Methods for estimating adjusted risk ratios

Peter Cummings
Department of Epidemiology
School of Public Health and
Harborview Injury Prevention and Research Center
University of Washington
Seattle, WA
peterc@u.washington.edu

Abstract. The risk ratio can be a useful statistic for summarizing the results of cross-sectional, cohort, and randomized trial studies. I discuss several methods for estimating adjusted risk ratios and show how they can be executed in Stata, including 1) Mantel–Haenszel and inverse-variance stratified methods; 2) generalized linear regression with a log link and binomial distribution; 3) generalized linear regression with a log link, normal distribution, and robust variance estimator; 4) Poisson regression with a robust variance estimator; 5) Cox proportional hazards regression with a robust variance estimator; 6) standardized risk ratios from logistic, probit, complementary log-log, and log-log regression; and 7) a substitution method. Advantages and drawbacks are noted for some methods.

Keywords: st0162, risk ratio, odds ratio

1 Introduction

The case–control study design is typically (but not always) used when outcomes are rare in the population from which study subjects are sampled. In 1951, Cornfield noted that when outcomes are sufficiently rare, the odds ratio from a case–control study will approximate the population risk ratio for the association of an exposure with a disease outcome. It was later realized that if controls are sampled as each case arises in time, the odds ratio will estimate the incidence-rate ratio even when outcomes are common (Greenland and Thomas 1982; Rodrigues and Kirkwood 1990; Rothman, Greenland, and Lash 2008, 113–114). In Stata, case–control data can be analyzed using Mantel–Haenszel stratified methods (cc, tabodds, mhodds), logistic regression (logistic), or conditional logistic regression (clogit) to estimate adjusted odds ratios that usually can be interpreted either as risk ratios (when outcomes are rare) or incidence-rate ratios (when incidence density sampling is used).

Cross-sectional, cohort, and randomized controlled trial designs with binary outcomes can often be summarized by estimating odds ratios or risk ratios. If the study outcome is sufficiently rare among exposed and unexposed study subjects, the odds ratio for the exposure–outcome association will closely approximate the risk ratio. But if the outcome is common and the risk ratio is not close to 1, the odds ratio will be further from 1 compared with the risk ratio. Even if the outcome is rare in the entire sample, if an adjustment is made for other variables, then the adjusted odds ratio will
Methods for estimating adjusted risk ratios

be further from 1 than the adjusted risk ratio if the outcome is common in adjustment variable subgroups that contribute a noteworthy portion of the outcomes (Greenland 1987).

When summary odds and risk ratios differ, there is debate regarding which is preferable. Some have argued that odds ratios are preferred because they are symmetric with regard to the outcome definition (Walter 1998; Olkin 1998; Senn 1999; Newman 2001, 35–40; Cook 2002). Furthermore, when outcomes are common, a constant (homogeneous) adjusted odds ratio for all subjects may be more plausible than a constant risk ratio (Levin 1991; Senn 1998; Cook 2002).

Some who favor risk ratios feel they are more easily understood by physicians (Sackett, Deeks, and Altman 1996). Others have noted that risk ratios have a desirable feature called collapsibility; in the absence of confounding, a weighted average of stratum-specific risk ratios will equal the ratio from one 2 × 2 table of the pooled (collapsed) counts from the stratum-specific tables (Miettinen and Cook 1981; Greenland 1987, 1991b; Greenland, Robins, and Pearl 1999; Newman 2001, 52–55; Rothman, Greenland, and Lash 2008, 62). This means that a crude (unadjusted) risk ratio will not change if we adjust for a variable that is not a confounder. In the absence of confounding, the risk ratio estimates the change in risk, on a ratio scale, for the entire exposed group due to exposure. Because of collapsibility, this risk ratio has a useful interpretation as the ratio change in the average risk in the exposed group due to exposure. It is not the average ratio change in risk (i.e., the average risk ratio) among exposed individuals, except in the unlikely event that the risk ratios for all individuals are the same (Greenland 1987).

Odds ratios lack the property of collapsibility and therefore the interpretation of an odds ratio is more limited; in the absence of confounding, it estimates the change in odds, on a ratio scale, in the exposed group due to exposure. But it does not estimate either the change in the average odds of the exposed due to exposure or the average change in odds (i.e., the average odds ratio) among exposed individuals, not even if all individuals had the same change in odds when exposed (Greenland 1987). The odds ratio will estimate the average change in odds for exposed individuals only if all individual odds ratios are the same and all individual risks without exposure are the same. Except in this unlikely situation, the crude odds ratio will be closer to 1 than the average of stratum-specific or individual odds ratios. Even in the absence of confounding, the adjusted (conditional) odds ratio will be further from 1 than the crude (unadjusted or marginal) odds ratio (Gail et al. 1984; Greenland 1987; Hauck et al. 1998; Steyerberg et al. 2000; Newman 2001, 52–55; Rothman, Greenland, and Lash 2008, 62; Cummings 2009).

