Simple exact analysis of the standardised mortality ratio

F D K LIDDELL

From the Department of Epidemiology and Health, McGill University, Montreal, PQ, Canada

SUMMARY   The standardised mortality ratio is the ratio of deaths observed, D, to those expected, E, on the basis of the mortality rates of some reference population. On the usual assumptions—that D was generated by a Poisson process and that E is based on such large numbers that it can be taken as without error—the long established, but apparently little known, link between the Poisson and χ^2 distributions provides both an exact test of significance and expressions for obtaining exact $(1-\alpha)$ confidence limits on the SMR. When a table of the χ^2 distribution gives values for $1-\frac{1}{2}\alpha$ and $\frac{1}{2}\alpha$ with the required degrees of freedom, the procedures are not only precise but very simple. When the required values of χ^2 are not tabulated, only slightly less simple procedures are shown to be highly reliable for $D>5$; they are more reliable for all D and α than even the best of three approximate methods. For small D, all approximations can be seriously unreliable. The exact procedures are therefore recommended for use wherever the basic assumptions (Poisson D and fixed E) apply.

Given that D deaths occurred where E would have been expected had the mortality rates of some reference population applied, the standardised mortality ratio (SMR) is D/E, or often $100(D/E)$. For half a century or more, consideration has been given to tests of significance of the SMR and to its interval estimation. The usual assumptions have been that E is without error (because based on sufficiently large numbers) and that D was generated by a Poisson process; this note adopts the same assumptions.

Most proposed procedures are approximate, although exact methods have been presented.[1] Because approximations are still being debated,[2][3] this note aims to unify the exact procedures and show their practicality. That they are quite simple arises from the remarkable link, known to Fisher,[4] between the Poisson and χ^2 distributions, well described on page 10 of volume 1 of the *Biometrika Tables for Statisticians*[5] hereinafter designated *BTS*1. Table 8 of *BTS*1 (pages 136–7) gives values of χ^2 according to the degrees of freedom (ν) for which the upper tail probability $Q(\chi^2|\nu)$, defined on page 136, takes specified values. The link is that $Q(2m|2c)$ is identical to the cumulative sum of the terms of the Poisson series with parameter $m=\frac{1}{2}\chi^2$, up to but excluding the term for which $c=\frac{1}{2}\nu$. Although many tables of the χ^2 distribution gives values only for selected ν, the use of the Wilson/Hilferty approximation[6]—that is, the first in the footnote on

page 137 of *BTS*1—allows almost unlimited extension, because it is satisfactory even for quite small ν (>10, say),[7] although *BTS*1 suggests its use only for $\nu \geqslant 100$.

Also evaluated are three approximate methods, based as follows: (*a*) on the square-root transformation of the Poisson distribution, well known to be both normalising and variance-stabilising, so that sd $(\sqrt{D})$ can be taken as $\frac{1}{2}$ for all D;[8] (*b*) on the asymptotic normality of the distribution of D with standard error $\sqrt{D}$;[9] and (*c*) on the asymptotic normality of the distribution of D with standard error $\sqrt{E}$.[8]

Test of significance

The rationale is that, on the stated assumptions, the true test of significance is provided by the probabilities accumulated in the appropriate tail of the Poisson distribution with parameter E. This is always so, and the accumulation is not unduly tedious for reasonably small D. For larger D, however, the shortcut from the χ^2 link provides exact probabilities as follows:

$$\Pr(D \text{ or more deaths}|E) = 1 - Q(2E|2D); \quad (1)$$
$$\Pr(D \text{ or fewer deaths}|E) = Q(2E|2D + 2). \quad (2)$$

If $D>E$, expressions (1) and (2) give probabilities less than or greater than 0.5, respectively; the reverse is true when $D<E$. A one-sided test would use either (1) or (2), depending on the alternative being examined. For a two-sided test, expression (1) is used if $D>E$ and expression (2) if $D<E$; the (one-tailed) probability obtained should then be doubled, following Armitage.[8]

