

Refer to the article "Perceived age as clinically useful biomarker of ageing: cohort study" by Christensen et al., published in the 2009 Christmas Edition of the BMJ, and available on the website www.biostat.mcgill.ca/hanley/c634/Rates. The site also has a dataset, with 1 record for each of the 18 covariate patterns ("cells"), containing the cell-specific mortality information in Tables 1 and 2, as well as some R code that can speed up some of the 'from scratch' calculations below.

1 One Rate

- i. Calculate the *overall* mortality rate, along with a 95%CI, for the (840) *men* in the full sample. Calculate the CI in 3 ways: using the identity link (the untransformed numerator), the log transform, and the sqrt transform.¹ Explain why the three CI's differ from each other as much/little as they do.

$$\hat{\lambda} = \widehat{ID} = c/PY = 348/4547.3 = 0.0765 = 7.65 \text{ per } 100PY$$

		95%CI for λ :	
- method:	Computation	λ_L	λ_U
- untransformed:	$\{348 \pm 1.96 \times 348^{1/2}\}/4547.3$	0.0685	0.0846
- log:	$\exp\{\log[\hat{\lambda}] \pm 1.96 \times (1/34)^{1/2}\}$	0.0689	0.0850
- sqrt:	$\{\sqrt{348 \pm 1.96 \times 0.5}\}^2/4547.3$	0.0687	0.0848

- ii. Calculate the mortality rate for the 119 *women* in the '1st third' in the *youngest age-category* (leftmost of the 9 data-columns of Table 2). Obtain the associated CI in 2 ways: using the exact method and using the identity link (the untransformed numerator - the one Rothman2002 uses on p132 of Ch 7). Suppose you are going to report the CI using 2 significant figures: do the two versions differ at the second significant figure? (see http://en.wikipedia.org/wiki/Significant_figures).

		95%CI for λ :	
- method:	Computation	λ_L	λ_U
- exact:	Table: $\{4.80, 18.39\}/781.2$	0.00614	0.0235
- exact:	Stata <code>cii 7812 10,poisson</code>	0.000614	-.00235
- exact:	R <code>epicalc ci.poisson(10,781.2)</code>	0.00613	0.0236
- Z SE's (as is):	$\{10 \pm 1.96 \times 10^{1/2}\}/781.2$	0.00487	0.0207

¹This first time, do the CI calculations 'from scratch'. Then check them against CI's obtained from a statistical package, such as the immediate command `cii` in **Stata**, the `pois.*` functions in the `epitools` package in R, `ci.poisson` in the `epicalc` R package, etc.

- iii. There were 675 deaths in the two sexes combined. From the age- and sex-specific mortality rates in Denmark's population during the same calendar period (<http://www.mortality.org>), some xxx.x deaths would have been expected in a cohort of this same age-sex composition followed for this same amount of time. Calculate an SMR, and test whether it is significantly different from 100. Suggest reasons why/why not.

We don't have sufficient data to calculate the xxx.x exactly, since to do so we would need to know the more specific age-distribution of those in the combined 80+ category, as well as (for everyone) the entry and last dates of each person, and to calculate from them how many person years each person spend in each of the (say) 5×5 age-calendar time "square" in the Lexis diagram², then multiply the total person years within the square by the population death rate in that square, and sum over all squares to get the xxx.x. To illustrate how it is done, the following calculations assume that on average, 2/3rds of the PY were lived in the calendar period 2000-2004 and 1/3rd in the period 2006-2007. With this assumption we can do exact calculations for the persons who started out in the 70-74 age category, spend 2/3rds of their PT in that 5×5 square, and then spend 1/3rd of their time in the 75-79 cell in 2005-2007. As is apparent in the calculations - they advance diagonally: that birth subcohort of the population experienced mortality rates of 0.02666 in the first square and 0.03934 in the next (75-79 2005-2007) one (see rates in Q4) . We can also do the calculations for those who were initially 75-79. For the sake of completeness (even though it leads to an underestimate of the expected numbers) we assume that those aged 80- at entry are in fact aged 80-84.

Age cat.	PY	Expected Deaths if 2/3 PY lived in initial Sex-Age 5 x 5 cell & 1/3 in next 5 x 5 cell.	E	O
F 70-	2244.2	$\frac{2}{3} \times PY \times 0.02666 + \frac{1}{3} \times PY \times 0.03934 =$	69.3	66
M 70-	2264.6	...	103.9	95
F 75-	1706.9	$\frac{2}{3} \times PY \times 0.04179 + \frac{1}{3} \times PY \times 0.06559 =$	84.9	76
M 75-	1362.6	...	103.5	111
F 80-	1763.2	...	148.7	185
M 80-	920.1	...	113.4	142
$\Sigma = \text{xxx.x} = \text{"Expected" number} = E_{Denmark}$			623.7	675

SMR = $675/623.7 = 1.08$ (108). This assumes that all of those in the 'Initially 80-' category were 80-84. If we restrict the calculations to the known numbers in the 70-74 and 75-79 categories, then , under the same assumptions about the 1/3 1/3 partition of the PT, we get SMR = $348/361.6 = 0.96$ (96).