For analysts who wish to estimate odds ratios for the association of exposure with disease in a cross-sectional study, cohort study, or randomized trial, the statistical methods in Stata’s `cc`, `tabods`, `mhosds`, and `logistic` commands can be used. If the goal is to estimate risk ratios, these same methods can be used if outcomes are sufficiently rare that odds ratios will closely approximate risk ratios. But if risk ratios are desired when outcomes are common, odds ratio estimates will not suffice. In this article, I describe methods for estimating adjusted risk ratios with confidence intervals (CIs) in Stata.
2 Data used to illustrate the methods

I will show how to reproduce the risk-ratio estimates and CIs that Greenland (2004a) gave in a review of risk-ratio estimation. The data (table 1) are from table 5.3 in Newman’s (2001, 98 and 126) textbook. Newman described 192 women who were diagnosed with breast cancer in Canada and followed for 5 years; 28% (54/192) of the women died, so the outcome was not rare. Greenland estimated the risk ratio for death at 5 years among women with low estrogen-receptor levels in their breast cancer tissue compared with women who had high receptor levels; these risk ratios were adjusted for cancer stage (I, II, or III) so that women with the same cancer stage were compared.

Table 1. Deaths, total subjects, and risk of death for 192 women with breast cancer followed for 5 years, by stage at diagnosis (I, II, III) and estrogen-receptor–level category (low, high). Also, risk ratios within each cancer stage for death among women with low versus high receptor levels.

<table>
<thead>
<tr>
<th>Receptor levels</th>
<th>Stage</th>
<th>Died</th>
<th>Total</th>
<th>Risk</th>
<th>Risk ratio for death comparing women with low versus high receptor levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>I</td>
<td>2</td>
<td>12</td>
<td>0.17</td>
<td>1.8</td>
</tr>
<tr>
<td>High</td>
<td>I</td>
<td>5</td>
<td>55</td>
<td>0.09</td>
<td>1.0 (reference group)</td>
</tr>
<tr>
<td>Low</td>
<td>II</td>
<td>9</td>
<td>22</td>
<td>0.41</td>
<td>1.8</td>
</tr>
<tr>
<td>High</td>
<td>II</td>
<td>17</td>
<td>74</td>
<td>0.23</td>
<td>1.0 (reference group)</td>
</tr>
<tr>
<td>Low</td>
<td>III</td>
<td>12</td>
<td>14</td>
<td>0.86</td>
<td>1.4</td>
</tr>
<tr>
<td>High</td>
<td>III</td>
<td>9</td>
<td>15</td>
<td>0.60</td>
<td>1.0 (reference group)</td>
</tr>
</tbody>
</table>

3 Method 1: Mantel–Haenszel and inverse-variance stratified methods

Mantel–Haenszel methods for odds ratios were described in 1959 (Mantel and Haenszel 1959) and extended to risk ratios in 1981 (Nurminen 1981; Tarone 1981; Kleinbaum, Kupper, and Morgenstern 1982; Newman 2001, 148–149; Rothman, Greenland, and Lash 2008, 274–275). We can estimate the adjusted risk ratio for death associated with low estrogen-receptor levels (the low variable) compared with high estrogen-receptor levels by using the cs command:

(Continued on next page)
Methods for estimating adjusted risk ratios

. use brcadat
   (Breast cancer data)
. cs died low, by(stage) pool

<table>
<thead>
<tr>
<th>Cancer stage</th>
<th>RR</th>
<th>[95% Conf. Interval]</th>
<th>M-H Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.833333</td>
<td>.4024728</td>
<td>8.35115</td>
</tr>
<tr>
<td>2</td>
<td>1.780749</td>
<td>.9269437</td>
<td>3.420991</td>
</tr>
<tr>
<td>3</td>
<td>1.428571</td>
<td>.8971062</td>
<td>2.274888</td>
</tr>
</tbody>
</table>

Crude          | 2.225806 | 1.449035               | 3.418974   |
Pooled (direct)| 1.554553 | 1.076372               | 2.245169   |
M-H combined   | 1.618421 | 1.093775               | 2.394719   |

Test of homogeneity (direct)  \( \chi^2(2) = 0.339 \) Pr>\( \chi^2 = 0.8443 \)
Test of homogeneity (M-H) \( \chi^2(2) = 0.385 \) Pr>\( \chi^2 = 0.8251 \)

I invoked the `pool` option so that the output shows both the Mantel–Haenszel combined risk ratio and the pooled risk ratio obtained using inverse-variance weights. These methods require that variables be treated as categorical, not continuous.

4 Method 2: Generalized linear regression with a log link and binomial distribution

Estimation of risk ratios using a generalized linear model with a log link and binomial distribution was proposed in 1986 (Wacholder 1986). This approach has been described in several articles (Robbins, Chao, and Fonseca 2002; McNutt et al. 2003; Barros and Hirakata 2003); it has been called log-binomial (Blizzard and Hosmer 2006) or binomial log-linear regression (Greenland 2004a). This approach can be implemented in Stata:
. glm died low stage2 stage3, family(binomial) link(log) eform difficult

Iteration 0:  log likelihood = -154.81266  (not concave)
Iteration 1:  log likelihood = -99.23223
Iteration 2:  log likelihood = -94.764125
Iteration 3:  log likelihood = -93.93508
Iteration 4:  log likelihood = -93.050613
Iteration 5:  log likelihood = -92.930394
Iteration 6:  log likelihood = -92.927072
Iteration 7:  log likelihood = -92.927069

