Title

> **reri** — Relative excess risk due to interaction

Description

reri reports three statistics that assess two-way interactions in an additive model of relative risk. The statistics are of interest for determining how two risk factors interact to produce a positive result for the outcome of interest. The estimates of the interactions themselves are reported as excess relative risks (ERRs). Statistics reported for the interaction are the relative excess risk due to interaction (RERI), attributable proportion (AP), and synergy index (SI); reri supports binomial generalized linear, Poisson, negative binomial, logistic, Cox, parametric survival, and interval-censored parametric and semiparametric survival models.

Quick start

Report ERRs, RERI, AP, and SI for the interaction of exposures exp1 and exp2 from a binomial generalized linear model for y fit by binreg

    reri binreg y exp1 exp2

Same as above, but use poisson to fit the model

    reri poisson y exp1 exp2

Same as above, but report incidence-rate ratios for the interactions rather than ERRs

    reri poisson y exp1 exp2, irr

Same as above, but include the covariate x, and specify noisily to view the full Poisson model

    reri poisson y exp1 exp2 x, irr noisily

Same as above, but report ERRs for the interactions as well as incidence-rate ratios for the full model

    reri poisson y exp1 exp2 x, irr err noisily

Fit a Cox model using stset data stratified by svar, and report ERRs, RERI, AP, and SI for the interactions of exp1 and exp2

    reri stcox exp1 exp2, strata(svar)

Same as above, but fit a Weibull model

    reri streg exp1 exp2, strata(svar) distribution(weibull)

Fit a Weibull model with interval-censored survival data, with variables t1 and t2 giving lower and upper endpoints for the censoring interval, and report ERRs, RERI, AP, and SI for the interactions

    reri stintreg exp1 exp2, distribution(weibull) interval(t1 t2)

Menu

Statistics > Epidemiology and related > Other > Relative excess risk due to interaction (RERI)

Syntax

> reri *cmd* *depvar* *var1* *var2* [*control_vars*] [*if*] [*in*] [*weight*]
>
> [, *options* *cmd_options*]

cmd defines the estimation command to be executed and may be one of the following:

> binreg, logistic, nbreg, poisson, stcox, stintcox, stintreg, and streg

depvar is required for all except stcox, stintcox, stintreg, and streg, which do not allow
depvar.

var1 and *var2* are binary variables (integer valued and nonnegative) whose interactions are to be
tested for additivity.

options	Description
Model	
baselevel1(*#*)	specify base level for *var1*
baselevel2(*#*)	specify base level for *var2*
Reporting	
noisily	display output from *cmd*
err	report interactions as ERR, the default
coef_type	report interactions as *coef_type* rather than ERR
level(*#*)	set confidence level; default is level(95)
nofvlabel	display numeric values rather than value labels
nolegend	do not display the legend for the interactions

var1 and *var2* cannot be factor variables, but *control_vars* may contain factor variables; see [U] **11.4.3 Factor variables**.

cmd_options are any options allowed by *cmd*, except options or, hr, rd, and vce(jackknife1) are not allowed
when *cmd* is binreg.

When *cmd* is poisson, vce(robust) is used by default when *depvar* is 0/1.

bootstrap, by, collect, jackknife, and statsby are allowed; see [U] **11.1.10 Prefix commands**.

Weights are allowed if *cmd* allows them; see [U] **11.1.6 weight**.

All postestimation commands behave as they would after *cmd*; see the postestimation manual entry for *cmd*.

Options

> [Model]

baselevel1(*#*) and baselevel2(*#*) specify base levels for *var1* and *var2*, respectively, the two
variables that give the interactions of interest.

If *var1* and *var2* are 0/1 variables, then by default, the RERI statistic is a test of whether the ERR of
interaction 01 (01 meaning *var1* $= 0$ and *var2* $= 1$) plus the ERR of interaction 10 equals the ERR
of interaction 11. In shorthand notation, we say we are testing $01 + 10 = 11$ on the ERR scale.

Specifying baselevel1(*#*) or baselevel2(*#*) allows the base levels of the interactions to be
changed. For example, with 0/1 variables, baselevel1(1) sets the base level for *var1* to 1, and
RERI tests whether $11 + 00 = 01$.