A test of whether the observed SMR is statistically significantly different from 100 is equivalent to testing if the 675 is statistically significantly different from 623.7, or 348 from 361.6, i.e. at the extreme of a Poisson

²Google Lexis diagram

distribution with mean $\mu = xxx.x$. Since we can quite safely use the Gaussian (Normal) approximation to a Poisson distribution with a 3-digit mean, we can calculate a p-value based on the z-statistic $\frac{675-xxx.x}{\sqrt{xxx.x}}$. The 95%CI for the SMR can be calculated by dividing the CI for the numerator (the CI based on the 675 or 348) by the corresponding E (the xxx.x). Since E is based on the entire Danish population, it is considered to have (relatively speaking) negligible sampling variation relative to the (Poisson) sampling variability in the numerator of the SMR.

You can think of the SMR as having a ‘2-sample’ structure where one is finite, and the other – the reference – effectively infinite but scaled down to have the same size and same age-sex-followuptime distribution as the (index) cohort.

Reasons for a non-null SMR: It may be than twins have different mortality rates than non-twins. Or it might be that to participate in this follow-up round (or earlier rounds) one would need to be healthier than average; note that there were some exclusions based on cognition.

2 Comparison of 2 Rates (unstratified)

- i. Calculate the *overall* mortality rate for the (986) *women* in the full sample.

$$\hat{\lambda}_W = \hat{ID}_W 327/5714.3 = 0.0572 \text{ per PY.}$$

Calculate the (crude) Men:Women mortality rate ratio (i.e., with men as the index, and women as the reference category), along with an associated 95%CI.³

$$\widehat{IDR}_{M:W} = 0.0765/0.0572 = 1.34.$$

- method:	Computation	95%CI	
		IDR_L	IDR_U
- log:	$\exp\{\log[\widehat{IDR}] \pm 1.96 \times (\frac{1}{348} + \frac{1}{327})^{1/2}\}$	1.15	1.56
Stata:	<code>iri 348 327 4547.3 5714.3</code>	1.15	1.56
R epitools:	<code>rateratio(c(327,348,5714.3,4547.3))</code>	1.15	1.56

- ii. Calculate the mortality rate ratio using as the *reference* category the 119 *women* in the ‘1st third’ in the *youngest age-category* and as the

³This first time, do the CI calculations ‘from scratch’. Then check them against those from a statistical package, such as the immediate command `iri` in **Stata** or the `rateratio.*` functions in `epitools`

index category the 119 *women* in the ‘2nd third’ in the *youngest age-category* (second of the 9 data-columns of Table 2). Calculate the associated 95%CI in 2 ways: using an exact method (e.g. using `iri` in **Stata**) and using the log link (the one Rothman2002 uses on p137-138 of Ch 7). Again, suppose you are going to report the CI using 2 significant figures: do the two versions differ at the second significant figure?

$$\widehat{IDR}_{2nd:1st} = (23/752.5)/(10/781.2) = 2.39.$$

- method:	Computation	95%CI	
		IDR_L	IDR_U
- log:	$\exp\{\log[\widehat{IDR}] \pm 1.96 \times (\frac{1}{23} + \frac{1}{10})^{1/2}\}$	1.14	5.02
Stata:	<code>iri 23 10 752.5 781.2</code>	1.09	5.62
R epitools:	<code>rateratio(c(10,23, 781.2,752.5))</code>	1.15	5.24

3 Comparison of 2 Rates - stratified data

- i. In section 2 part i, you calculated a crude (crude) Men:Women mortality rate ratio. Before you do any further calculations, how much (and in which direction) do you expect an age-standardized or otherwise age-adjusted M:W mortality rate ratio to differ from the crude rate ratio? Explain.
- ii. Compute an age-adjusted M:W mortality rate ratio. Do so in four ways: (a) informally, using the median of the three age-category-specific M:W mortality rate ratios; (b) using classical age-standardization; (c) using a weighed average of the three log-rate-ratios (similar to Woolf’s method, except that each variance only involves the sums of the reciprocals of the numbers of cases in the index and reference categories, rather than the four reciprocals when dealing with log-odds-ratios); and (d) using the Mantel-Haenszel summary rate ratio.⁴

(a) From the table below, we see that the 3 stratum-specific \widehat{IDR} ’s are 1.43, 1.83 and 1.47. The **median** of these is 1.47 (the mean is a bit higher– 1.58). In practice, we would put more weight on the rate ratios that were more precisely measured, and less on those less well measured (see below).

