{2} = P\left(\sum_{j=1}^{J} w_{ij} D_{ij} > r_i | \mu_i, \nu_i\right) \geq P\left(\sum_{j=1}^{J} w_{ij} D_{ij} \geq r_i + \bar{w}_i | \mu_i, \nu_i\right)$$
$$= P\left(\sum_{j=1}^{J} w_{ij} D_{ij} \geq \tilde{r}_i | \mu_i, \nu_i\right)$$

Now, assuming that $(d_{ij} + 1/J)$ have means equal to their variances, similar to the Poisson distribution, so that

$$\text{Var}\left(\sum_{j=1}^{J} w_{ij}\left(d_{ij} + \frac{1}{J}\right)\right) = \sum_{j=1}^{J} w_{ij}^2 \left(d_{ij} + \frac{1}{J}\right)$$

using the gamma approximation, the upper confidence limit for μ_i is given by

$$U(\tilde{r}_i; \tilde{\nu}_i; \bar{w}_i; \alpha) = \frac{\tilde{\nu}_i}{2\tilde{r}_i} (\chi^2_{(2\tilde{r}_i^2/\tilde{\nu}_i)})^{-1} \left(1 - \frac{\alpha}{2}\right)$$

Therefore, the proposed gamma CI for μ_i is $(L(r_i; \hat{\nu}_i; \alpha),\ U(\tilde{r}_i; \tilde{\nu}_i; \bar{w}_i; \alpha))$. Another approximation of the upper confidence limit based on the mean $\bar{w}_i$ can be obtained by using $(\hat{\nu}_i + \bar{w}_i^2)$ instead of $\tilde{\nu}_i$. This results in the following CI: $(L(r_i; \hat{\nu}_i; \alpha), U(\tilde{r}_i; \hat{\nu}_i + \bar{w}_i^2; \bar{w}_i; \alpha))$. Through simulations (not shown here), we found that these two intervals performed very similarly. Therefore, we will focus only on $(L(r_i; \hat{\nu}_i; \alpha), U(\tilde{r}_i; \tilde{\nu}_i; \bar{w}_i; \alpha))$. Note that the lower limits of the gamma interval of Fay and Feuer[4] and the modified gamma intervals are the same. We shall define the CI for μ_i when $r_i = 0$ as $(0, U(\tilde{r}_i; \tilde{\nu}_i; \bar{w}_i; \alpha))$, thus ensuring a coverage probability of at least $(1 - \alpha)$.

Fay[9] developed a confidence interval for the ratio of two age-adjusted rates, $\mu_i/\mu_{i'}$, for $\mu_i, \mu_{i'} > 0$, based on $R_i = \sum_{j=1}^{J} w_{ij} D_{ij}$ and $R_{i'} = \sum_{j=1}^{J} w_{i'j} D_{i'j}$, where D_{ij} and $D_{i'j}$ are independent. Assuming the gamma approximations for R_i and $R_{i'}$, Fay[9] used the result that, conditional on $D_{ij} + D_{i'j} = t_j$, the distribution of D_{ij} is a binomial distribution with parameters t_j and $n_{ij}(\lambda_{ij}/\lambda_{i'j})/n_{ij}(\lambda_{ij}/\lambda_{i'j}) + n_{i'j}$. For constructing the lower confidence limit for $\mu_i/\mu_{i'}$, Fay[9] assumed that μ_i is distributed as gamma with mean r_i and variance

$\hat{\nu}_i$ and that $\mu_{i'}$ is distributed as gamma with mean $(r_{i'} + W_{i'})$ and variance $(\hat{\nu}_{i'} + W_{i'}^2)$ and used the result that, conditional on t_j,

$$\left(\frac{r_{i'} + W_{i'}}{r_i}\right)\frac{\mu_i}{\mu_{i'}} \sim F_{((2r_i^2/\hat{\nu}_i),(2(r_{i'}+W_{i'})^2/(\hat{\nu}_{i'}+W_{i'}^2)))}$$

where $W_{i'} = \max_{j:d_{i'j}<t_j}\{w_{i'j}\}$ and for independent $\chi_m^2 =^d G(m/2, 2)$ and $\chi_n^2 =^d G(n/2, 2)$, $F_{(m,n)} =^d (\chi_m^2/m)/(\chi_n^2/n)$ denotes the F distribution with numerator degrees of freedom (d.f.) m and the denominator d.f. n with density given by

$$g(x|m,n) = \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}\left(\frac{m}{n}\right)^{m/2}\frac{x^{(m/2)-1}}{(1+(m/n)x)^{(m+n)/2}}, \quad 0 < x < \infty$$

Since the numerator and the denominator chi-squared random variables in $F_{((2r_i^2)/\hat{\nu}_i),(2(r_{i'}+W_{i'})^2/(\hat{\nu}_{i'}+W_{i'}^2))}$ depend on t_j, the unconditional distribution of $\mu_i/\mu_{i'}$ is a mixture of F distributions, and not an F distribution.

The lower confidence limit is

$$\frac{r_i}{r_{i'} + W_{i'}} F^{-1}_{((2r_i^2)/\hat{\nu}_i,(2(r_{i'}+W_{i'})^2)/(\hat{\nu}_{i'}+W_{i'}^2))}\left(\frac{\alpha}{2}\right)$$

where $F^{-1}_{(a,b)}(p)$ is the pth percentile of $F_{(a,b)}$. Now, assuming that μ_i is distributed as gamma with mean $(r_i + W_i)$ and variance $(\hat{\nu}_i + W_i^2)$ and that $\mu_{i'}$ is distributed as gamma with mean $r_{i'}$ and variance $\hat{\nu}_{i'}$, Fay[9] derived the upper confidence limit to be

$$\frac{r_i + W_i}{r_{i'}} F^{-1}_{((2(r_i+W_i)^2)/(\hat{\nu}_i+W_i^2),(2r_{i'}^2)/\hat{\nu}_i'))}\left(1 - \frac{\alpha}{2}\right)$$

Note that this approximation cannot be readily applied for constructing CIs for the ratios μ_i/μ, that is, the ratios of the age-adjusted rates for the regions i to the US age-adjusted rates, as the latter depends on the former ones.

We propose a modification in the above CI for $\mu_i/\mu_{i'}$. For the lower limit, we assume that μ_i is distributed as gamma with mean r_i and variance $\hat{\nu}_i$ and that $\mu_{i'}$ is distributed as gamma with mean $\tilde{r}_{i'}$ and variance $\tilde{\nu}_{i'}$ and since the two distributions are independent chi-squares, we have

$$\left(\frac{\tilde{r}_{i'}}{r_i}\right)\frac{\mu_i}{\mu_{i'}} \sim F_{((2r_i^2)/\hat{\nu}_i,(2\tilde{r}_{i'}^2)/\tilde{\nu}_{i'})}$$

This results in the lower limit to be $r_i/\tilde{r}_{i'}\, F^{-1}_{((2r_i^2)/\hat{\nu}_i,(2\tilde{r}_{i'}^2)/\tilde{\nu}_{i'})}(\alpha/2)$. Similarly, the upper limit can be obtained. The proposed CI for $\mu_i/\mu_{i'}$ is

$$\left(\frac{r_i}{\tilde{r}_{i'}} F^{-1}_{((2r_i^2)/\hat{\nu}_i,(2\tilde{r}_{i'}^2)/\tilde{\nu}_{i'})}\left(\frac{\alpha}{2}\right), \frac{\tilde{r}_i}{r_{i'}} F^{-1}_{((2\tilde{r}_i^2)/\tilde{\nu}_i,(2r_{i'}^2)/\hat{\nu}_{i'})}\left(1 - \frac{\alpha}{2}\right)\right)$$

Another CI for $\mu_i/\mu_{i'}$ using $(r_i + \bar{w}_i)$ and $(\hat{\nu}_i + \bar{w}_i^2)$ instead of r_i and $\tilde{\nu}_i$ is given by

$$\left(\frac{r_i}{r_{i'} + \bar{w}_{i'}} F^{-1}_{((2r_i^2)/\hat{\nu}_i,(2(r_{i'}+\hat{w}_{i'})^2)/(\hat{\nu}_{i'}+\bar{w}_{i'}^2))}\left(\frac{\alpha}{2}\right), \frac{r_i + \bar{w}_i}{r_{i'}} F^{-1}_{((2(r_i+\bar{w}_i)^2)/(\hat{\nu}_i+\bar{w}_i^2),(2r_{i'}^2)/\hat{\nu}_i')}\left(1 - \frac{\alpha}{2}\right)\right)$$

Once again, we mention that this interval performs similarly to the above one, and we will not focus on this. We further remark that, unlike as in Fay,[9] these intervals do not assume the dependence of the w_{ij} on t_j.