If baselevel1(*#*) is not specified, the base level of *var1* is the default base level of the factor
variable i.*var1*. The default base level is the smallest of the two values of *var1*, unless the

default has been changed by fvset. This is similarly true for baselevel2(#) and *var2*. See [U] **11.4.3.2 Base levels** and [R] **fvset**. Note that base levels cannot be set using factor-variable notation when specifying *var1* and *var2*.

───
Reporting
───

noisily displays all output from *cmd*.

err specifies that the coefficients of the interactions be displayed as ERRs. This is the default. ERR is equal to relative risk minus one. See *Remarks and examples* below.

coef_type specifies an alternative scale for the coefficients of the interactions instead of ERRs. Allowed *coef_type* depends on *cmd* and can be rr, or, irr, or hr, whichever is allowed by *cmd*. In addition, *coef_type* can be coef, which displays the interaction terms as unscaled coefficients.

For binreg, *coef_type* rr displays the interactions as relative-risk ratios. The other models fit by binreg using options or, hr, and rd are not available with reri. For logistic, *coef_type* or displays the interactions as odds ratios. For poisson and nbreg, *coef_type* irr displays the interactions as incidence-rate ratios. For stcox, stintcox, stintreg, and streg, *coef_type* hr displays the interactions as hazard ratios. For the survival-time commands, nohr may be used as a synonym for coef.

When noisily is specified, err and a *coef_type* may be specified simultaneously. In this case, the table of interactions displayed by reri reports ERRs, and the coefficient table produced by *cmd* displays coefficients on the *coef_type* scale.

level(#) specifies the confidence level, as a percentage, for confidence intervals. The default is level(95) or as set by set level; see [R] **level**.

nofvlabel specifies that numeric values rather than value labels be displayed in the legend and *cmd* output.

nolegend suppresses the display of the legend for the interactions.

Remarks and examples

stata.com

Remarks are presented under the following headings:

> *Introduction*
> *Additive versus multiplicative interactions*
> *Incidence-rate ratios, hazard ratios, and odds ratios*

Introduction

reri reports three statistics that assess additive interactions of relative risk: RERI, AP, and SI (Lash et al. 2021). These measures are typically of interest when working with epidemiologic data. When one assesses the effect of multiple exposures on an outcome, determining whether one exposure modifies the effect of another can be an important piece of the puzzle.

When there are two exposures that affect an outcome, we often want to model how the exposures interact. That is, for binary exposures A and B, we want to model the risk of a positive outcome for subjects having both A and B compared with subjects having only A and those having only B.

Typically, a multiplicative model for the interaction is fit to the data: the risk for those with A and B is hypothesized to equal the risk for only A times the risk for only B. If the risk for A and B equals the multiplicative risk, then we say there is no interaction of A and B. If the risk is greater than the multiplicative risk, we say there is a positive interaction of A and B. Multiplicative models

are computationally convenient because they are simple to specify, and it is easy to test whether the interaction term is greater (or less) than multiplicative.

Inherent to this model is the hypothesis that the risk is multiplicative. For many biological processes, however, an additive model of risk may model the process better than a multiplicative one (see, for example, Andersson et al. [2005]). That is, the risk for subjects with A and B is hypothesized to equal the risk for only A plus the risk for only B. For this hypothesis, we want to look at how much the risk for A and B is greater (or less) than the additive risk.

The RERI statistic is this risk difference formulated using relative risks. A RERI statistic that is 0 means the risk is additive. One that is positive means the risk is superadditive, and one that is negative means the risk is subadditive. The test of statistical significance of the RERI statistic is straightforward, but it is not as simple as the test for the multiplicative model (see *Methods and formulas*).

Besides the RERI statistic, there are two other related statistics for additive models: AP and SI. The AP is the proportion of risk of A and B due to the superadditivity of exposures A and B. The SI recasts the RERI statistic as a ratio, rather than a difference.