(b) **Age-standardization**: this also usually implies some type of weighted average, but with the weights pre-specified and not driven by

⁴This first time, do the CI calculations ‘from scratch’. Then check the answers from the second and fourth ways against those from a statistical package, such as the `ir` command in **Stata** or the appropriate functions in `epicalc` and `Epi`. Rothman2002Ch8 covers pooled rate ratios starting on p 153.

precision considerations.

One option is to use an *external* set of weights (e.g when comparing cancer incidence, as in the WHO book Cancer Incidence in Five Continents, standardized cancer incidence rates are computing using a “World standard population” set of weights: see <http://www.who.int/whosis/indicators/compendium/2008/1mst/en/>). The weights are thus somewhat arbitrary. For example we could decide, for our purposes to take weights close to the world standard, and say to use weights of say 220 150 and 110 (i.e. 46%, 32% and 22%).

This weighted average of the 3 age-specific rates 0.0294, 0.0445 and 0.1049 for females is 0.0508; the corresponding one for the 3 rates for men 0.0419, 0.0814 and 0.1543 is 0.0793. Thus their ratio is $0.0793/0.0508 = 1.56$.

Another option is an *internal* set of weights, such as a composite of the age or PY distribution of the aggregated male and female cohorts: 4508.8, 3069.5 and 2683.3 or approx. 0.44 0.30 and 0.26. These gives weighted averages of 0.0537(F) and 0.0832(M), and a M:F ratio of 1.55.

JH prefers to describe these averages by the nature of the weights, rather than by the name ‘*direct*’ standardization that is commonly associated with this method. ‘*Indirect*’ standardization is a term used to describe the process by which we arrived at the xxx.x in the earlier question, i.e., by multiplying the *PY in the cohort* by the rates in the reference population. The terms ‘*direct*’ and ‘*indirect*’ are not very explanatory, and also, with a bit of algebra one can show that the SMR is a weighed average of the stratum-specific rate ratios (if interested, see the excellent 1960 article on the Measurement of Occupational Mortality by Liddell and the comments on standardization in JH’s notes on stratified data (both under resources for stratified data in c634).

(c) **Antilog of (precision)-weighted average of $\log[\widehat{IDR}]$ ’s :**

Applying the formula $Var[\log\{\widehat{IDR}\}] = 1/c_1 + 1/c_0$ to the data in each age-stratum, we obtain the 3 variances $V = 0.0257, 0.0222$ and 0.0124 . The smaller variance in the 3rd (oldest) stratum reflects the larger numbers of deaths in that stratum, and the correspondingly more stable rate ratio. Thus the $Var[\log\{\widehat{IDR}\}]$ in that stratum should get more weight in the weighted average. Using the inverses (i.e., reciprocals) of the variances and scaling them so they add to 1, we get the weights 0.24, 0.27 and 0.49. Applying these to the 3 log-rate-ratios of 0.355, 0.604 and 0.386, we obtain the weighted average $w.a. = 0.438$. The antilog of this is $\exp(0.438) = \widetilde{IDR}_{w.a.} = 1.55$.

The variance of the weighed average is $Var(w.a.) = \frac{1}{1/V_1+1/V_2+1/V_3} =$

$\frac{1}{\frac{1}{Info_1+Info_2+Info_3}} = 0.006083$ so the (multiplicative) 95%*ME* to accompany the 1.55 is $M.E. = \exp[1.96 \times \sqrt{0.006083}] = 1.165$. Thus, the limits are $IDR_L = 1.55 \div 1.165 = 1.33$ and $IDR_U = 1.55 \times 1.165 = 1.81$.

Note: you can also use the stratum specific variances to get stratum-specific CI’s :- {1.04 to 1.95}, {1.37 to 2.45} and {1.18 to 1.83}. Each of these contains the summary rate ratio, so we have no *statistical* contra-indications to merging the 3 ratios.

(d) **Mantel-Haenszel summary \widetilde{IDR} :**

The ‘crude’ M:F ratio of 1.34 is a mixed-up or confused ratio, since it contrasts (somewhat younger) men and (somewhat older) women.

The magnitudes of the 3 ‘numerator’ and 3 ‘denominator’ contributions tell us that the Mantel-Haenszel summary \widetilde{IDR} is also a precision-based summary: the largest contribution to the ratio is from the oldest stratum.