Generalized linear models  No. of obs = 192
Optimization : ML  Residual df = 188
Scale parameter = 1
Deviance = 185.8541388  (1/df) Deviance = .9885858
Pearson = 190.2212968  (1/df) Pearson = 1.011815
Variance function:  V(u) = u*(1-u)  [Bernoulli]
Link function :  g(u) = ln(u)  [Log]
AIC = 1.009657
Log likelihood = -92.92706941  BIC = -802.555

<table>
<thead>
<tr>
<th></th>
<th>OIM</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>died</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>1.558321</td>
<td>.3148624</td>
<td>2.20</td>
<td>0.028</td>
</tr>
<tr>
<td>stage2</td>
<td>2.538159</td>
<td>.9991488</td>
<td>2.37</td>
<td>0.018</td>
</tr>
<tr>
<td>stage3</td>
<td>5.868042</td>
<td>.273768</td>
<td>4.57</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Above Stata reported that for iteration 0, the likelihood region was not concave. When the command is run without the `difficult` option, Stata 10.0 will repeatedly report a not-concave region and fail to converge. The `difficult` option changed Stata’s convergence algorithm and solved the problem in this example, but that option may not always work. The convergence problem arose because among women with Stage III cancer and low estrogen-receptor levels, the risk of death was close to 1: 12/14 = 0.86. When the risk is close to 1 in a stratum of the data, maximum-likelihood convergence may fail. This problem has been discussed in several articles (Carter, Lipsitz, and Tilley 2005; Blizzard and Hosmer 2006; Lumley, Kronmal, and Ma 2006; Localio, Margolis, and Berlin 2007).

Wacholder (1986; Lumley, Kronmal, and Ma 2006) described a method that modified the convergence by truncating estimated risks to values slightly greater than 0 and less than 1. This is implemented in Stata’s `binreg` command:

(Continued on next page)
Methods for estimating adjusted risk ratios

```stata
.xi: binreg died low i.stage, rr
i.stage  _Istage_1-3 (naturally coded; _Istage_1 omitted)
Iteration 1:  deviance = 309.6253
Iteration 2:  deviance = 190.79
Iteration 3:  deviance = 185.9416
Iteration 4:  deviance = 185.8543
Iteration 5:  deviance = 185.8541
Iteration 6:  deviance = 185.8541
Generalized linear models  No. of obs =       192
Optimization : MQL Fisher scoring  Residual df =       188
             (IRLS EIM)  Scale parameter =        1
Deviance = 185.8541388  (1/df) Deviance = .98858588
Pearson = 190.2164034  (1/df) Pearson = 1.011789
Variance function: V(u) = u*(1-u) [Bernoulli]
Link function  : g(u) = ln(u) [Log]
                 BIC =      -802.555
```

| died     | Risk Ratio | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|----------|------------|-----------|------|------|----------------------|
| low      | 1.558326   | 0.3067419 | 2.25 | 0.024 | 1.059515            | 2.291972 |
| _Istage_2| 2.538158   | 0.9976133 | 2.37 | 0.018 | 1.174781            | 5.483782 |
| _Istage_3| 5.868047   | 2.259727  | 4.60 | 0.000 | 2.755899            | 12.48196 |

The risk ratios and standard errors estimated by `glm, family(binomial) link(log)` and by `binreg, rr` are similar, but not identical, because the convergence methods differ. The default for `glm` is maximum likelihood and the default for `binreg` is iterated, reweighted least squares; by changing the default optimization, either command can produce the estimates obtained by the other, provided that both methods achieve convergence. Because `binreg` constrains risk estimates to be greater than 0 and less than 1, it may converge when maximum likelihood will not. Sometimes both methods will fail to converge.

Another approach to convergence difficulty is to make several copies of the original data and append them into one data file (Deddens, Petersen, and Lei 2003; Petersen and Deddens 2006; Deddens and Petersen 2008). Then make one more copy, but recode all the 0 outcomes to 1 and all the 1 outcomes to 0 in that copy, and append this recoded copy to all the other copies. Then analyze all these data together; including one set of data with reversed outcomes may help the maximum-likelihood algorithm converge. If the number of copies is sufficiently large, the risk-ratio estimates will approximate those from maximum-likelihood methods. Because a set of records larger than the original data is used, corrections must be made to the standard errors and CIs. This extra step of correcting the standard errors can be avoided by using just two copies of the data, one with recoded outcomes, with appropriate weights (Lumley, Kronmal, and Ma 2006). Below I used importance weights of 0.999 and 0.001, and produced the risk ratio I would get by analyzing 999 copies of the original data with just 1 copy of recoded data. The `difficult` option was still required.
. generate iweight = .999
. append using brcadat
(label noyes already defined)
. recode died 0=1 1=0 if iweight==.
(died: 192 changes made)
. replace iweight = .001 if iweight==.
(192 real changes made)
. glm died low stage2 stage3 [iweight=iweight], family(binomial) link(log)
> eform difficult
Iteration 0: log likelihood = -155.22946 (not concave)
Iteration 1: log likelihood = -99.401459
Iteration 2: log likelihood = -94.953215
Iteration 3: log likelihood = -94.13882
Iteration 4: log likelihood = -93.23838
Iteration 5: log likelihood = -93.113209
Iteration 6: log likelihood = -93.109614
Iteration 7: log likelihood = -93.109611