2.2 Normal approximations

Define $R_{ij} = D_{ij}/n_{ij}(= D_{ij}/(n\xi_{ij}\xi_j))$. Let $n \to \infty$ so that $0 < \xi_{ij}, \xi_j < 1$. Note that $0 < \lambda_{ij} < \infty$. Then as $\min\{n_{ij}\lambda_{ij}\} \to \infty$,

$$\sqrt{n}\left(\frac{\xi_{ij}\xi_j}{\lambda_{ij}}\right)^{1/2}(R_{ij} - \lambda_{ij}) \longrightarrow^{\text{ind}} N(0,1), \quad i = 1,\ldots,I;\; j = 1,\ldots,J$$

That is, R_{ij} are independent and asymptotically normally distributed, $R_{ij} \sim AN(\lambda_{ij}, \lambda_{ij}/(n\xi_{ij}\xi_j))$. The other asymptotic results based on R_{ij}, $100(1-\alpha)\%$ CIs for $\mu_i, \mu, \mu_i/\mu_{i'}, \mu_i/\mu, \mu_i/\mu_{(-i)}$ and $\mu_i - \mu$, and their logarithmic and logit transformations are presented in Appendix C. In particular, the $100(1-\alpha)\%$ CIs for μ_i/μ, and $\mu_i - \mu$, based on the correlated age-adjusted rates, are given by

$$\frac{\mu_i}{\mu} = \left\{\frac{\hat{\mu}_i}{\hat{\mu}} \pm z_{\alpha/2}\frac{\sqrt{(\hat{\sigma}_i^2\hat{\mu}^2 + \hat{\sigma}^2\hat{\mu}_i^2 - 2\hat{\rho}_i\hat{\mu}_i\hat{\mu})}}{\sqrt{n\hat{\mu}^4}}\right\} \vee 0$$

$$\mu_i - \mu = \hat{\mu}_i - \hat{\mu} \pm z_{\alpha/2}\frac{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}^2 - 2\hat{\rho}_i}}{\sqrt{n}}$$

where $a \vee b = \max(a, b)$. When $\rho_i = 0$, which is true iff $\lambda_{ij} = 0$ for all j, these CIs reduce to (see, e.g., Ries *et al.*[10] for the CI of $\mu_i - \mu$ when $\rho_i = 0$)

$$\frac{\mu_i}{\mu} = \left\{\frac{\hat{\mu}_i}{\hat{\mu}} \pm z_{\alpha/2}\frac{1}{\sqrt{n}}\sqrt{\frac{1}{\hat{\mu}^4}(\hat{\sigma}_i^2\hat{\mu}^2 + \hat{\sigma}^2\hat{\mu}_i^2)}\right\} \vee 0, \quad \mu_i - \mu = \hat{\mu}_i - \hat{\mu} \pm z_{\alpha/2}\frac{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}^2}}{\sqrt{n}}$$

Since $\rho_i > 0$, the length of the CI for $\mu_i - \mu$, ignoring the adjustment for ρ_i, is wider, and hence the interval is more conservative.

2.3 Beta approximations

In general, the age-adjusted rates are less than 1 and equal to 1 if and only if there is one age-group with both the values of cancer counts and population at risk for that age group equal to 1, which is not a practical case. A rationale for the beta approximation is as follows. Let $R_i = \sum_{j=1}^{J} w_j R_{ij}$, where $R_{ij} = D_{ij}/n_{ij}$. Let D_{ij} and $\bar{D}_{ij}$ be independent Poisson random variables with means $n_{ij}\lambda_{ij}$ and $n_{ij}(1-\lambda_{ij})$, respectively. Then the distribution of $D_{ij}|_{D_{ij}+\bar{D}_{ij}=n_{ij},\lambda_{ij}} \sim \text{Bin}(n_{ij}, \lambda_{ij})$, a binomial distribution with parameters n_{ij} and λ_{ij}.[12] Using the result, given in Appendix D, we can approximate the distribution of R_i by a beta distribution with parameters $\hat{a}_i$ and $\hat{b}_i$, $\text{Be}(\hat{a}_i, \hat{b}_i)$, where

$$\hat{a}_i = \tilde{r}_i \left(\frac{\tilde{r}_i(1-\tilde{r}_i)}{\tilde{\nu}_i} - 1 \right), \quad \hat{b}_i = (1-\tilde{r}_i) \left(\frac{\tilde{r}_i(1-\tilde{r}_i)}{\tilde{\nu}_i} - 1 \right)$$

We define an approximate $100(1-\alpha)\%$ CI for μ_i as $(L_{\bar{R}_i}, U_{\bar{R}_i})$, where $L_{\bar{R}}$ and $U_{\bar{R}}$ are obtained by solving the following incomplete beta integrals:

$$\int_0^{L_{\bar{R}_i}} B(x|\hat{a}_i, \hat{b}_i)\,\mathrm{d}x = \frac{\alpha}{2}, \quad \int_0^{U_{\bar{R}_i}} B(x|\hat{a}_i, \hat{b}_i)\,\mathrm{d}x = 1 - \frac{\alpha}{2}$$

Here, $B(x|a, b)$ is the density of a beta distribution, $\text{Be}(a, b)$, with parameters a and b.

3 Examples and simulations

As an illustration, age-adjusted tongue cancer mortality rates were calculated for each of the regions. Tongue cancer occurs mostly among the elders. The 2002 mortality data for tongue cancer, even though available from the NCHS, were obtained from the SEER Program of the NCI (see the web site: www.seer.cancer.gov). We carried out two different simulation studies to evaluate the performance of the proposed gamma, beta and normal (with lower limits truncated at 0) intervals with the gamma interval of Fay and Feuer.[4] In the first simulation study, we took the true means of the Poisson distributions of D_{ij} to be the observed values of deaths in the (i, j)th cell, where i stands for the 51 regions of the US (50 states and Washington DC) and j stands for the 19 age-groups, to be $(i = 1, \ldots, 51;\ j = 1, \ldots, 19)$. Therefore, the true value of μ_i is the observed value of the age-adjusted rate for each i. From the Poisson distributions, we generated 10 000 values of d_{ij}, and obtained the observed values of the age-adjusted rates R_i using the normalized weights w_j, based on the 2000 US standard population, so that $\sum_{j=1}^{J} w_j = 1$. We computed approximate 95% CIs for μ_i for each of the 51 regions using the gamma intervals of Fay and Feuer[4] and the proposed gamma, beta and normal intervals. Additionally, we compared the F interval of Fay[9] for $\mu_i/\mu_{(-i)}$ with the proposed F and normal intervals (with left limits truncated at 0). We compared the age-adjusted rate of each of the 51 regions with the rest of US age-adjusted rate. Once again, we chose the year 2002 tongue cancer mortality age-adjusted rates for the 51 regions. The simulations were carried out assuming the 2000 standard population generating d_{ij} from independent Poisson with mean equal to the observed d_{ij}.

Table 1 gives the results of the first simulation study. Columns 2 and 3 of the table give the observed (true) tongue cancer mortality counts and age-adjusted rates (per 100 000 mid-year population) for the 50 states, the District of Columbia, and the four Census Bureau Regions (Northwest, Midwest, West, and South). Column 3 presents the empirical coverage probabilities of the 95% CIs for the (simulated) age-adjusted rates based on the gamma, modified gamma, beta, and normal approximations. Column 5 shows the observed (true) ratios of age-adjusted rates of each of the 51 regions with the rest of the US Column 6 gives the empirical coverage probabilities of the 95% CIs for the (simulated) rate ratios based on *F* modified *F* and normal approximations.