Let's formally define these statistics. Let p_{A+B+} be the probability of having a positive outcome for subjects positive for binary exposure A and positive for binary exposure B. Define p_{A-B+}, p_{A+B-}, and p_{A-B-} similarly. An additive model for the probabilities is

$$p_{A+B+} = p_{A-B-} + (p_{A-B+} - p_{A-B-}) + (p_{A+B-} - p_{A-B-})$$

Said in words, the probability of a positive outcome for $A+B+$ equals the probability for the base category $A-B-$ plus the amounts that probabilities for $A-B+$ and $A+B-$ exceed the probability for the base category.

If we divide this equation by p_{A-B-} and note that $p_{A+B+}/p_{A-B-} = \mathrm{RR}_{A+B+}$, the relative risk for $A+B+$ (and similarly for RR_{A-B+} and RR_{A+B-}), we get

$$\mathrm{RR}_{A+B+} = \mathrm{RR}_{A-B+} + \mathrm{RR}_{A+B-} - 1$$

ERR is defined as $\mathrm{ERR} = \mathrm{RR} - 1$. In terms of ERR, the additive model becomes

$$\mathrm{ERR}_{A+B+} = \mathrm{ERR}_{A-B+} + \mathrm{ERR}_{A+B-}$$

The RERI statistic is defined as

$$\mathrm{RERI} = \mathrm{ERR}_{A+B+} - \mathrm{ERR}_{A-B+} - \mathrm{ERR}_{A+B-}$$

Hence, RERI is 0 for a perfectly additive model. If it is negative, the effect of the $A+B+$ interaction is less than additive (subadditive). If positive, the effect is more than additive (superadditive).

AP is simply RERI scaled by the relative risk (not ERR) of $A+B+$. That is,

$$\mathrm{AP} = \frac{\mathrm{RERI}}{\mathrm{RR}_{A+B+}}$$

When AP is positive, it represents the proportion of risk of $A+B+$ due to the superadditivity of exposures A and B (that is, the amount greater than additivity). Note that it is a proportion of risk, not relative risk, as one can see by multiplying both numerator and denominator by the reference probability p_{A-B-}.

The SI is a ratio measure:

$$\mathrm{SI} = \frac{\mathrm{ERR}_{A+B+}}{\mathrm{ERR}_{A-B+} + \mathrm{ERR}_{A+B-}}$$

The SI is 1, rather than 0, for a perfectly additive model. It is <1 for a subadditive model and >1 for a superadditive model. Note that because ERRs can be negative (they are always ≥ -1), the SI can be negative, making interpretation difficult. The SI is intended as a measure for cases in which all ERRs of the model are positive, but this will not always be the case. ERRs will be determined by the data and the specification of the model. See Knol et al. (2011) and example 4 for discussions about negative ERRs.

Additive versus multiplicative interactions

Typically, interactions are fit and tested by specifying main effects and interactions of the main effects (Rothman, Greenland, and Walker 1980). Suppose x_{A+} is a 0/1 variable that is 1 when exposure A is positive and 0 when it is negative. The variable x_{B+} is defined similarly for exposure B. Let x_{A+B+} be the 0/1 variable that is 1 when both A and B are positive and 0 otherwise.

When we specify a model and include terms

$$b_1 x_{A+} + b_2 x_{B+} + b_3 x_{A+B+}$$

we call $b_1 x_{A+}$ and $b_2 x_{B+}$ main effects and $b_3 x_{A+B+}$ their interaction. The test of the interaction is the test of $b_3 = 0$.

It is a model of interaction that is additive in these terms. However, when the statistics of interest are risk ratios, odds ratios, or hazard ratios, the model is generally called multiplicative because it is multiplicative in these statistics. The risk ratios (or odds ratios or hazard ratios) are given by the exponentiated coefficients; that is, $\mathrm{RR}_{A+} = e^{b_1}$, $\mathrm{RR}_{B+} = e^{b_2}$, etc. The risk ratio for $A+B+$ is $\mathrm{RR}_{A+B+} = e^{b_1} e^{b_2} e^{b_3}$. If the interaction is multiplicative, $e^{b_3} = 1$ and

$$\mathrm{RR}_{A+B+} = \mathrm{RR}_{A+} \times \mathrm{RR}_{B+}$$

As mentioned earlier, many biological processes more closely follow an additive model of risk than a multiplicative one. The additive model of ERR is

$$\mathrm{ERR}_{A+B+} = \mathrm{ERR}_{A-B+} + \mathrm{ERR}_{A+B-}$$

Note that unlike the multiplicative model with main effects, the three states in the additive model, $A-B+$, $A+B-$, and $A+B+$, are disjoint.