Q can be obtained from tables, such as table 8 of $BTS1$ for $D<15$ or *Documenta Geigy*[10] for all $D<100$. Tables of the χ^2 distribution usually provide only critical values. If the true p-value is required interpolation will be necessary; or it can be obtained accurately from many calculator or computer algorithms. For larger D, the above expressions can be replaced by the following, each yielding a standardised normal deviate z:[6]

$$z = 3(D)^{\frac{1}{3}}[1 - (9D)^{-1} - (D/E)^{-1/3}]; \qquad (1a)$$

$$z = 3(D')^{\frac{1}{3}}[(D'/E)^{-1/3} + (9D')^{-1} - 1], \qquad (2a)$$

where $D' = D + 1$. If $D>E$, z from (1a) is positive, the deviate from (2a) negative; if $D<E$, (1a) and (2a) yield negative and positive z, respectively. The p-values usually required are, as for Q, upper-tail so that positive z gives probabilities <0.5, and vice versa. The (one-tailed) p-value of z can readily be obtained—for instance, from table 1 of $BTS1$, pages 110–6, subtracting from unity for the upper-tail probability.

The approximate methods lead to standardised normal deviates as follows: (a) $z = 2(\sqrt{D} - \sqrt{E})$; (b) $z = (D - E)/\sqrt{D}$; and (c) $z = (D - E)/\sqrt{E}$.

Interval estimation

The basis for obtaining the $(1 - \alpha)$ confidence interval on the SMR is to determine the means, E_L and E_U, of two Poisson processes such that both $\Pr(D$ or more deaths$|E_L)$ and $\Pr(D$ or fewer deaths$|E_U) = \frac{1}{2}\alpha$; then $SMR_L = E_L/E$, $SMR_U = E_U/E$.

The exact limits are found as follows:

Lower limit: find χ^2_L for which $Q(\chi^2_L|2D) = 1 - \frac{1}{2}\alpha$;

then $E_L = \frac{1}{2}\chi^2_L$ and $SMR_L = \frac{1}{2}\chi^2_L/E$; $\qquad (3)$

Upper limit: find χ^2_U for which $Q(\chi^2_U|2D + 2) = \frac{1}{2}\alpha$;

then $E_U = \frac{1}{2}\chi^2_U$ and $SMR_U = \frac{1}{2}\chi^2_U/E$. $\qquad (4)$

For $D<15$, table 8 of $BTS1$ provides both χ^2_L and χ^2_U for six levels of $(1 - \alpha)$. For $D = 0(1)30(5)50$, table 40 of $BTS1$ (p 227) gives $\frac{1}{2}\chi^2_L$ and $\frac{1}{2}\chi^2_U$ for five levels of $(1 - \alpha)$. For other values of D and any α, the limits are:

$$SMR_L = D\{1 - (9D)^{-1} - z\cdot(9D)^{-\frac{1}{2}}\}^3/E; \qquad (3a)$$

$$SMR_U = D'\{1 - (9D')^{-1} + z\cdot(9D')^{-\frac{1}{2}}\}^3/E, \qquad (4a)$$

where z is the standardised normal deviate corresponding to $\frac{1}{2}\alpha$.

The limits from the approximate methods are: (a) $\{\sqrt{D} \pm (\frac{1}{2}z)\}^2/E$; (b) $\{D \pm z\sqrt{D}\}/E$; and (c) $\{D \pm z\sqrt{E}\}/E$; all using the same z.