Generalized linear models No. of obs = 384
Optimization : ML Residual df = 380
Scale parameter = 1
Deviance = 186.2192211 (1/df) Deviance = .4900506
Pearson = 190.253604 (1/df) Pearson = .5006674
Variance function: V(u) = u*(1-u) [Bernoulli]
Link function : g(u) = ln(u) [Log]
AIC = .5057792
Log likelihood = -93.10961053 BIC = -2075.025

| died  | Risk Ratio | Std. Err. | z   | P>|z|  | [95% Conf. Interval] |
|-------|------------|-----------|-----|------|---------------------|
| low   | 1.556724   | .3143539  | 2.19| 0.028| 1.047915            |
| stage2| 2.523499   | .9897355  | 2.36| 0.018| 1.169918            |
| stage3| 5.823722   | 2.248418  | 4.56| 0.000| 2.732558            |

5 Method 3: Generalized linear regression with a log link, normal distribution, and robust variance estimator

Convergence problems for generalized linear regression with a log link can also be resolved by using a Gaussian (normal) distribution (Lumley, Kronmal, and Ma 2006). The resulting standard errors may be too big or small, but a robust variance estimator will correct the standard errors. Below convergence was achieved without the difficult option:

(Continued on next page)
Methods for estimating adjusted risk ratios

. use brocadat, clear
(Breast cancer data)
. glm died low stage2 stage3, family(gaussian) link(log) eform robust

Iteration 0: log pseudolikelihood = -169.42788
Iteration 1: log pseudolikelihood = -119.04978
Iteration 2: log pseudolikelihood = -98.365962
Iteration 3: log pseudolikelihood = -94.332246
Iteration 4: log pseudolikelihood = -94.242364
Iteration 5: log pseudolikelihood = -94.242364

Generalized linear models

No. of obs = 192
Optimization : ML
Residual df = 188
Scale parameter = .1595924
(1/df) Deviance = .1595924
Pearson = 30.00336652 (1/df) Pearson = .1595924

Variance function: V(u) = 1
Link function : g(u) = ln(u)

AIC = 1.023358
Log pseudolikelihood = -94.24236355

Robust

| Robust | exp(b) | Std. Err. | z | P>|z| | [95% Conf. Interval] |
|--------|--------|-----------|---|------|------------------|
|        | died   |           |   |      |                  |
| low    | 1.553274 | .3193155 | 2.14 | 0.032 | 1.038154 2.323992 |
| stage2 | 2.524622 | 1.005045 | 2.33 | 0.020 | 1.157006 5.50803 |
| stage3 | 5.86819 | 2.312817 | 4.49 | 0.000 | 2.710329 12.70534 |

6 Method 4: Poisson regression with a robust variance estimator

Poisson regression is a generalized linear model with a log link and a Poisson distribution. When the outcome is binary, the exponentiated coefficients are risk ratios instead of incidence-rate ratios (Gourieroux, Monfort, and Trognon 1984a,b; Lloyd 1999, 85–86; Wooldridge 2002, 648–649; Greenland 2004a; Zou 2004; Carter, Lipsitz, and Tilley 2005). Methods that rely on the Poisson distribution assume that the mean count and its variance are equal. In the breast cancer data, the mean count of deaths per woman was 54/192 = 0.28125. If the variance of the mean count is also 0.28125, then the standard error of the mean count is the square root of the variance divided by the square root of the number of women (192) = 0.0382733. This is indeed the standard error that Stata reports using the ci, poisson command or using lincom after the poisson command:

. ci died, poisson

<table>
<thead>
<tr>
<th>Variable</th>
<th>Exposure</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>died</td>
<td>192</td>
<td>.28125</td>
<td>.0382733</td>
<td>.2112837 .3669702</td>
</tr>
</tbody>
</table>
Poisson regression Number of obs = 192
LR chi2(0) = 0.00
Prob > chi2 = .
Log likelihood = -122.49961 Pseudo R2 = 0.0000

| died   | Coef.  | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|------|------|----------------------|
| _cons | -1.268511 | .1360828  | -9.32| 0.000 | -1.535229 -1.001794 |

. lincom _cons, irr
   ( 1) [died]_cons = 0

| died   | IRR    | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|--------|--------|-----------|------|------|----------------------|
| (1)    | .28125 | .0382733  | -9.32| 0.000 | .2154064 .3672201   |

If the deaths were from a Poisson distribution, women would have nonnegative integer counts of 0, 1, 2, 3, ..., or more deaths. The data cannot be Poisson, because no woman dies more than once. The data are from a binomial distribution, and the binomial variance is assumed to be the proportion that died multiplied by 1 minus that proportion. The standard error of the mean proportion is the square root of the variance divided by the square root of the number of women, which is 0.0324477. This is the standard error reported by `ci`, `binomial`:

. ci died, binomial

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>— Binomial Exact —</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>died</td>
<td>192</td>
<td>.28125</td>
<td>.0324477</td>
<td>.2188833</td>
<td>.3505085</td>
</tr>
</tbody>
</table>

If we use Poisson methods for these binomial data, the standard error for the outcome proportion (risk) is too large: 0.03827 instead of 0.03245. As the outcome becomes less common, the Poisson standard error will converge toward the binomial standard error (Armitage, Berry, and Matthews 2002, 71–76). But in the breast cancer data, use of Poisson regression to estimate risk ratios will produce standard errors, p-values, and CIs that are too large:

. poisson died low stage2 stage3, irr nolog

| died | IRR    | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|------|--------|-----------|------|------|----------------------|
| low  | 1.630775 | .4688634  | 1.70 | 0.089 | .9282513 2.864987    |
| stage2 | 2.520742 | 1.074375  | 2.17 | 0.030 | 1.093288 5.811955    |
| stage3 | 5.913372 | 2.645148  | 3.97 | 0.000 | 2.460814 14.20992    |
Above, the 95% CI for the low variable is wide, 0.93 to 2.86, compared with the CIs from other methods. We can obtain standard errors and CIs that are approximately correct by using a robust variance estimator, which can relax the assumption that the data are from a Poisson distribution (Wooldridge 2002, 650–651; Greenland 2004a; Zou 2004; Carter, Lipsitz, and Tilley 2005). The robust variance estimator is sometimes called the Huber, White, Huber–White, sandwich, or survey estimator, as well as other names (Hardin and Hilbe 2007, 35–36). In Stata, we can invoke this estimator with the vce(robust) option and the CI for the low variable becomes narrower, 1.07 to 2.48:

```
. poisson died low stage2 stage3, irr nolog vce(robust)
Poisson regression
Number of obs = 192
Wald ch2(3) = 53.61
Prob > ch2 = 0.0000
Log pseudolikelihood = -109.14601 Pseudo R2 = 0.1090

<table>
<thead>
<tr>
<th></th>
<th>Robust</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
<td>z</td>
</tr>
<tr>
<td>died</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>1.630775</td>
<td>.3480542</td>
<td>2.29</td>
<td>0.022</td>
<td>1.073305 2.477792</td>
</tr>
<tr>
<td>stage2</td>
<td>2.520742</td>
<td>.9937819</td>
<td>2.35</td>
<td>0.019</td>
<td>1.16399 5.458932</td>
</tr>
<tr>
<td>stage3</td>
<td>5.913372</td>
<td>2.28568</td>
<td>4.60</td>
<td>0.000</td>
<td>2.772187 12.61386</td>
</tr>
</tbody>
</table>
```

7 Method 5: Cox proportional hazards regression with a robust variance estimator

Cox proportional hazards regression can estimate risk ratios if we set the follow-up time to 1, or any quantity that is the same for all subjects, and use the Breslow method to break ties. The robust variance estimator should be used, because otherwise the standard errors will be too large:

```
. generate byte time = 1
. stset time, failure(died) noshow
  failure event: died != 0 & died < .
  obs. time interval: (0, time]
  exit on or before: failure
  192 total obs.
  0 exclusions
  192 obs. remaining, representing
  54 failures in single record/single failure data
  192 total analysis time at risk, at risk from t = 0
     earliest observed entry t = 0
     last observed exit t = 1
```
The results above reproduce exactly the results from Poisson regression with the robust variance estimator; the Poisson and Cox methods are identical when implemented in this way. Options other than the Breslow method for dealing with tied survival times will produce risk-ratio estimates for exposure to a low estrogen-receptor–level tumor that are too large: 1) the \texttt{efron} option produces a risk ratio of 1.91, 2) the \texttt{exactm} option yields 2.04, and 3) the \texttt{exactp} option risk ratio is 2.49.

8 Method 6: Regression-based standardized risk ratios

Any regression model for binomial outcomes can estimate the probability (risk) of death for women with high and low estrogen-receptor–level tumors within each cancer stage. With this information, we can estimate the average risk of death that would be expected if all 192 women had low estrogen-receptor–level tumors and the distribution of cancer stages observed in the data. This estimate is said to be standardized to the distribution of the other variables, cancer stage in this example, in the regression model (Lane and Nelder 1982; Flanders and Rhodes 1987; Joffe and Greenland 1995; Greenland 1991a, 2004a; Localio, Margolis, and Berlin 2007; Rothman, Greenland, and Lash 2008, 442–446). The average risk can also be estimated assuming that all 192 women had a high estrogen-receptor–level tumor. The ratio of the low estrogen-receptor–level average risk divided by the high estrogen-receptor–level average risk can be calculated and the standard error for this risk ratio can be estimated using the delta method (Casella and Berger 2002, 240–245). The word “standardized” is used just as for standardized mortality rates or any statistic standardized to a given population distribution. We can estimate this standardized risk ratio by using logistic regression:

(Continued on next page)
Methods for estimating adjusted risk ratios

The adjusted odds ratio for death among women with a low estrogen-receptor level–tumor, compared with women with a high estrogen-receptor–level tumor, was 2.5. Because the outcome of death was common, this odds ratio does not closely approximate the risk ratio.