reri is called with the syntax

 reri *cmd depvar var1 var2* ...

reri uses *var1* and *var2* to create the three terms for the additive model using factor-variable notation. Suppose *var1* is xa, a 0/1 variable, indicating whether a subject has exposure A. Suppose *var2* is xb for exposure B. Then, the terms needed to test for additive ERR are, in factor-variable notation,

 `0.xa#1.xb 1.xa#0.xb 1.xa#1.xb`

reri creates these terms and calls *cmd* to fit the model. Then, it uses nlcom to calculate the RERI, AP, and SI statistics and their standard errors. See *Methods and formulas*.

Incidence-rate ratios, hazard ratios, and odds ratios

Incidence rates and hazards are measures of risk. An incidence rate is a probability adjusted for time at risk. A hazard is a conditional probability when the distribution is discrete and an incremental conditional probability when the distribution is continuous (Kalbfleisch and Prentice 2002). Hence, incidence-rate ratios and hazard ratios are measures of relative risk. So using RERI, AP, and SI statistics with models that estimate incidence-rate ratios (poisson and nbreg) or hazard ratios (stcox, stintcox, streg, stintreg) presents no difficulties.

Odds ratios, however, are different. As is well known, the odds ratio approximates the risk ratio in the population when the prevalence of the outcome in the population is low (see, for example, Greenland and Thomas [1982]). So if your data are from a case–control study and the population prevalence of a positive outcome is low, then odds ratios from a logistic regression will approximate the risk ratios in the population, and the RERI, AP, and SI statistics are appropriate. On the other hand, if the population prevalence is high, the odds ratio is typically a poor approximation to the risk ratio (unless it is close to 1), and interpretation of the RERI, AP, and SI statistics may be problematic (see, for example, Skrondal [2003]).

If the study you are analyzing is a cohort study, observational study, or a randomized controlled trial with a binary outcome, you can estimate risk directly rather than fitting a logistic regression. A binomial model of risk or a Poisson model can be fit (using an appropriate variance estimator; see Cummings [2009] and example 2).

If time to outcome is measured, then a survival model may be fit, and there are no issues interpreting the RERI and SI statistics. Interpreting the AP statistic is a little trickier because it relates to the hazard at a point in time rather than risk in a population, so the proportion is the proportion of the hazard at a given time.

▷ Example 1: Binomial model of risk

We have simulated data on birth defects that mimic the results from Brender et al. (2013), an observational study of birth defects. Three binary outcomes are included in these data and are tube, palate, and heart, representing neural tube defects, palate defects, and heart defects, respectively. Predictors of birth defects are two measures of nitrate intake: drug, with values 0/1, indicating nitrosatable drug exposure; and nitrate, also with values 0/1, representing either low or high daily consumption of nitrates from drinking water.

We load the dataset and describe it.

```
. use https://www.stata-press.com/data/r18/nitrates
(Simulated prenatal nitrate intake and birth defects data)

. describe

Contains data from https://www.stata-press.com/data/r18/nitrates.dta
 Observations:           1,367                   Simulated prenatal nitrate
                                                   intake and birth defects data
    Variables:               6                   4 Oct 2022 16:05
                                                 (_dta has notes)

Variable      Storage   Display    Value
    name         type    format    label      Variable label

id              int      %9.0g                Identification number
tube            byte     %9.0g     case       Neural tube defect
heart           byte     %9.0g     case       Conotruncal heart defect
palate          byte     %9.0g     case       Cleft palate
drug            byte     %9.0g     yesno      Nitrosatable drug exposure
nitrate         byte     %9.0g     nitrate    Total nitrate intake

Sorted by: id
```

We first examine the outcome tube. We want to see whether the interactions of the two predictors drug and nitrate follow an additive model of risk. Because this is an observational study, a binomial model of risk is appropriate, and we can fit it using binreg.