Examples

Several examples are presented to show the universality and precision of the simple methods described above, and the comparative, but anticipated,[2] [8] [9] failure of the approximations when D is small. In the first example,[11] $D = 8$, $E = 3.59$, SMR $= 8/3.59 = 2.23$. For the test of significance we need $Q(7.18|16)$. From $BTS1$, table 8, $0.975 > Q > 0.950$ or, by the simplest interpolation recommended in $BTS1$ (p 16, footnote†), $Q \simeq 0.968$; from a calculator algorithm $Q = 0.969634$. Thus $\Pr(8$ or more deaths when 3.59 were expected$) = 1 - Q = 0.030366$ [or approximately 0.032, or lies between 0.025 and 0.05]. The first eight terms (from $r = 0$ to $r = 7$) of the Poisson series $e^{-3.59}(3.59)^r/(r!)$ accumulate to 0.969634, confirming the accuracy of the shortcut of expression (1).

The 90% confidence interval on 2.23 requires $1 - \frac{1}{2}\alpha = 0.95$, $\frac{1}{2}\alpha = 0.05$. For the lower and upper limits, $\nu = 16$ and 18, respectively, so that (from table 8 of $BTS1$) $\chi^2_L = 7.96165$ and $\chi^2_U = 28.8693$. Thus $SMR_L = 3.98082/3.59 = 1.1089$ and $SMR_U = 14.43465/3.59 = 4.0208$. Thus the 90% confidence interval on SMR does not embrace unity, and is in accord with a two-sided p-value close to 0.06. Confirmation of the accuracy of the confidence limits can be obtained from the Poisson series with expectations E_L and E_U; the relevant accumulations give both $\Pr(8$ or more deaths$|3.98082)$ and $\Pr(8$ or fewer deaths$|14.43465)$ as 0.05 (to at least 6 dp).

The first portion of the table summarises these results, and those from application of the three approximate methods.

The second example[2] had $D = 23$ and $E = 17.83$, with SMR $= 23/17.83 = 1.29$. Expression (1a) gives the test statistic (normal standardised deviate):

$$z = 3(23)^{\frac{1}{3}}[1 - (207)^{-1} - (23/17.83)^{-1/3}]$$
$$= (14.387495)[1 - 0.004831 - 0.918631]$$
$$= (14.387495)(0.076538) = 1.101189,$$

the one-sided p-value of which is 0.135408 by interpolation from table 1 of $BTS1$, or 0.135407 from two calculator algorithms. This differs by less

Comparison of exact and approximate methods

Method	Exact	Square-rooot transform	$sd(D) = \sqrt{D}$	$sd(D) = \sqrt{E}$
Example 1: $D = 8$, $E = 3\cdot59$, $(1 - \alpha) = 0\cdot9$				
p-value	0·0304	0·0309	0·0595	0·0100
SMR_L*	1·1089	1·1209	0·9325	1·3603
	(0·0500)	(0·0526)	(0·0213)	(0·1215)
SMR_U	4·0208	3·7128	3·5244	3·0966
	(0·0500)	(0·0856)	(0·1167)	(0·2218)
Example 2: $D = 23$, $E = 17\cdot83$, $(1 - \alpha) = 0\cdot95$				
p-value	0·1354†	0·1258	0·1405	0·1104
SMR_L	0·8175	0·8166	0·7628	0·8258
	(0·0249)	(0·0247)	(0·0124)	(0·0275)
SMR_U	1·9357	1·8710	1·8172	1·7541
	(0·0250)	(0·0381)	(0·053	(0·0773)
Example 3: $D = 210$, $E = 180$, $(1 - \alpha) = 0\cdot95$				
p-value	0·0156‡	0·0158	0·0192	0·0127
SMR_L	1·0142	1·0142	1·0089	1·0206
	(0·0250)	(0·0250)	(0·0210)	(0·0305)
SMR_U	1·3356	1·3298	1·3245	1·3128
	(0·0250)	(0·0291)	(0·0334)	(0·0447)

*Obtained by specified method; figures in brackets are the tail area probabilities, $Pr(D$ or more$|E_L)$ and $Pr(D$ or less$|E_U)$, found from the appropriate Poisson series.
†From 1(a); the true value is 0·1356.
‡Confirmed (to 4 dp) from the Poisson series.

than 0·0002 from the true (one-tailed) p-value found directly from the Poisson series as 0·135576.