The single ratio, i.e., the sum of the ‘numerator contributions’ divided by the sum of the ‘denominator contributions’ is $202.3/130.3=1.55$. The variance of the log of Mantel-Haenszel summary \widetilde{IDR} , calculated in same way as in Rothman2002 page 156, is 0.006071, is just about the same as that in (c), and so the CI will also be the same to 2 decimal places.

The weighted log-ratios method (c) does not work well for *sparse* stratum-specific data, such as those in Mantel and Haenszel’s own 1959 paper, whereas the MH method works well all the way down to the *finest possible strata*, such as the twin pairs analyzed in Figure 2 in the BMJ article.

M	age	deaths	PY	rate	\widehat{IDR}	MH <i>den</i> *	$\frac{c \times PT_1 \times PT_0}{PT^2}$
	cat	(c)		(c/PY)		<i>num</i> *	
0	1	66	2244.2	0.0294	Ref	33.1	
1	1	<u>95</u>	<u>2264.6</u>	0.0419	1.43	47.3	
		161	4508.8				40.24
0	2	76	1706.9	0.0445	Ref	33.7	
1	2	<u>111</u>	<u>1362.6</u>	0.0814	1.83	61.7	
		187	3069.5				46.16
0	3	185	1763.2	0.1049	Ref	63.4	
1	3	<u>142</u>	<u>920.1</u>	0.1543	1.47	93.3	
		327	2683.3				73.67
0	Σ	327	5714.3	0.0572	Ref	130.3	
1	Σ	348	4547.3	0.0765	1.34	202.3	
							160.09
	Σ				$\widehat{IDR}_{MH} = \frac{202.3}{130.3} = 1.55$		
						$V_{log} = \frac{160.09}{202.3 \times 130.3} = 0.006071$	

**num* = $c_1 \times PY_0/PY$; *den* = $c_0 \times PY_1/PY$;

iii. In section 2 part ii, you calculated the ‘1st-third versus 2nd third’ mortality rate ratio for just one of the six sex-age strata. Calculate the corresponding ratios for the other 5 strata. Comment.

- The six rate ratios, from 70- M to 80- M to 70- W to 80- W, are 0.85, 0.75, 1.12, 2.39, 1.56 and 1.98. Whereas one might be tempted to think the pattern represented ‘effect-modification’ by gender, it may be nothing than random variation. One way to formally check this out is to calculate and plot the 6 CIs and see if they share ‘common ground’ i.e., if the summary ratio is contained within each of the 6 (one might allow one to ‘miss’).

Obtain a single (summary) age-and-sex-adjusted M:W mortality rate ratio from the six stratum-specific ones, along with the associated 95%CI.

- The Mantel-Haenszel summary rate-ratio is 1.31. The (multiplicative) ME is 1.22. So we can see that the 1.31 is statistically significantly above unity: the 95%CI is $1.31 \div 1.22$ to 1.31×1.22 .

The antilog of the weighted average of the log-rate-ratios is 1.30. The (multiplicative) ME is again 1.22.

Before you had done these calculations, how much (and in which direction) would you have expected the age-and-sex-adjusted M:W mortality rate ratio to differ from the crude⁵ rate ratio? Explain.

- The crude IDR is $(221/3495.7)/(176/3623.2) = 0.0632/0.0486 = 1.30$. The fact that the crude and the age-sex-adjusted rate ratios agree in this instance should not be surprising, since within each age-sex stratum the (person) denominators are balanced: within each stratum there are as many persons in the ‘2nd third’ as there are in the ‘1st third’. The experience in the aggregated 6 ‘1st thirds’ has the same age-sex distribution as the experience in the aggregated 6 ‘2nd thirds’, so it is a fair comparison.

4 Sex-Age-CalendarTime Patterns in population mortality rates in Denmark

...Year.....Age...Female...Male.....Total... Observed rates

1980-1984	70-74	0.02725	0.05213	0.03814
1980-1984	75-79	0.04592	0.08235	0.06042
1980-1984	80-84	0.08098	0.12163	0.09561
1980-1984	85-89	0.13680	0.18202	0.15193

2000-2004	70-74	0.02666	0.03972	0.03261
2000-2004	75-79	0.04179	0.06586	0.05189
2000-2004	80-84	0.06923	0.10584	0.08279
2000-2004	85-89	0.11970	0.16773	0.13480

2005-2007	70-74	0.02359	0.03468	0.02874
2005-2007	75-79	0.03934	0.05815	0.04750
2005-2007	80-84	0.06559	0.09622	0.07730
2005-2007	85-89	0.11462	0.15808	0.12860

⁵The one computed by computing the rate in the ‘1st-third’ as the sum of the 6 numerators divided by the sum of the 6 PT denominators, and similarly for the rate in the ‘2nd-third’.