The syntax of reri is reri *cmd depvar var1 var2*, where *var1* and *var2* are the binary predictors of interest. Note that *var1* and *var2* are specified without using factor-variable notation. reri will create the interactions automatically. Here's the result:

```
. reri binreg tube drug nitrate
Fitting binreg ...
Interaction of drug and nitrate on an additive scale
Model: Binomial generalized linear            Number of observations = 1,284
drug#nitrate
         - +      No#High
         + -      Yes#Low
         + +      Yes#High
```

	ERR	EIM std. err.	z	P>|z|	[95% conf. interval]	
drug#nitrate						
- +	.1161419	.1571024	0.78	0.435	-.1529507	.4707206
+ -	.2385481	.2392316	1.11	0.268	-.1517982	.8085336
+ +	.7769424	.407042	2.51	0.012	.1342	1.783922
RERI	.4222523	.453286	0.93	0.352	-.466172	1.310677
Attr. prop.	.2376286	.2187062	1.09	0.277	-.1910276	.6662848
Synergy ind.	2.190483	2.034328	0.84	0.398	.3548285	13.52263

```
Note: P>|z| for synergy index (SI) is for test H0: SI = 1.
```

Note that the interactions created by reri represent disjoint groups in the data. The -+ interaction is for drug = No and nitrate = High. The +- and ++ interactions are defined similarly as shown in the legend. The interaction -- is, of course, the reference category.

By default, reri displays the model coefficients for the -+, +-, and ++ interactions as ERR, which is relative risk minus one. So ERR greater than 0 means that relative risk is greater than 1. If the model is additive, we would have

$$\text{ERR}_{++} = \text{ERR}_{-+} + \text{ERR}_{+-}$$

The RERI statistic is the difference between the two sides of this equation:

$$\text{RERI} = \text{ERR}_{++} - \text{ERR}_{-+} - \text{ERR}_{+-}$$

The estimates are $\text{ERR}_{-+} = 0.116$ and $\text{ERR}_{+-} = 0.239$, so $\text{ERR}_{-+} + \text{ERR}_{+-} = 0.355$. The estimate of ERR_{++} is 0.777, which is greater than $\text{ERR}_{-+} + \text{ERR}_{+-}$. So the model is superadditive. The RERI statistic is 0.422, the difference between 0.777 and 0.355.

The p-value for RERI is a test of the null hypothesis $\text{RERI} = 0$. That is, a small p-value provides evidence to reject the additive model for the interactions. In this case, $p = 0.352$, so we do not have enough evidence to reject the additive model, despite the observed superadditivity.

AP is

$$\text{AP} = \frac{\text{RERI}}{\text{RR}_{++}}$$

It is the proportion of risk for the ++ interaction that is due to the risk that is above additive. In this case, AP is 0.238 or 23.8%. Its p-value is a test of the null hypothesis $\text{AP} = 0$. This null hypothesis is equivalent to the null hypothesis $\text{RERI} = 0$, so p-values will be similar.

The SI is

$$\text{SI} = \frac{\text{ERR}_{++}}{\text{ERR}_{-+} + \text{ERR}_{+-}}$$

It is how many times larger (or smaller) is $\text{ERR}_{++} = 0.777$ than $\text{ERR}_{-+} + \text{ERR}_{+-} = 0.355$. In this example, it is about twice as large, or more precisely, 2.19 times as large. The null hypothesis is $\text{SI} = 1$, rather than 0. The SI is a rather odd statistic because it is a test of an additive model using a multiplicative scale!

`reri` can optionally display the interaction coefficients from `binreg` as relative risk (that is, risk ratios) rather than ERR. This is done by specifying the `rr` option. We do so and also specify the `noisily` option to see the output from `binreg`.

```
. reri binreg tube drug nitrate, rr noisily
Fitting binreg:
Iteration 1:  Deviance =  1176.736
Iteration 2:  Deviance =  1176.736

Generalized linear models                    Number of obs    =       1,284
Optimization      : MQL Fisher scoring       Residual df      =       1,280
                    (IRLS EIM)               Scale parameter  =           1
Deviance        =  1176.735717              (1/df) Deviance  =    .9193248
Pearson         =  1283.999016              (1/df) Pearson   =    1.003124

Variance function: V(u) = u*(1-u/1)         [Binomial]
Link function    : g(u) = ln(u)             [Log]
                                             BIC              =   -7985.166
```

tube	Risk ratio	EIM std. err.	z	P>|z|	[95% conf. interval]	
drug#nitrate						
No#High	1.116142	.1571024	0.78	0.435	.8470493	1.470721
Yes#Low	1.238548	.2392316	1.11	0.268	.8482018	1.808534
Yes#High	1.776942	.407042	2.51	0.012	1.1342	2.783922
_cons	.157969	.013697	-21.28	0.000	.1332805	.1872307