Both the modified gamma and modified *F* intervals are more efficient than their counterparts because their empirical coverage probabilities are at least 95%, but are lower than for the gamma and the *F* intervals. The beta and normal intervals are slightly liberal as they do not perform well as they have empirical coverage probabilities less than 95% for a number of regions.

In the second simulation study, we considered the effect of randomly generated values of w_{ij} and d_{ij} on the performance of the gamma, beta and normal intervals. Here, the subscript i does not play any role, and is treated as a dummy variable, but it is kept for the sake of notational consistency. We generated 19 numbers, corresponding to the $J = 19$ age-groups, from the uniform $U(0, 1)$ distribution and standardized them (and called them $w_{ij}; j = 1, \ldots, 19$) so that $\sum_{j=1}^{19} w_{ij}$ is a very small number, say, equal to 5.0×10^{-6}. We again generated 19 numbers from $U(0, 1)$ and standardized them (and called them $d_{ij}; j = 1, \ldots, 19$) so that their sum is small, $\sum_{j=1}^{19} d_{ij} = 20$. These standardized numbers were taken to be the values of the true means $\lambda_{ij}, j = 1, \ldots, 19$.

Then, we simulated 10 000 values of d_{ij} from the Poisson distributions with means $\lambda_{ij}, j = 1, \ldots, 19$. From these, we calculated the age-adjusted rates r_i and the 95% CIs for μ_i using the gamma and normal intervals. We also calculated the variance of w_{ij}. We repeated the entire process 500 times. Note that we could have standardized the sum $\sum_{j=1}^{19} w_{ij}$ to any other small number, but we chose it to be 5.0×10^{-6} so that it was similar to what we have based on the 2000 US standard population and the 2002 age-adjusted rates. We also could have standardized the sum $\sum_{j=1}^{19} d_{ij}$ to any other number than 20 possibly to 50, but we kept it to 20 to see the effect of small sample size; that is, the small number of total mortality counts for the region, i.

Note that out of 10 000 intervals, corresponding to each one of the 500 replications, it is expected that approximately 9500 intervals would contain the true mean μ_i and 500 would not; that is, it is expected that approximately 250 values of the lower limits would be above the true mean μ_i and about the same number of the upper limits would be below the true mean μ_i. In Figures 1 and 2, we plotted the 500 values of the variance of the normalized weights w_{ij} on the x-axis, and the frequencies of the lower and upper limits of μ_i for the Fay and Feuer[4] intervals, modified gamma, beta and normal intervals that fell, respectively, above and below the true mean μ_i, were plotted on the y-axis. In Figure 3, we plotted both the lower and the upper limits against the variance of w_{ij}. Note that the two solid lines in Figure 3 correspond to the lower and upper 95% confidence limits for true proportion, p, based on Bin(10 000, 0.05), and then rescaled by multiplying by 10 000; that is $10\,000(0.05 \pm 1.96\sqrt{0.05 \times 0.95/10000}) \approx (457, 543)$. Thus the expected

Table 1 Comparisons of empirical coverage probabilities for 95% CIs for the age-adjusted mortality rates of states/Census Bureau Regions and ratios of these rates to the rest of the US for tongue cancer

State/region	True count	True rate (per 100 000)	Coverage of 95% CI (rate)				True ratio	Coverage of 95% CI (ratio)		
			Gamma	Modified gamma	Beta	Normal		*F*	Modified *F*	Normal
Alaska	1	0.25	97.0	97.0	97.0	99.3	0.38	97.0	97.0	99.3
Wyoming	3	0.56	98.8	98.8	96.5	98.9	0.87	98.8	98.8	99.0
Montana	4	0.41	98.0	98.0	95.3	99.2	0.63	98.1	98.1	99.1
Vermont	4	0.58	98.1	98.1	95.0	99.3	0.89	98.3	98.3	99.2
Delaware	5	0.60	98.8	98.8	96.9	96.1	0.92	98.7	98.7	96.1
Rhode Island	5	0.45	98.6	98.6	96.6	95.2	0.69	98.4	98.4	96.1
Washington DC	6	1.06	97.9	96.4	94.1	97.2	1.62	97.9	96.8	97.0
Utah	6	0.34	97.3	96.7	94.8	96.8	0.52	97.7	96.8	96.6
Nebraska	8	0.46	96.8	96.8	95.1	94.9	0.70	96.8	96.7	95.2
South Dakota	8	0.92	97.7	97.1	95.1	96.2	1.42	97.7	97.1	96.2
New Mexico	9	0.48	96.8	96.2	94.4	96.5	0.73	97.0	96.4	96.2
West Virginia	9	0.41	97.5	97.5	95.7	96.4	0.62	97.8	97.6	96.5
North Dakota	10	1.51	97.6	97.6	96.1	95.7	2.32	97.2	97.0	95.7
Hawaii	12	0.89	96.7	96.3	94.9	95.8	1.37	96.8	96.5	95.9
Iowa	12	0.36	96.7	96.2	94.6	95.4	0.56	96.7	96.2	95.5
Idaho	13	1.05	96.5	96.1	94.6	95.6	1.61	96.5	96.0	95.5
Kansas	13	0.46	96.8	96.5	95.0	95.3	0.70	96.7	96.4	95.4
Maine	14	0.93	97.8	97.1	95.9	95.7	1.43	97.5	97.0	95.9
New Hampshire	14	1.10	97.1	96.6	95.1	95.7	1.69	97.0	96.6	95.7
Mississippi	15	0.53	96.4	96.2	95.0	95.0	0.81	96.7	96.4	95.4
South Carolina	16	0.40	96.6	96.4	95.1	95.5	0.61	96.5	96.2	95.4
Colorado	18	0.51	96.7	96.3	95.4	95.3	0.77	96.7	96.3	95.4
Oklahoma	19	0.52	96.5	96.2	95.3	95.0	0.80	96.7	96.2	95.2
Alabama	20	0.43	96.8	96.4	95.1	95.9	0.65	96.8	96.6	95.8
Arkansas	22	0.74	96.7	96.5	95.3	95.7	1.14	96.6	96.3	95.5
Kentucky	22	0.52	96.6	96.4	95.2	95.5	0.80	96.4	96.3	95.5
Louisiana	25	0.58	95.7	95.5	94.4	95.0	0.89	95.8	95.6	94.9
Arizona	26	0.47	96.4	96.1	95.0	95.7	0.72	96.3	96.2	95.5
Nevada	26	1.21	96.4	95.6	94.7	94.9	1.88	96.5	95.6	94.9
Connecticut	27	0.69	96.3	96.0	94.6	95.3	1.06	96.1	95.9	95.3
Oregon	27	0.75	96.2	96.0	95.0	95.2	1.15	96.3	96.1	95.2
Minnesota	31	0.63	96.1	95.9	95.0	95.1	0.96	95.9	95.8	95.0
Missouri	36	0.59	96.0	95.9	94.9	95.2	0.91	96.1	95.9	95.2
Georgia	38	0.52	96.3	95.9	95.2	95.1	0.79	96.3	96.0	95.1
Virginia	38	0.53	96.3	96.1	95.2	95.3	0.81	96.2	96.1	95.3
Massachusetts	39	0.56	96.2	96.0	95.2	95.3	0.85	96.3	96.1	95.3
Maryland	40	0.75	96.1	96.0	95.2	95.5	1.16	96.4	96.2	95.5
Indiana	42	0.67	96.0	95.8	95.0	95.3	1.03	96.0	95.9	95.3
Wisconsin	43	0.75	95.6	95.5	94.5	94.8	1.16	95.6	95.5	94.8
Washington	47	0.80	95.9	95.7	94.9	95.2	1.23	95.9	95.7	95.2
Tennessee	50	0.83	96.0	95.9	94.9	95.2	1.28	96.0	95.9	95.2
North Carolina	53	0.64	96.0	95.9	95.2	95.3	0.99	96.2	96.1	95.4
New Jersey	59	0.65	96.0	95.8	95.2	95.3	1.00	96.1	95.9	95.3
Illinois	65	0.53	95.6	95.6	94.9	95.0	0.80	95.5	95.5	94.9
Ohio	68	0.56	95.9	95.8	95.1	95.2	0.86	96.0	95.9	95.3
Michigan	76	0.75	95.4	95.3	94.5	94.6	1.16	95.4	95.3	94.8
Pennsylvania	86	0.60	95.9	95.8	95.0	95.3	0.91	95.9	95.8	95.2
New York	118	0.59	95.9	95.8	95.3	95.3	0.90	95.7	95.6	95.2
Texas	140	0.76	95.8	95.7	95.2	95.2	1.19	95.5	95.4	95.2
Florida	145	0.70	96.1	96.0	95.5	95.5	1.09	95.8	95.7	95.3
California	254	0.81	95.7	95.6	95.3	95.3	1.28	95.8	95.7	95.4
Northeast	366	0.62	95.7	95.6	95.3	95.2	0.94	95.7	95.7	95.2
Midwest	412	0.61	95.6	95.5	95.2	95.2	0.93	95.2	95.2	94.8
West	446	0.74	95.6	95.6	95.4	95.3	1.18	95.7	95.6	95.4
South	663	0.64	95.0	95.0	94.8	94.9	0.98	95.2	95.1	95.0