Above I used `predictnl` to estimate the ln of the risk ratio (\texttt{lnrr} variable); this command can estimate nonlinear comparisons from regression coefficients. The `se` option estimated the standard error for the ln risk ratio using the delta method. To make the output less cluttered, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line. To estimate the risk or probability of death, I used `delimit` to change how Stata recognizes the end of a command line.
in the expression, after the line with only a division sign, is the sum of all estimated risks if all women had a high estrogen-receptor–level tumor, again standardized to the observed cancer-stage distribution. The first sum is divided by the second and the ln taken of this ratio so that the ln of the risk ratio is estimated. I then estimated the risk ratio, which is exp(lnrr), and the 95% upper and lower confidence limits for the risk ratio, and used the display command to show these results for the last record by using the subscript \[N\]: risk ratio = 1.7, 95% CI is [1.1, 2.6].

We can use simpler commands to estimate the risk ratio, but they do not provide a CI. Still, these commands show how the risks and risk ratio may be estimated and are shown below to clarify how Stata is using the regression estimates. After fitting the logistic model, the commands are

```
. replace low=0
   (48 real changes made)
. predict risk0
   (option pr assumed; Pr(died))
. summ risk0, meanonly
   . local avrisk0 = r(mean)
. replace low=1
   (192 real changes made)
. predict risk1
   (option pr assumed; Pr(died))
. summ risk1, meanonly
   . local avrisk1 = r(mean)
   . local rr = `avrisk1´/`avrisk0´
   . display "Risk1 = " `avrisk1´ " Risk0 = " `avrisk0´ " Risk ratio = " `rr´
   Risk1 = .40087948 Risk0 = .23924549 Risk ratio = 1.6755989
```

Risks for binomial outcomes can also be estimated after probit regression. In the probit model, the outcome risk estimate applies the cumulative standard normal distribution function (normal()) to the linear predictor, instead of the ln odds function used in logistic regression:

```
. use brcadat, clear
   (Breast cancer data)
. probit died low stage2 stage3, nolog

   Probit regression
   Number of obs = 192
   LR chi2(3) = 42.21
   Prob > chi2 = 0.0000
   Log likelihood = -92.968357 Pseudo R2 = 0.1850

                     died
                      Coef. Std. Err.      z    P>|z|     [95% Conf. Interval]
  --------------------- ------------- ------------- ------- ------ ------------------
     low               .5386148    .2343501     2.30   0.022       0.079297    .9979327
    stage2             .6290485    .2503224     2.51   0.012      0.1384256    1.119671
    stage3             1.739085    .3302966     5.27   0.000      1.091715    2.386454
     _cons            -1.376363    .2165793    -6.36   0.000     -1.800851    -.9518752
```
Methods for estimating adjusted risk ratios

Nelder (2001) has suggested that when a dichotomous outcome is common, the complementary log-log regression model may fit the data well:

```
.delimit ;
delimiter now;
.predictnl lnrr = ln(
  sum(normal(_b[_cons]+_b[stage2]*stage2+_b[stage3]*stage3+_b[low]))
  / 
  sum(normal(_b[_cons]+_b[stage2]*stage2+_b[stage3]*stage3)))
  , se(lnrr_se);
.delimit cr
delimiter now cr
.scalar rr = exp(lnrr[_N])
.scalar upper = exp(lnrr[_N] + invnormal(1-.05/2)*lnrr_se[_N])
.scalar lower = exp(lnrr[_N] - invnormal(1-.05/2)*lnrr_se[_N])
.display "Risk ratio = " rr " 95% CI = " lower " , " upper
Risk ratio = 1.6751332 95% CI = 1.0913484, 2.5711965
```

Hardin and Hilbe (2007, 147) note that if most subjects either have or do not have the outcome, the complementary log-log and log-log models may fit better than logistic or probit models. We can fit the log-log model using the `glm` command:
P. Cummings

. use breast.dat, clear
(Breast cancer data)
. glm died low stage2 stage3, nolog family(bin) link(loglog) eform

Generalized linear models
No. of obs = 192
Optimization : ML
Residual df = 188
Scale parameter = 1

Deviance = 186.5949538 (1/df) Deviance = .9925263
Pearson = 192.8460376 (1/df) Pearson = 1.025777

Variance function: V(u) = u*(1-u) [Bernoulli]
Link function : g(u) = -ln(-ln(u)) [Log-log]

AIC = 1.013515
Log likelihood = -93.2974769 BIC = -801.8142

| died   | exp(b)       | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|--------|--------------|-----------|-------|--------|-----------------------|
| low    | 1.642276     | .3938706  | 2.07  | 0.039  | 1.023682, 2.627796    |
| stage2 | 1.695259     | .3494023  | 2.56  | 0.010  | 1.131876, 2.539063    |
| stage3 | 6.133979     | 2.415435  | 4.61  | 0.000  | 2.835022, 13.27174    |