Note: **_cons** estimates baseline risk.

```
Interaction of drug and nitrate on an additive scale

drug#nitrate
        - +     No#High
        + -     Yes#Low
        + +     Yes#High
```

	Risk ratio	EIM std. err.	z	P>|z|	[95% conf. interval]	
drug#nitrate						
- +	1.116142	.1571024	0.78	0.435	.8470493	1.470721
+ -	1.238548	.2392316	1.11	0.268	.8482018	1.808534
+ +	1.776942	.407042	2.51	0.012	1.1342	2.783922
RERI	.4222523	.453286	0.93	0.352	-.466172	1.310677
Attr. prop.	.2376286	.2187062	1.09	0.277	-.1910276	.6662848
Synergy ind.	2.190483	2.034328	0.84	0.398	.3548285	13.52263

```
Note: P>|z| for synergy index (SI) is for test H0: SI = 1.
```

We see the output of binreg followed by the output of reri. In this case, reri displays the interaction coefficients as risk ratios, matching the output of binreg. Risk ratios are just ERRs plus one, so point estimates and confidence intervals for the interactions are just shifted by one from the previous results, and standard errors and p-values are unchanged. The RERI, AP, and SI statistics are all exactly the same as they were in the previous results.

For comparison, we can fit a binreg model specifying the interaction in the standard manner: main effects for drug and nitrate and a single interaction term for drug = Yes and nitrate = High.

```
. binreg tube drug##nitrate, rr

Iteration 1:  Deviance =  1617.128
Iteration 2:  Deviance =  1213.623
Iteration 3:  Deviance =  1177.487
Iteration 4:  Deviance =  1176.736
Iteration 5:  Deviance =  1176.736
Iteration 6:  Deviance =  1176.736

Generalized linear models                      Number of obs    =      1,284
Optimization      : MQL Fisher scoring         Residual df      =      1,280
                    (IRLS EIM)                  Scale parameter  =          1
Deviance         =    1176.735717              (1/df) Deviance  =   .9193248
Pearson          =    1283.99895               (1/df) Pearson   =   1.003124

Variance function: V(u) = u*(1-u/1)            [Binomial]
Link function    : g(u) = ln(u)                [Log]
                                               BIC              =  -7985.166
```

tube	Risk ratio	EIM std. err.	z	P>|z|	[95% conf. interval]	
drug						
Yes	1.238548	.2392315	1.11	0.268	.8482019	1.808533
nitrate						
High	1.116142	.1571023	0.78	0.435	.8470494	1.47072
drug#nitrate						
Yes#High	1.285408	.3952681	0.82	0.414	.7035471	2.348491
_cons	.157969	.013697	-21.28	0.000	.1332805	.1872307

```
Note: _cons estimates baseline risk.
```

This model is not as easy to interpret as the previous ones and shows no evidence that `drug`, `nitrate`, or their interaction has any effect on neural tube defects. However, the previous model fit with disjoint interactions provided strong evidence that the relative risk of neural tube defects was greater for the `++` group when compared with the reference group; in that model, the risk ratio for the `++` group was greater than 1, and the *p*-value was 0.012.

Note that when running `binreg` directly, we must specify the `rr` model option. When `reri` calls `binreg`, it automatically fits an `rr` model. Specifying `rr` with `reri` is merely a display option; it reports relative risk rather than ERR.

◁

❑ Technical note

As mentioned earlier, when `reri` is called with, say,

> `reri` *cmd* `y x1 x2 ...`

it creates disjoint interactions and runs

> *cmd* `y 0.x1#1.x2 1.x1#0.x2 1.x1#1.x2 ...`

(assuming that `x1` and `x2` are 0/1 and that 0 is taken as the base category for both).