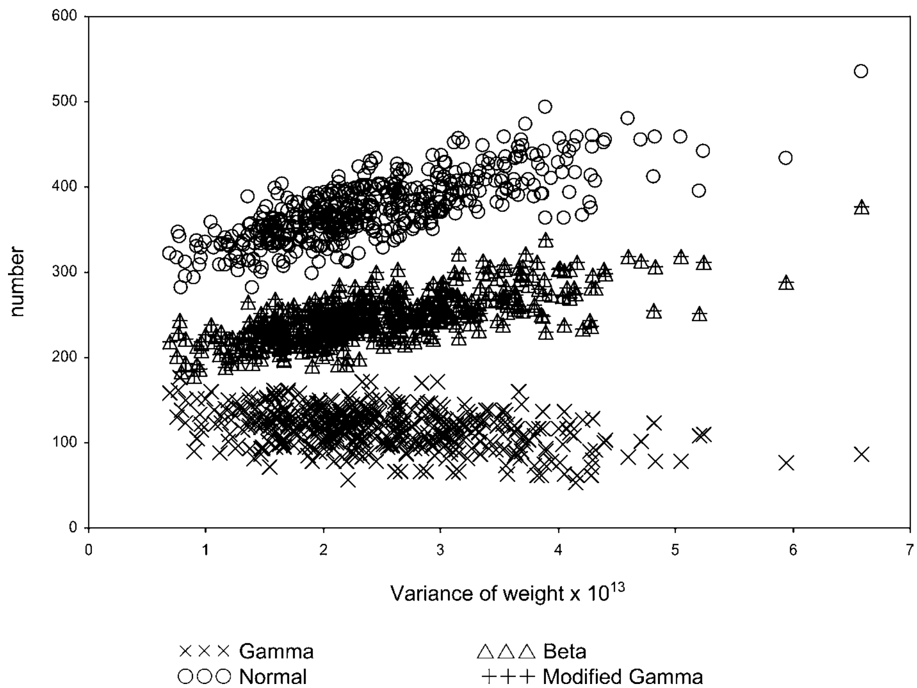

Figure 1 Number of upper limits below true mean (over 10 000 replications).

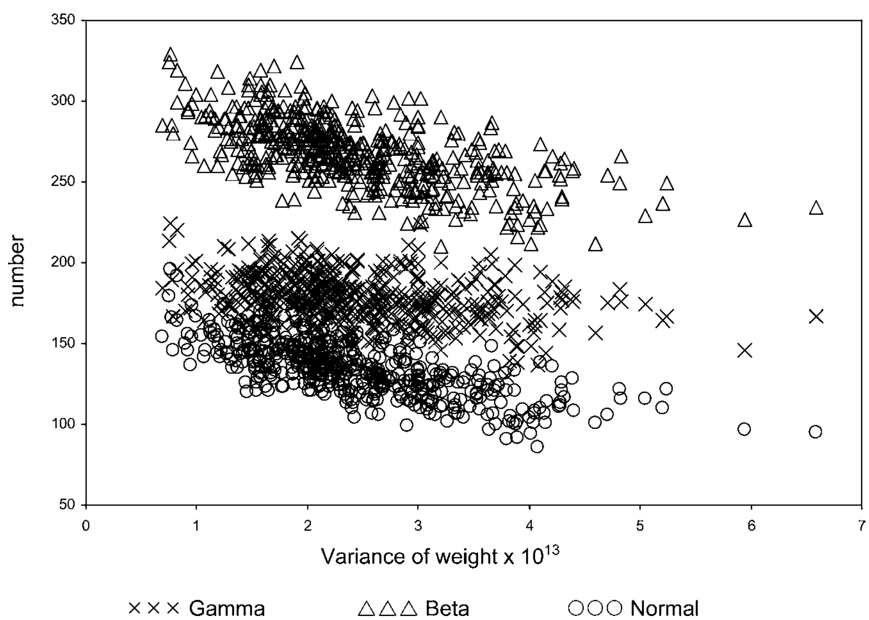

Figure 2 Number of lower limits above true mean (over 10 000 replications).

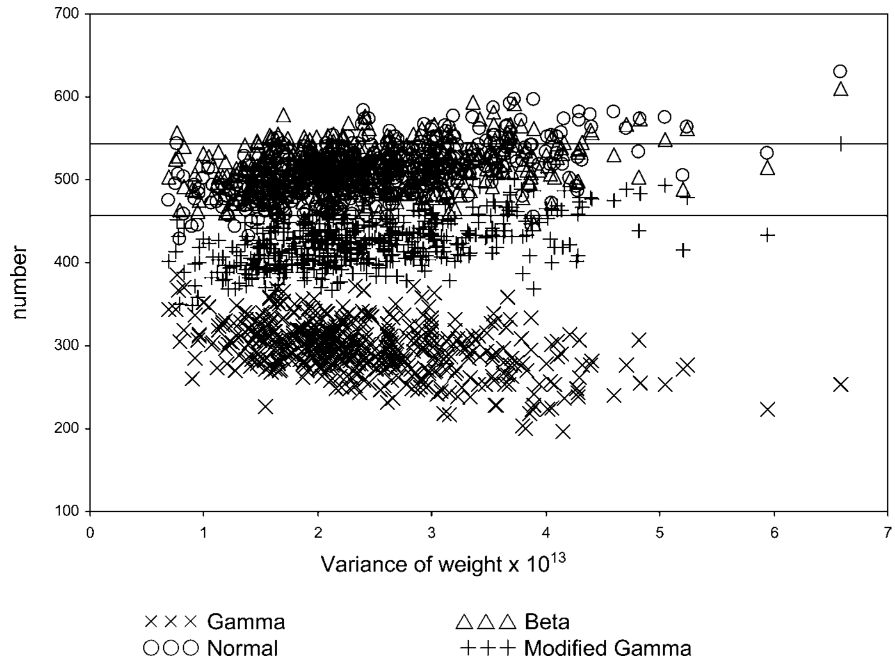

Figure 3 Number of CIs not covering true mean (over 10 000 replications).

numbers of the lower and upper limits of μ_i that fall above and below the true mean are between 457 and 543. In Figure 4, we plotted the lengths of the simulated intervals against the variance of w_{ij}.