For the desired 95% confidence limits, $z = 1\cdot96$, and expressions (3a) and (4a) lead to:

$$SMR_L = 23\{1 - (207)^{-1} - (1\cdot96)(207)^{-\frac{1}{2}}\}^3/17\cdot83$$
$$= 23(0\cdot858940)^3/17\cdot83 = 14\cdot575243/17\cdot83$$
$$= 0\cdot817456;$$

$$SMR_U = 24\{1 - (216)^{-1} + (1\cdot96)(216)^{-\frac{1}{2}}\}^3/17\cdot83$$
$$= 24(1\cdot128731)^3/17\cdot83 = 34\cdot513035/17\cdot83$$
$$= 1\cdot935672.$$

In this example it would have been possible to use table 40 of *BTS*1, for $c = D = 23$, bearing in mind the different convention over α; the values of $\frac{1}{2}\chi^2_L$ and $\frac{1}{2}\chi^2_U$ can be read as 14·58 and 34·51—that is, as above when rounded to 2 dp. The appropriate accumulations of the Poisson series with parameters E_L and E_U (to 6 dp) were 0·975079 and 0·024984. These results and the corresponding findings from the approximate methods are in the second part of the table in this paper.

The full comparison of methods for the third example[9] is presented at the foot of the table. In all these three examples, D was greater than E and the approximations taking $sd(D) = \sqrt{D}$ and $sd(D) = \sqrt{E}$ have been conservative and liberal, respectively. This was to have been expected, and the reverse has been found to hold in several further examples in which $D<E$. It was also to have been foreseen that the square-root transformation of D leads to the best of the approximate methods, and this is confirmed from the table. Indeed, the other two approximations can be unreasonable except for very large D.

When D is small, all the approximations are unreliable, and occasionally nonsensical. Consider $D = 1, E = 5$, for which the true p-value is $Q(10|4)$ which can be evaluated as 0·0404 $[= e^{-5}(1+5)$ also$]$. The exact 90% confidence limits on the SMR are found to be 0·0103 and 0·9488, the latter a little less than unity in conformity with the (one-tailed) p-value; these limits can easily be validated from the first one and two terms of the relevant Poisson series, provided all possible dp are carried in the calculation. Even the 99% confidence limits on the SMR—that is, 0·0010 and 1·4860—remain exact, as can be confirmed similarly. The square-root transformation is much too liberal, yielding P = 0·0067 and 90% confidence interval 0·0063 to 0·6643. The calculated value of "SMR_L" would be meaningless were $z>2\sqrt{D}$, or with $bd = 1$ when $\frac{1}{2}\alpha<0\cdot0228$. The approximations based on $sd(D) = \sqrt{D}$ are ridiculously liberal and with negative, hence impossible, "SMR_L": P = 0·00003; and $SMR_U = 0\cdot5290$. The p-value (0·0368) and SMR_U(0·9356) based on $sd(D) = \sqrt{E}$ might seem more acceptable in this example, but the "SMR_L" is also negative. On the other hand, even when $D = 0$, the p-value, that is $Q(2E|2) = e^{-E}$, and SMR_U can be evaluated precisely, while $SMR_L \equiv 0$.

Discussion

This paper has been concerned solely with the situation in which E can be taken as without sampling error, perhaps the most important circumstance in the study of SMRs, but certainly not the only one. (The recent statement[3] that Ederer and Mantel[12] had taken account of the possibility of errors in E as well as in D is misleading.)