If `reri` gives an error message and you cannot figure out the reason why, run

> *cmd* `y 0.x1#1.x2 1.x1#0.x2 1.x1#1.x2 ...`

and see whether this gives any problems such as dropped terms or dropped observations.

❑

▷ Example 2: Poisson model

We continue with our previous example. We again fit a model for the outcome `tube`, representing neural tube birth defects. Our predictors, as in the previous example, are `drug` (nitrosatable drug exposure) and `nitrate` (drinking water nitrate intake). The only difference is that we specify `poisson` rather than `binreg`.

```
. reri poisson tube drug nitrate
Fitting poisson ...
Interaction of drug and nitrate on an additive scale
Model: Poisson                                    Number of observations = 1,284
drug#nitrate
       - +      No#High
       + -      Yes#Low
       + +      Yes#High
```

	ERR	Robust std. err.	z	P>|z|	[95% conf. interval]	
drug#nitrate						
- +	.1161419	.1571636	0.78	0.435	-.1530418	.4708787
+ -	.2385481	.2393248	1.11	0.268	-.1519233	.8088004
+ +	.7769424	.4072075	2.51	0.012	.133993	1.78443
RERI	.4222523	.4534688	0.93	0.352	-.4665302	1.311035
Attr. prop.	.2376286	.2187937	1.09	0.277	-.1911993	.6664564
Synergy ind.	2.190483	2.035131	0.84	0.399	.3545735	13.53235

```
Note: P>|z| for synergy index (SI) is for test H0: SI = 1.
```

We see that the reported point estimates for the interactions, RERI, AP, and SI are identical to the ones estimated when we used `binreg`. The standard errors are almost the same, differing only in the fourth significant digit.

When the outcome is 0/1, `reri poisson` by default reports standard errors calculated using `vce(robust)`. This is because when the outcome is 0/1, the distribution is binomial (Bernoulli), not Poisson, and `vce(robust)` gives standard errors that are valid in this case. The other *vcetype*s allowed with `poisson` are available and can be selected by specifying the `vce()` option with `reri`. See [R] **poisson** and [R] *vce_option*.

`binreg` by default fits models using iterated, reweighted least-squares optimization and uses the expected information matrix for the variance estimator. `poisson` uses maximum likelihood optimization. Point estimates will theoretically be the same but will have numerical differences because of the different optimizers used. Different variance estimators are used, so slight differences, as we observed here, are to be expected.

One advantage of using `poisson` with `reri` rather than `binreg` is that maximum likelihood optimization is more robust than iterated, reweighted least-squares optimization. There can be convergence problems with iterated, reweighted least squares when maximum likelihood will converge without any difficulty.

`binreg` does have an `ml` option for maximum likelihood optimization. We can also specify `vce(robust)` to reproduce the results given by `poisson`.

```
. reri binreg tube drug nitrate, ml vce(robust)
Fitting binreg ...
Interaction of drug and nitrate on an additive scale
Model: Binomial generalized linear          Number of observations = 1,284
drug#nitrate
       - +     No#High
       + -     Yes#Low
       + +     Yes#High
```

	ERR	Robust std. err.	z	P>|z|	[95% conf. interval]	
drug#nitrate						
- +	.1161419	.1571636	0.78	0.435	-.1530418	.4708787
+ -	.2385481	.2393248	1.11	0.268	-.1519233	.8088004
+ +	.7769424	.4072075	2.51	0.012	.133993	1.78443
RERI	.4222523	.4534688	0.93	0.352	-.4665302	1.311035
Attr. prop.	.2376286	.2187937	1.09	0.277	-.1911993	.6664564
Synergy ind.	2.190483	2.035131	0.84	0.399	.3545735	13.53235

```
Note: P>|z| for synergy index (SI) is for test H0: SI = 1.
```

`reri binreg` with `ml` and `vce(robust)` gives the same results as `reri poisson`. But if we want maximum likelihood with the robust variance estimator, we might as well just use `reri poisson`.

◁

▷ Example 3: Logistic model

We continue with the previous example. These data were simulated to mimic an observational study with full data from several sites over a set period of time. So risk ratios can be estimated directly. But let's pretend it was a case–control study. The outcomes, birth defects, have low prevalences, so odds ratios from a logistic regression should be good approximations to the risk ratios.