From Figures 1–4, we observe that the modified gamma intervals have empirical coverage at least 95%, but slightly lower than the gamma intervals of Fay and Feuer,[4] the beta and normal intervals (with lower limits truncated at 0 if they were negative) also have empirical coverage probabilities very close to 95%, and their widths are lower than the gamma intervals. The coverage probabilities of the upper limits of both the beta and modified gamma intervals are identical and at least 97.5%, but slightly lower than the gamma intervals of Fay and Feuer.[4] The lower limits of the normal intervals are slightly more conservative than those for gamma, while the upper limits of the normal intervals are least conservative. The advantage of using modified gamma intervals over the gamma intervals is clear from Figure 3, wherein the gamma intervals show a coverage probability of around 97% as the variance w_{ij} increases, the modified intervals show the coverage probability staying slightly higher than 95%. Overall, from these simulation studies, the gamma intervals of Fay and Feuer[4] are more conservative than the proposed gamma. The beta intervals are slightly more liberal than both the modified gamma and the gamma intervals of Fay and Feuer.[4] The normal intervals are more liberal than the beta intervals.

In simulations, when the Poisson means were 0, as the observed d_{ij} were 0, we set the simulated values of D_{ij} to be equal to 0. This is because D_{ij} are non-negative random variables with the means and variances equal and if the mean of a D_{ij} is 0 then that D_{ij}

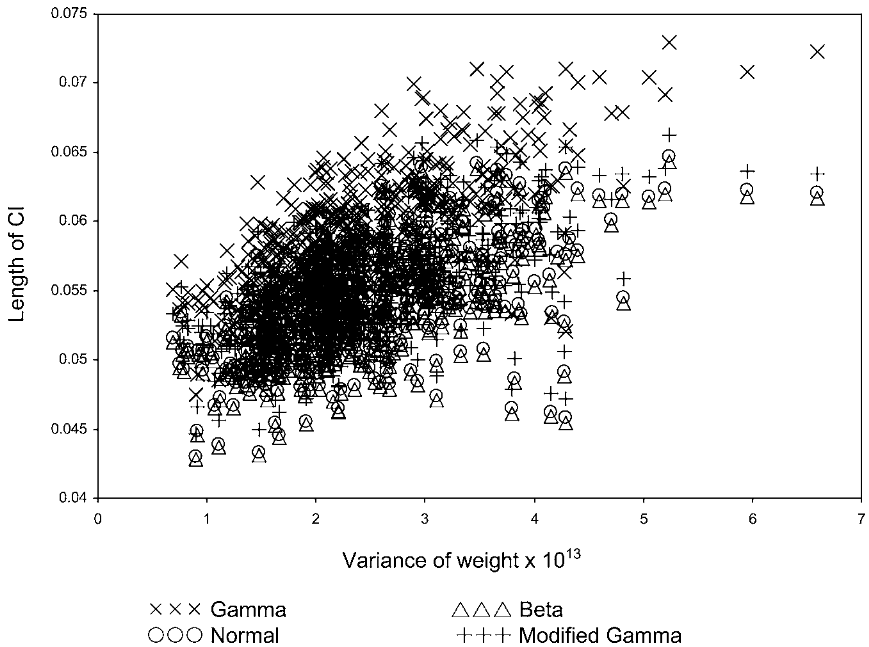

Figure 4 Length of CIs.

is 0 with probability 1. Of course, when D_{ij} have positive means, there is a good chance that the simulations could still result in 0 for the simulated values of D_{ij}. We considered another simulation study where we took the D_{ij} to be Poisson with means $n_{ij}r_{a,j}$, with $r_{a,j} = \sum_{i=1}^{I} d_{ij}/n_j$ as the observed 2002 US age-specific mortality rates for tongue cancer. Note that in this case, $\mu_i = \sum_{j=1}^{J} w_{ij}(n_{ij}r_{a,j})$ is a constant, independent of both i and j, and the ratio of the means of two age-adjusted rates is 1. The results of this study were very similar to those given in Table 1.

Next, we studied the performance of the 95% normal intervals for the ratios μ_i/μ and the differences $\mu_i - \mu$, and their coverage probabilities were close to 0.95. As an illustration, Figure 5 gives the plots of the number of CIs that do not contain the ratio of the observed age-adjusted mortality rates for Arkansas to the US, for the normal intervals, with lower limits truncated at 0 and with $\ln(\mu_i/\mu)$ transformation. For comparison, we also plotted these numbers for both the F and modified F intervals, ignoring the dependence of R_i on R. The Figure 5 shows that the F intervals are very conservative, the modified F intervals and the normal intervals based on the logarithmic transformation have coverage probabilities close to 0.95, and the normal intervals with lower limits truncated at 0 are slightly liberal. Of course, both the F and modified F intervals do not incorporate the crucial assumption of the dependence between R_i and R, and may not be appropriate in this context.

We also applied the normal CIs for $\mu_i - \mu$, to compare if the 2002 esophagus age-adjusted mortality rates for each of the 51 regions were equal to or not to the US

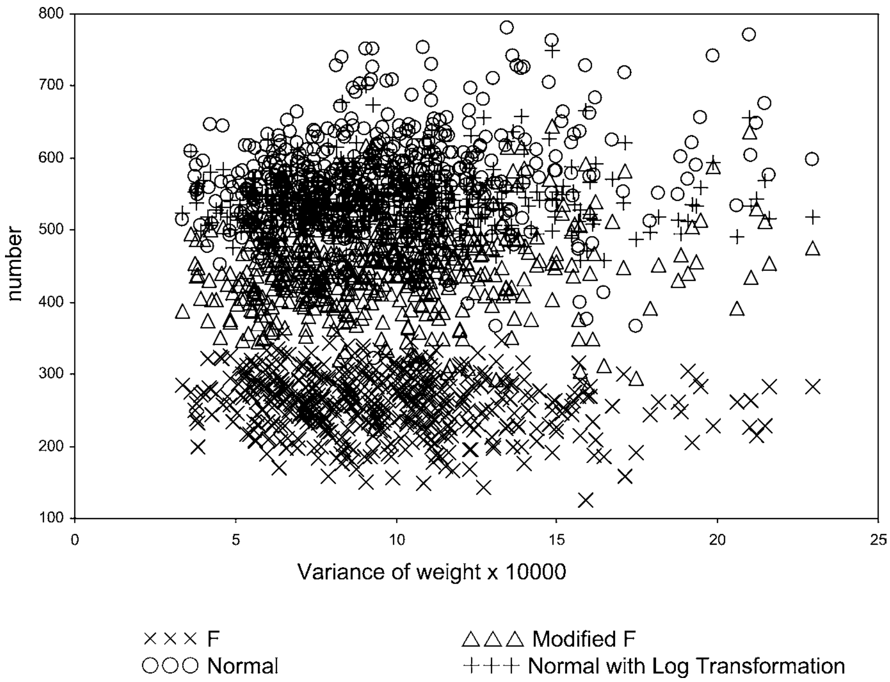

Figure 5 Number of CIs not covering true ratio of 2002 age-adjusted mortality rates for Arkansas to US for tongue cancer (over 10 000 replications).

age-adjusted rate using the 2002 esophagus mortality data. We found that the age-adjusted rates for Ohio and Pennsylvania were different from the US when we applied the normal CIs for $\mu_i - \mu$ with correlated R_i and R, as the CIs did not contain 0; whereas when we applied the CIs for $\mu_i - \mu$ based on the uncorrelated R_i and R, the age-adjusted rates for the two states were equal to that of the US as the CIs contained 0. For the other 49 regions, the two CIs produced results that were in agreement.

4 Discussion

The advantage of the modified gamma and F intervals is that they depend on all w_{ij} rather than just the largest value $w_{i(J)}$. The F intervals are based on the ratio of two chi-squared distributions that are independent and, unlike Fay,[9] do not depend on the restrictions $d_{ij} + d_{i'j} = t_j$ for all j. Also, the advantage of using the estimates of $\tilde{\mu}_i$, $\tilde{\sigma}_i^2$ and $\tilde{\rho}_i$, based on the continuity correction factor, over their counterparts $\hat{\mu}_i$, $\hat{\sigma}_i^2$ and $\hat{\rho}_i$, is more for the rare cancer sites. Without the adjustment for the continuity correction, the normal and beta CIs for μ_i, for the tongue cancer site, were observed to be liberal, especially for the regions with small mortality counts.