. #delimit ;
delimiter now ;
. predictnl lnrr = ln(
    > sum(exp(-exp(-(_b[_cons]+_b[stage2]*stage2+_b[stage3]*stage3+_b[low])))))
    > / 
    > sum(exp(-exp(-(_b[_cons]+_b[stage2]*stage2+_b[stage3]*stage3)))))
    > , se(lnrr_se);
. #delimit cr
delimiter now cr
. scalar rr = exp(lnrr[_N])
. scalar upper = exp(lnrr [_N] + invnormal(1-.05/2)*lnrr_se[_N])
. scalar lower = exp(lnrr [_N] - invnormal(1-.05/2)*lnrr_se[_N])
. display "Risk ratio = " rr " 95% CI = " lower " , " upper
Risk ratio = 1.6312705 95% CI = 1.0545183, 2.5234682

9  Method 7: A substitution method

A crude (unadjusted) odds ratio can be converted to a risk ratio: crude risk ratio = (crude odds ratio) / \{1 – Po + (Po × crude odds ratio)\}, where Po is the proportion of all unexposed subjects who had the outcome in the data. Zhang and Yu (1998) suggested that, in a cohort study, one can use this same formula to convert an adjusted odds ratio to an adjusted risk ratio. This substitution method was described by Holland (1989), who used it to estimate an adjusted risk difference from a Mantel–Haenszel summary odds ratio. Greenland and Holland (1991) reported that this will produce ratio estimates biased away from 1 when outcomes are common and risk among those not exposed varies substantially. The bias occurs because the summary odds ratio is not a weighted average of stratum-specific odds ratios; odds ratios lack the property of collapsibility (Greenland 1987). The method can be implemented in Stata:
Methods for estimating adjusted risk ratios

Methods for estimating adjusted risk ratios

. use brcadat, clear
(Breast cancer data)
. summ died if low==0, meanonly
. local p0 = r(mean)
. display `p0´
.21527778
. logistic died low stage2 stage3, nolog

Logistic regression Number of obs = 192
LR chi2(3) = 42.27
Prob > chi2 = 0.0000
Log likelihood = -92.939847 Pseudo R2 = 0.1853

| died    | Odds Ratio | Std. Err. | z  | P>|z| | [95% Conf. Interval] |
|---------|------------|-----------|----|------|----------------------|
| low     | 2.508065   | .9916923  | 2.33| 0.020 | 1.155507             | 5.443836 |
| stage2  | 3.109772   | 1.44851   | 2.44| 0.015 | 1.248087             | 7.748406 |
| stage3  | 18.8389    | 11.03231  | 5.01| 0.000 | 5.978344             | 59.36498 |

. scalar rr = exp(_b[low])/
[1-(`p0´)*exp(_b[low])]
. scalar lower = exp(_b[low]-invnormal(1-.05/2)*_se[low])/
[1-(`p0´)*exp(_b[low]-invnormal(1-.05/2)*_se[low])]
. scalar upper = exp(_b[low]+invnormal(1-.05/2)*_se[low])/
[1-(`p0´)*exp(_b[low]+invnormal(1-.05/2)*_se[low])]
. display "Risk ratio = " rr " 95% CI = " lower ", upper
Risk ratio = 2.508065 95% CI = 1.155507, 5.443836

10 Bootstrap CIs

In some examples above, approximately correct CIs were obtained using robust or delta methods. Bootstrap methods can also be used for CIs. Here are commands to estimate bootstrap CIs for the risk ratio by using a logistic model:

. use brcadat, clear
(Breast cancer data)
. program stlogit, rclass
 1. version 10
 2. logistic died low stage2 stage3, nolog
 3. preserve
 4. replace low=0
 5. predict risk0
 6. summ risk0, meanonly
 7. scalar avrisk0 = r(mean)
 8. replace low=1
 9. predict risk1
10. summ risk1, meanonly
11. scalar avrisk1 = r(mean)
12. return scalar lnrr = ln(avarisk1/avarisk0)
13. restore
14. end
. set seed 93514
P. Cummings

.stat bootstrap lnrr=r(lnrr), saving(bsanrr3a, replace) reps(400001) nowarn nodots:
> stlogit
(output omitted)
. estat bootstrap, all eform
Bootstrap results Number of obs = 192
command: stlogit
Replications = 400001

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Bootstrap</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>exp(b)</td>
<td>Bias Std. Err.</td>
<td></td>
</tr>
<tr>
<td>lnrr</td>
<td>1.6755989</td>
<td>-.0049751 .37575458</td>
<td>1.079661 2.600476 (N)</td>
</tr>
<tr>
<td></td>
<td>1.068839</td>
<td>2.587837 (P)</td>
<td>1.071832 2.594932 (BC)</td>
</tr>
</tbody>
</table>

(N) normal confidence interval
(P) percentile confidence interval
(BC) bias-corrected confidence interval

Stata’s bootstrap command simplifies the task of estimating bootstrap CIs by using four methods: 1) normal, 2) percentile, 3) bias corrected, and 4) bias corrected and accelerated. Other methods are available (Carpenter and Bithell 2000). For the risk ratios estimated in this article, the choice among Stata’s four methods makes little difference. But for some epidemiologic data, the normal and percentile methods should be used with caution because they may have substantial coverage error (Efron and Tibshirani 1993; Carpenter and Bithell 2000; Greenland 2004b).