All the theory for simple comprehensive exact procedures has long been available. Even expressions (1a), (2a), (3a), and (4a) have been quoted,[13] but only as an unpublished communication (from Byar), and without derivation, while there is some confusion over one-tailed and one-sided tests. Mulder[14] has recently drawn attention to part of the material of reference (1) and to table 40 (and the discussion) in *BTS*1. However, the theory never seems to have been presented together. This note has

shown how it is entirely practicable in all circumstances, except—unimportantly—for finding confidence limits with quite unusual $(1 - \alpha)$ when $D<15$. To avoid any possible confusion, the one approach (through the link between Poisson and χ^2) has been adopted throughout. For small enough D, however, the test by means of the Poisson tail area can be "quicker." The term "exact" is not strictly applicable when the Wilson/Hilferty approximation is used, but the lack of precision is considered negligible, unless D is smaller than say 5. It is substantially less than with the best of the approximate methods.

Where $D<15$, there are no difficulties in practice. For larger D, all the expression (1a) or (2a), (3a), and (4a) can be evaluated on a pocket calculator in about 90 seconds, provided that it has a means of providing cube-roots. A program for the HP–67 to provide (in 40 seconds) both test statistic and confidence intervals can be made available by the author on request; it is obviously more attractive than one proposed elsewhere which, for one process, requires several minutes of iteration.[13]

For $D<15$, the approximate methods are all "dirty" and not even as "quick" as the exact methods. For larger D, the only approximate method that is reasonably reliable is the one based on the square-root transformation of D; this is "quicker" than the use of the exact methods only when each of the three expressions has to be calculated without program. Even so, such calculation is not excessively "slow." Thus there seem no good reasons for not adopting the exact methods.

Finally, it should be remembered that tests should almost always be two-sided[8]; also that even when D is close to E, this does not necessarily obtain in every stratum over which standardisation has been carried out. The misleading nature of an SMR calculated in any inappropriate circumstances cannot be overemphasised.[15]

I thank Dr James A Hanley for many helpful discussions during the preparation of this paper.

References

1 Bailar JC, Ederer F. Significance factors for the ratio of a Poisson variable to its expectation. *Biometrics* 1964; **20:** 639–43.

2 Vandenbroucke JP. A shortcut method for calculating the 95 per cent confidence interval of the standardized mortality ratio. [Letter.] *Am J Epidemiol* 1982; **115:** 303–4.

3 Frentzel-Beyme, R. Re: 'A shortcut method for calculating the 95 per cent confidence interval of the standardized mortality ratio'. [Letter.] *Am J Epidemiol* 1982; **116:** 873–4.

4 Fisher RA. The mathematical distributions used in common tests of significance. *Econometrica* 1935; **3:** 353–65.

5 Pearson ES, Hartley HO. *Biometrika tables for statisticians.* Vol I. 3rd ed. London: Biometrika Trust, 1976.

6 Wilson EB, Hilferty MM. The distribution of chi-square. *Proc Natl Acad Sci* 1931; **17:** 684–8.

7 Liddell FDK. On the distributions of χ^2 and F. *Bulletin in Applied Statistics* 1983; **10:** 240–3.

8 Armitage P. *Statistical methods in medical research.* Oxford: Blackwell, 1971.

9 Registrar General. *Decennial supplement England and Wales 1951. Occupational mortality part II.* Vol 1. *Commentary.* London: HMSO, 1958: 18–9.

10 Diem K, ed. *Documenta Geigy: scientific tables.* 6th ed. Montreal: Geigy Pharmaceuticals, 1962: 36–9.

11 Dement JM, Harris RL, Symons MJ, Shy CE. Estimates of dose-response for respiratory cancer among chrysotile asbestos textile workers. *Ann Occup Hyg* 1982; **26:** 869–87.

12 Ederer F, Mantel N. Confidence limits on the ratio of two Poisson variables. *Am J Epidemiol* 1974; **100:** 165–7.

13 Rothman KJ, Boice JD. *Epidemiologic analysis with a programmable calculator.* Washington: US Government Printing Office, 1979: 29–31.

14 Mulder PGH. An exact method for calculating a confidence interval of a Poisson parameter. [Letter.] *Am J Epidemiol* 1983; **117:** 377.

15 Liddell FDK. The measurement of occupational mortality. *Br J Ind Med* 1960; **17:** 228–33.