In Figures 1–4, we reported the performance of the gamma, modified gamma, beta and normal (with lower limits truncated at 0) intervals. We also studied, but did not report, the performance of the ABC, DKES, and normal intervals for μ_i based on the transformations $\ln(-\ln(R_i))$ and $\ln(R_i/(1-R_i))$. We observed that both the gamma and modified gamma intervals always retained the nominal coverage of at least 0.95, with the modified gamma intervals being less conservative than the gamma intervals. None of the other intervals retained the nominal coverage. The DKES intervals were next with the empirical coverage probabilities closer to the nominal value of 0.95, and then the beta intervals, the ABC intervals, the normal (with lower limits truncated at 0) intervals, the normal intervals based on $\ln(-\ln(R_i))$, and the normal intervals based on $\ln(R_i/(1-R_i))$, in that order. Similarly, for the CIs for the ratios of the means of two (uncorrelated) age-adjusted rates, both the F intervals of Fay[9] and the modified F intervals retained the nominal coverage of at least 0.95, with the modified F being less conservative of the two. The normal intervals (with lower limits truncated at 0) have coverage probabilities very close to 0.95 followed by the normal intervals based on the transformation $\ln(R_i/R_{(-i)})$.

We may mention that the beta intervals can be viewed as approximation to Bayesian credible intervals for μ_i. Assume that $0 < \lambda_{ij} < 1$ are small so that $D_{ij} \sim^{\text{ind}} \text{Bin}(n_{ij}, \lambda_{ij})$. Further assume that λ_{ij} are independent with prior $\pi(\lambda_{ij}) \propto 1$, $0 < \lambda_{ij} < 1$. Then the posterior distributions are given by

$$\lambda_{ij}|n_{ij}, r_{ij} \sim^{\text{ind}} \text{Be}(n_{ij}r_{ij}+1, n_{ij}(1-r_{ij})+1) \approx \text{Be}(n_{ij}r_{ij}, n_{ij}(1-r_{ij}))$$

and we can approximate the posterior means and variances of $\mu_i = \sum_{j=1}^{J} w_{ij}\lambda_{ij}$ by $\tilde{r}_i$ and $\tilde{v}_i$. Now, the credible intervals can be obtained as follows. Generate G^* (large) Markov chain Monte Carlo (MCMC) values on $\lambda_{ij}^{(g)}$, $g = 1, \ldots, G^*$, using Gibbs sampler, from the posterior distributions of λ_{ij}, and compute the G^* values of μ_i, namely, $\mu_i^{(g)} = \sum_{g=1}^{G^*} w_{ij}\lambda_{ij}^{(g)}$, and then construct the $100(1-\alpha)\%$ credible interval for μ_i from the empirical distribution of $\{\mu_i^{(g)}\}$, by ordering these values from the smallest to the largest and taking the credible interval to be the $100(\alpha/2)$th and $100(1-\alpha/2)$th ordered values. We performed MCMC simulations and constructed the credible intervals for the 2002 age-adjusted mortality rates for the tongue cancer for the 51 regions of the US and found that the credible intervals were more liberal than the beta intervals in Table1.

The assumption that the mortality or incidence counts are independent Poisson is used by many, for example, see Brillinger,[3] and is perhaps a consequence of the underlying birth/death (continuous) Poisson process model. We have not seen any analyses for the age-adjusted rates for the case of correlated D_{ij}. However, as pointed out by a referee, it is quite possible for neighboring states to have common socio-economic and other factors resulting in correlated D_{ij}s. This is an important topic for future research.

5 Conclusion

We presented CIs for the means of the cancer age-adjusted rates for the 51 regions, μ_i, for the US μ, for the ratios of the means $\mu_i/\mu_{i'}$, $\mu_i/\mu_{(-i)}$, μ_i/μ and for the differences $\mu_i - \mu$. We developed modifications of the gamma interval of Fay and Feuer,[4] and the *F* interval of Fay,[9] and proposed new CIs based on the beta and normal intervals. Simulations were carried out to compare the performance of these intervals in terms of their empirical coverage probabilities, and results showed that the modified gamma and *F* intervals performed better than the gamma interval of Fay and Feuer[4] and the *F* interval of Fay[9] in terms of retaining the nominal coverage. The other intervals such as the DKES, ABC, beta, and the truncated normal intervals were shown to be good competitors. The modified gamma and *F* intervals are going to replace the gamma and *F* intervals in the SEER Program. In addition, for comparing μ_i and μ, the normal intervals for $\mu_i - \mu$ that incorporate the correlation between R_i and R are also recommended to replace the ones that are based on the uncorrelated R_i and R in the SEER Program.[10] Even though the results of this paper are presented in the context of constructing the CIs for the (true) age-adjusted mortality rates based on data from the SEER Program, they can be applied to similar data from other countries as well.

Acknowledgements

The authors would like to thank the editor and the two referees for their valuable comments that led to a significant improvement of the original manuscript.

References

1 Jemal A, Murray T, Ward E, Samuels A, Tiwari RC, Ghafoor A, Feuer EJ, Thun MJ. Cancer Statistics 2005. *CA A Cancer Journal for Clinicians* 2005; **55:** 1–22.
2 American Cancer Society, *Cancer Facts & Figures* 2005.
3 Brillinger DR. The natural variability of vital rates and associated statistics (with discussion), *Biometrics* 1986; **42:** 693–734.
4 Fay MP, Feuer EJ. Confidence intervals for directly standardized rates: a method based on the gamma distribution. *Statistics in Medicine* 1997; **16:** 791–801.
5 DiCiccio T, Efron B. More accurate confidence intervals in exponential families. *Biometrika* 1992; **79:** 231–45.
6 Efron B, Tibshirani RJ. *An introduction to the bootstrap*. Chapman & Hall, 1993.
7 Swift MB. Simple confidence intervals for standardized rates based on the approximate bootstrap method. *Statistics in Medicine* 1995; **14:** 1875–88.
8 Dobson AJ, Kuulasmaa K, Ederle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. *Statistics in Medicine* 1991; **10:** 457–62.
9 Fay MP. Approximate confidence intervals for rate ratios from directly standardized rates with sparse data. *Communications in Statistics – Theory and Methods* 1999; **28:** 2141–60.
10 Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg LX, Edwards BK. *SEER Cancer Statistics Review, 1973–1997*. National Cancer Institute (NIH Pub. No. 00-2789), 2000.
11 Johnson NL, Kotz S. *Continuous univariate distributions – I*. Wiley, 1969.
12 Bickel PJ, Doksum KA. *Mathematical statistics: basic ideas and selected topics*, Holden-Day, Inc., 1977.
13 Breslow NE, Day NE. *Statistical methods in cancer research, volume II – the design and analysis of cohort studies*. Oxford University Press, 1987.
14 Casella G, Berger RL. *Statistical inference*. Wadsworth & Brooks/Cole Advanced Books & Software, 1990.

Appendix A: Means and variances of R_i, $R_{(-i)}$ and R, and of their ratios, and the covariance between R_i and R

We can rewrite R_i, $R_{(-i)}$ and R as

$$R_i = \frac{1}{n}\sum_{j=1}^{J} w_j \frac{D_{ij}}{\xi_j \xi_{ij}}; \quad R_{(-i)} = \frac{1}{n}\sum_{j=1}^{J} w_j \frac{D_{(-i)j}}{\xi_j \xi_{(-i)j}}; \quad R = \frac{1}{n}\sum_{j=1}^{J} w_j \frac{D_j}{\xi_j}$$

Let

$$\sigma_i^2 = \sum_{j=1}^{J} w_j^2 \frac{\lambda_{ij}}{\xi_j \xi_{ij}}; \quad \sigma_{(-i)}^2 = \sum_{j=1}^{J} w_j^2 \left(\frac{\sum_{i' \neq i}^{I} \xi_{i'j} \lambda_{i'j}}{\xi_j \xi_{(-i)j}^2} \right);$$

$$\sigma^2 = \sum_{j=1}^{J} w_j^2 \left(\frac{\sum_{i=1}^{I} \xi_{ij} \lambda_{ij}}{\xi_j} \right); \quad \rho_i = \sum_{j=1}^{J} w_j^2 \frac{\lambda_{ij}}{\xi_j}$$