11 Risk-ratio methods for matched data

Adjusted risk ratios for matched data can be estimated using conditional Poisson regression, which Stata implements in the xtpoisson, fe command. I have previously reviewed the analysis of matched cohort data in the Stata Journal (Cummings and McKnight 2004) and elsewhere (Cummings, McKnight, and Weiss 2003; Cummings, McKnight, and Greenland 2003).

12 Summary

When the risk-ratio estimates in this article are rounded to one decimal, nearly all the methods produced estimates of 1.6 or 1.7 (table 2). They also differed little with regard to estimated CIs: the 95% lower bound was 1.0 or 1.1 and the upper bound, 2.3 to 2.6.

The risk ratio that stands out as different came from the substitution method: risk ratio = 1.9 and 95% CI is [1.1, 3.1]. The substitution method has nothing to recommend it; it will usually produce estimates biased away from 1 when outcomes are common, and in Stata it offers little advantage in terms of simplicity. Stata users who wish to estimate an adjusted risk ratio have better methods that they can use, all of which are fairly easy to implement.
Methods for estimating adjusted risk ratios

Table 2. Risk-ratio estimates for death within 5 years among 192 women with breast cancer, comparing women with low estrogen-receptor–level tumors with women with high estrogen-receptor–level tumors, adjusted for cancer stage at diagnosis. Results are shown using the methods described in this article.

<table>
<thead>
<tr>
<th>Method</th>
<th>Risk ratio</th>
<th>95% CI</th>
<th>Bootstrap 95% CI†</th>
<th>Akaike information criteria‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Stratified methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mantel–Haenszel</td>
<td>1.62</td>
<td>1.09, 2.39</td>
<td>1.07, 2.48</td>
<td>.</td>
</tr>
<tr>
<td>Inverse-variance weights</td>
<td>1.55</td>
<td>1.08, 2.25</td>
<td>. . . . . .</td>
<td></td>
</tr>
<tr>
<td>2. Generalized linear regression with log link and binomial distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum likelihood</td>
<td>1.56</td>
<td>1.05, 2.32</td>
<td>. . . 193.9</td>
<td></td>
</tr>
<tr>
<td>Wacholder’s truncated method</td>
<td>1.56</td>
<td>1.06, 2.29</td>
<td>1.03, 2.44</td>
<td>. . . . . .</td>
</tr>
<tr>
<td>Copy method</td>
<td>1.56</td>
<td>1.05, 2.31</td>
<td>. . . . . .</td>
<td></td>
</tr>
<tr>
<td>3. Generalized linear regression with log link, Gaussian distribution, robust variance estimator</td>
<td>1.55</td>
<td>1.04, 2.32</td>
<td>1.04, 2.43</td>
<td>196.5</td>
</tr>
<tr>
<td>4. Poisson regression with robust variance estimator</td>
<td>1.63</td>
<td>1.07, 2.48</td>
<td>1.06, 2.55</td>
<td>226.3</td>
</tr>
<tr>
<td>5. Cox proportional hazards with robust variance estimator</td>
<td>1.63</td>
<td>1.07, 2.48</td>
<td>1.06, 2.55</td>
<td>547.1</td>
</tr>
<tr>
<td>6. Regression-based standardized risk ratios</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic</td>
<td>1.68</td>
<td>1.09, 2.57</td>
<td>1.07, 2.59</td>
<td>193.9</td>
</tr>
<tr>
<td>Probit</td>
<td>1.68</td>
<td>1.09, 2.57</td>
<td>1.07, 2.59</td>
<td>193.9</td>
</tr>
<tr>
<td>Complementary log-log</td>
<td>1.67</td>
<td>1.10, 2.52</td>
<td>1.08, 2.56</td>
<td>193.5</td>
</tr>
<tr>
<td>Log-log</td>
<td>1.63</td>
<td>1.05, 2.52</td>
<td>1.03, 2.54</td>
<td>194.6</td>
</tr>
<tr>
<td>7. Substitution method</td>
<td>1.89</td>
<td>1.12, 2.78</td>
<td>1.08, 3.11</td>
<td>. . . . . .</td>
</tr>
</tbody>
</table>

† Bias-corrected bootstrap CIs based upon 400001 replications. Not estimated for the inverse-variance stratified method because the `cs` command does not return the pooled risk ratio from this method. Not estimated for the maximum-likelihood version of the generalized linear model with a log link and binomial distribution because convergence failed in many bootstrap samples. Convergence also failed in 100 bootstrap samples (0.025%) using Wacholder’s truncated method (`binreg`), 10 samples (0.0025%) using the generalized linear model with a log link and Gaussian distribution, and 5 samples (0.00125%) using the complementary log-log method.

‡ Akaike information criteria statistic for models fit using maximum likelihood. The statistic compares the fitted model with a model that has only the outcome variable. In Stata, smaller Akaike information criteria statistics indicate better fit.
13 Acknowledgment

This work was supported by grant R49/CE000197-04 from the Centers for Disease Control and Prevention, Atlanta, GA.

14 References


Methods for estimating adjusted risk ratios


**About the author**

Peter Cummings is a professor in the Department of Epidemiology in the School of Public Health and a core faculty member at the Harborview Injury Research and Prevention Research Center, University of Washington, Seattle, WA.