Then,

$$\mu_i \equiv E(R_i) = \sum_{j=1}^{J} w_j \lambda_{ij}; \quad \mu_{(-i)} \equiv E(R_{(-i)}) = \sum_{j=1}^{J} w_j \frac{\sum_{i' \neq i}^{I} \xi_{i'j} \lambda_{i'j}}{\xi_{(-i)j}};$$

$$\mu \equiv E(R) = \sum_{j=1}^{J} w_j \left(\sum_{i=1}^{I} \xi_{ij} \lambda_{ij} \right)$$

$$v_i \equiv \mathrm{Var}(R_i) = \frac{\sigma_i^2}{n}; \quad v_{(-i)} = \frac{\sigma_{(-i)}^2}{n};$$

$$v \equiv \mathrm{Var}(R) = \frac{\sigma^2}{n}; \quad \mathrm{Cov}(R_i, R) = \frac{\rho_i}{n}$$

Using the delta-method, the means and variances of the ratios $R_i/R_{i'}$, $R_i/R_{(-i)}$ and R_i/R are given by

$$E\left(\frac{R_i}{R_{i'}}\right) \approx \frac{\mu_i}{\mu_{i'}}; \quad E\left(\frac{R_i}{R_{(-i)}}\right) \approx \frac{\mu_i}{\mu_{(-i)}}; \quad E\left(\frac{R_i}{R}\right) \approx \frac{\mu_i}{\mu}$$

$$\mathrm{Var}\left(\frac{R_i}{R_{i'}}\right) \approx \frac{\sigma_i^2 \mu_{i'}^2 + \sigma^2 \mu_i^2}{n \mu_{i'}^4}; \quad \mathrm{Var}\left(\frac{R_i}{R_{(-i)}}\right) \approx \frac{\sigma_i^2 \mu_{(-i)}^2 + \sigma^2 \mu_i^2}{n \mu_{(-i)}^4}$$

$$\mathrm{Var}\left(\frac{R_i}{R}\right) \approx \frac{\sigma_i^2 \mu^2 + \sigma^2 \mu_i^2 - 2\rho_i \mu_i \mu}{n \mu^4}$$

Appendix B: ABC and DKES intervals

The ABC intervals are[4]

$$L_{\mathrm{ABC}}(\mu_i;\alpha) = \hat{\mu}_i + \frac{z_{0i} - z_{\alpha/2}}{\{1 - a_i[z_{0i} - z_{\alpha/2}]\}^2}\frac{\hat{\sigma}_i}{\sqrt{n}}$$

$$U_{\mathrm{ABC}}(\mu_i;\alpha) = \hat{\mu}_i + \frac{z_{0i} + z_{\alpha/2}}{\{1 - a_i[z_{0i} + z_{\alpha/2}]\}^2}\frac{\hat{\sigma}_i}{\sqrt{n}}$$

where $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$ is the upper $\alpha/2$th percentile point of the standard normal distribution function, Φ, $a_i = z_{0i} = (\sum_{j=1}^{J} w_{ij}^3 d_{ij})/(6\hat{v}_i^{3/2})$. The DKES intervals are[4]

$$L_{\mathrm{DKES}}(\mu_i;\alpha) = \hat{\mu}_i + \frac{\hat{\sigma}_i}{\sqrt{n\sum_{j=1}^{J} d_{ij}}}\left[\frac{1}{2}\left(\chi^2_{2(\sum_{j=1}^{J} d_{ij})}\right)^{-1}\left(\frac{\alpha}{2}\right) - \sum_{j=1}^{J} d_{ij}\right]$$

$$U_{\mathrm{DKES}}(\mu_i;\alpha) = \hat{\mu}_i + \frac{\hat{\sigma}_i}{\sqrt{n\sum_{j=1}^{J} d_{ij}}}\left[\frac{1}{2}\left(\chi^2_{2(1+\sum_{j=1}^{J} d_{ij})}\right)^{-1}\left(1-\frac{\alpha}{2}\right) - \sum_{j=1}^{J} d_{ij}\right]$$

Appendix C: Asymptotic normality and confidence intervals based on R_{ij}

Let $\mathbf{R} = (R_{11},\ldots,R_{1J},\ldots,R_{I1},\ldots,R_{IJ})^{\mathrm{T}}$, $\bar{\mathbf{R}} = (R_1,\ldots,R_I,R)^{\mathrm{T}}$, $\mu = (\mu_1,\ldots,\mu_I,\mu)^{\mathrm{T}}$ and let $\sum = ((\sigma_{ij}))$ be $(I+1)\times(I+1)$ matrix with $\sigma_{ii} = \sigma_i^2$, $\sigma_{i,I+1} = \sigma_{I+1,i} = \rho_i$ and $\sigma_{ii'} = 0$ for $i \neq i'$. Here the superscript T denotes the transpose. Since $\mathbf{R}$ can be expressed as $\bar{\mathbf{R}} = A\mathbf{R}$ for an appropriately defined matrix A, we have

$$\sqrt{n}(\bar{\mathbf{R}} - \mu) \longrightarrow N_{(I+1)}\,(0,\Sigma)$$

where $\sum = A[\mathrm{Cov}(\mathbf{R})]A^{\mathrm{T}}$ and $N_p(\mathbf{b},B)$ denotes a p-dimensional multivariate normal distribution.

Thus for any non-null $(I+1)$-column vector $\mathbf{a}$,

$$\sqrt{n}\mathbf{a}^{\mathrm{T}}(\bar{\mathbf{R}} - \mu) \longrightarrow N(0,\mathbf{a}^{\mathrm{T}}\Sigma\mathbf{a})$$

In particular, by choosing **a** appropriately, we have

$$R_i = \sum_{j=1}^{J} w_j R_{ij} \sim^{\text{ind}} AN\left(\mu_i, \frac{\sigma_i^2}{n}\right)$$

$$R_{(-i)} = \sum_{j=1}^{J} w_j \left(\frac{\sum_{i' \neq i}^{I} \xi_{i'j} R_{i'j}}{\xi_{(-i)'j}}\right) \sim AN\left(\mu_{(-i)}, \frac{\sigma_{(-i)}^2}{n}\right)$$

$$R = \sum_{j=1}^{J} w_j \left(\sum_{i=1}^{I} \xi_{ij} R_{ij}\right) \sim AN\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\frac{R_i}{R_{i'}} \sim AN\left(\frac{\mu_i}{\mu_{i'}}, \frac{\sigma_i^2 \mu_{i'}^2 + \sigma_{i'}^2 \mu_i^2}{n\mu_{i'}^4}\right); \quad \frac{R_i}{R_{(-i)}} \sim AN\left(\frac{\mu_i}{\mu_{(-i)}}, \frac{\sigma_i^2 \mu_{(-i)}^2 + \sigma_{(-i)}^2 \mu_i^2}{n\mu_{(-i)}^4}\right)$$

$$\frac{R_i}{R} \sim AN\left(\frac{\mu_i}{\mu}, \frac{\sigma_i^2 \mu^2 + \sigma^2 \mu_i^2 - 2\rho_i \mu_i \mu}{n\mu^4}\right)$$

$$(R_i - R) \sim AN\left(\mu_i - \mu, \frac{\sigma_i^2 + \sigma^2 - 2\rho_i}{n}\right)$$

$$\mu_i = \left\{\hat{\mu}_i \pm z_{\alpha/2} \frac{\hat{\sigma}_i}{\sqrt{n}}\right\} \vee 0; \quad \mu = \left\{\hat{\mu} \pm z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}}\right\} \vee 0$$

$$\frac{\mu_i}{\mu_{i'}} = \left\{\frac{\hat{\mu}_i}{\hat{\mu}_{i'}} \pm z_{\alpha/2} \frac{\sqrt{(\hat{\sigma}_i^2 \hat{\mu}_{i'}^2 + \hat{\sigma}_{i'}^2 \hat{\mu}_i^2)}}{\sqrt{n\hat{\mu}_{i'}^4}}\right\} \vee 0$$

$$\frac{\mu_i}{\mu} = \left\{\frac{\hat{\mu}_i}{\hat{\mu}} \pm z_{\alpha/2} \frac{\sqrt{(\hat{\sigma}_i^2 \hat{\mu}^2 + \hat{\sigma}^2 \hat{\mu}_i^2 - 2\hat{\rho}_i \hat{\mu}_i \hat{\mu})}}{\sqrt{n\hat{\mu}^4}}\right\} \vee 0$$

$$\frac{\mu_i}{\mu_{(-i)}} = \left\{\frac{\hat{\mu}_i}{\hat{\mu}_{(-i)}} \pm z_{\alpha/2} \frac{\sqrt{(\hat{\sigma}_i^2 \hat{\mu}_{(-i)}^2 + \hat{\sigma}_{(-i)}^2 \hat{\mu}_i^2)}}{\sqrt{n\hat{\mu}_{(-i)}^4}}\right\} \vee 0$$

$$\mu_i - \mu = \hat{\mu}_i - \hat{\mu} \pm z_{\alpha/2} \frac{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}^2 - 2\hat{\rho}_i}}{\sqrt{n}}$$

where $a \vee b = \max(a, b)$.

Since $0 \leq R_i \leq 1$ and $0 \leq R_i/R_{(-i)} \leq \infty$ with probability 1, the following transformations are commonly used to transform the range of these random variables to $(-\infty, \infty)$ and their results on the asymptotic normality yield:

$$\ln(-\ln R_i) \sim AN\left(\ln(-\ln(\mu_i)), \frac{\sigma_i^2}{n(\mu_i \ln \mu_i)^2}\right)$$

$$\log\mathrm{it}(R_i) \equiv \ln\left(\frac{R_i}{1-R_i}\right) \sim AN\left(\ln\left(\frac{\mu_i}{1-\mu_i}\right), \frac{\sigma_i^2}{n(\mu_i(1-\mu_i))^2}\right)$$

$$\ln\left(\frac{R_i}{R_{(-i)}}\right) \sim AN\left(\ln\left(\frac{\mu_i}{\mu_{(-i)}}\right), \frac{1}{n}\left[\frac{\sigma_i^2}{\mu_i^2} + \frac{\sigma_{(-i)}^2}{\mu_{(-i)}^2}\right]\right)$$

Based on these transformations, the CIs for μ_i, $\mu_i/\mu_{(-i)}$ and μ_i/μ are given as follows:

I)

$$\mu_i = \exp\left\{-\exp\left[\ln(-\ln(\hat{\mu}_i)) \pm z_{\alpha/2}\frac{\hat{\sigma}_i}{(\hat{\mu}_i \ln \hat{\mu}_i)\sqrt{n}}\right]\right\}$$

II)

$$\mu_i = \left[1 + \exp\left\{-\left[\ln\left(\frac{\hat{\mu}_i}{1-\hat{\mu}_i}\right) \pm z_{\alpha/2}\frac{\hat{\sigma}_i}{(\hat{\mu}_i(1-\hat{\mu}_i))\sqrt{n}}\right]\right\}\right]^{-1}$$

III)

$$\frac{\mu_i}{\mu_{(-i)}} = \exp\left\{\ln\left(\frac{\hat{\mu}_i}{\hat{\mu}_{(-i)}}\right) \pm z_{\alpha/2}\left[\frac{1}{n}\left[\frac{\hat{\sigma}_i^2}{\hat{\mu}_i^2} + \frac{\hat{\sigma}_{(-i)}^2}{\hat{\mu}_{(-i)}^2}\right]\right]^{1/2}\right\}$$

IV)

$$\frac{\mu_i}{\mu} = \exp\left\{\ln\left(\frac{\hat{\mu}_i}{\hat{\mu}}\right) \pm z_{\alpha/2}\frac{\hat{\mu}}{\hat{\mu}_i}\left[\frac{1}{n}\frac{\hat{\sigma}_i^2\hat{\mu}^2 + \hat{\sigma}^2\hat{\mu}_i^2 - 2\hat{\rho}_i\hat{\mu}\hat{\mu}_i}{\hat{\mu}^4}\right]^{1/2}\right\}$$

The CIs in III) above, were also derived by Breslow and Day.[13] Note that we will use $\tilde{\mu}_i$, $\tilde{\nu}_i$ and $\tilde{\rho}_i$ instead of $\hat{\mu}_i$, $\hat{\nu}_i$ and $\hat{\rho}_i$.

Appendix D: Beta approximations of R_{ij} and R_i

Using the relation that[14]

$$\begin{aligned}\sum_{k=0}^{x} \binom{n}{k} p^k(1-p)^{n-k} &= \frac{\Gamma(n+1)}{\Gamma(x+1)\Gamma(n-x)} \int_0^{1-p} t^{n-x-1}(1-t)^x \, \mathrm{d}t \\ &= \int_0^{1-p} B(t|n-x, x+1) \, \mathrm{d}t \\ &= \int_p^1 B(t|x+1, n-x) \, \mathrm{d}t\end{aligned}$$

It then follows that

$$P(R_{ij} \geq r_{ij} | (D_{ij} + \bar{D}_{ij}) = n_{ij}, \lambda_{ij}) = \int_0^{\lambda_{ij}} B(t|n_{ij}r_{ij}+1, n_{ij}(1-r_{ij})) \, \mathrm{d}t$$

Another heuristic argument for the beta approximation for R_{ij} is based on the gamma or chi-squared approximation of a Poisson distribution. Let χ_k^2 and $\alpha\chi_k^2$ denote a chi-squared random variable with k degrees of freedom and a re-scaled (by a factor $\alpha > 0$) χ_k^2 random variable. Note that $\chi_k^2 =^d G(k/2, 1)$, and if χ_r^2 and χ_s^2 are independent, $\chi_r^2/(\chi_r^2 + \chi_s^2) =^d \chi_r^2/(\chi_{r+s}^2) \sim \mathrm{Be}(r/2, s/2)$, and that $\chi_r^2/(\chi_r^2 + \chi_s^2)$ and $\chi_r^2 + \chi_s^2$ are independent with $\chi_r^2 + \chi_s^2 =^d \chi_{r+s}^2$.

Since D_{ij} and $\bar{D}_{ij}$ are independent, distributed as $\mathrm{Po}(n_{ij}\lambda_{ij})$ and $\mathrm{Po}(n_{ij}(1-\lambda_{ij}))$, respectively, and their distributions can be approximated by independent chi-squared distributions $1/2\chi^2_{2([n_{ij}r_{ij}]+1)}$ and $1/2\chi^2_{2(n_{ij}-[n_{ij}r_{ij}])}$, where $[x]$ denotes the integer value of x, we have

$$\begin{aligned}\frac{D_{ij}}{D_{ij}+\bar{D}_{ij}} &\simeq \frac{1/2\chi^2_{2([n_{ij}r_{ij}]+1)}}{1/2\chi^2_{2([n_{ij}r_{ij}]+1)} + 1/2\chi^2_{2(n_{ij}-[n_{ij}r_{ij}])}} \\ &= \frac{\chi^2_{2([n_{ij}r_{ij}]+1)}}{\chi^2_{2([n_{ij}r_{ij}]+1)} + \chi^2_{2(n_{ij}-[n_{ij}r_{ij}])}} \sim \mathrm{Be}([n_{ij}r_{ij}]+1, n_{ij}-[n_{ij}r_{ij}]).\end{aligned}$$

Thus, $R_{ij} \sim \mathrm{Be}([n_{ij}r_{ij}]+1, n_{ij}-[n_{ij}r_{ij}])$. We can now approximate the distribution of $R_i = \sum_{j=1}^{J} w_j R_{ij}$ by a beta distribution, $\mathrm{Be}(\hat{a}_i, \hat{b}_i)$, where

$$\hat{a}_i = \tilde{r}_i \left(\frac{\tilde{r}_i(1-\tilde{r}_i)}{\tilde{\nu}_i} - 1 \right), \quad \hat{b}_i = (1-\tilde{r}_i) \left(\frac{\tilde{r}_i(1-\tilde{r}_i)}{\tilde{\nu}_i} - 1 \right)$$