We specify `logistic` with `reri`:

```
. reri logistic tube drug nitrate

Fitting logistic ...

Interaction of drug and nitrate on an additive scale

Model: Logistic                              Number of observations = 1,284

drug#nitrate
         - +    No#High
         + -    Yes#Low
         + +    Yes#High
```

	ERR	Std. err.	z	P>|z|	[95% conf. interval]	
drug#nitrate						
- +	.141003	.1933785	0.78	0.436	-.1814898	.5905577
+ -	.2965734	.3085999	1.09	0.275	-.1867932	1.067251
+ +	1.080139	.6495028	2.35	0.019	.1280107	2.835939
RERI	.642563	.691611	0.93	0.353	-.7129695	1.998096
Attr. prop.	.3089038	.2594074	1.19	0.234	-.1995252	.8173329
Synergy ind.	2.468459	2.456832	0.91	0.364	.3509409	17.36273

Note: P>|z| for synergy index (SI) is for test H0: SI = 1.

The ERRs reported are the odds ratios minus one. Calculating ERR using odds ratios assumes that the odds ratios are approximations to the risk ratios.

Here we see that RERI = 0.643, AP = 0.309, and SI = 2.47. When we used `poisson` and `binreg`, we got RERI = 0.422, AP = 0.238, and SI = 2.19. Results are similar but not that close numerically.

If we have data from an observational study, a cohort, or a randomized controlled trial, we can estimate risk ratios directly, and we should do so using `poisson` or `binreg`. If it is a case–control study, we have no choice but to use `logistic`.

◁

▷ Example 4: Negative ERRs

We present data from example 2 in [ST] **stcox**. The data represent 48 participants in a cancer drug trial. Some receive treatment (`drug` = 1) and some receive a placebo (`drug` = 0). The data contain the participants' age and time until death. It has already been `stset` and is ready to be analyzed using a survival model.

We load the dataset, `describe` it, and run `st` to see the survival-time settings.

```
. use https://www.stata-press.com/data/r18/drugtr
(Patient survival in drug trial)

. describe

Contains data from https://www.stata-press.com/data/r18/drugtr.dta
 Observations:            48                  Patient survival in drug trial
    Variables:             8                  3 Mar 2022 02:12
─────────────────────────────────────────────────────────────────────────────
Variable      Storage   Display    Value
    name         type    format    label      Variable label
─────────────────────────────────────────────────────────────────────────────
studytime        byte    %8.0g                Months to death or end of exp.
died             byte    %8.0g                1 if patient died
drug             byte    %8.0g                Drug type (0=placebo)
age              byte    %8.0g                Patient's age at start of exp.
_st              byte    %8.0g                1 if record is to be used; 0
                                                otherwise
_d               byte    %8.0g                1 if failure; 0 if censored
_t               byte    %10.0g               Analysis time when record ends
_t0              byte    %10.0g               Analysis time when record begins
─────────────────────────────────────────────────────────────────────────────
Sorted by:

. st
-> stset studytime, failure(died)

Survival-time data settings

         Failure event: died!=0 & died<.
Observed time interval: (0, studytime]
     Exit on or before: failure
```

The RERI statistic requires two 0/1 predictors. These data have two predictors, drug and age, but age is not 0/1. Let's create a 0/1 variable from age. First, we summarize age:

```
. summarize age, detail

                  Patient's age at start of exp.
─────────────────────────────────────────────────────────────────────────────
       Percentiles      Smallest
 1%          47             47
 5%          48             48
10%          49             48         Obs                 48
25%        50.5             49         Sum of wgt.         48

50%          56                        Mean            55.875
                          Largest      Std. dev.     5.659205
75%          60             65
90%          65             67         Variance        32.0266
95%          67             67         Skewness       .3161066
99%          67             67         Kurtosis       2.125197
```

The age of the participants ranges from 47 to 67 years. Median age is 56. Second, we create the variable older with older = 1 representing those with ages > 56 and older = 0 those with ages ≤ 56. We give it a value label as well.

```
. generate older = cond(age > 56, 1, 0)

. label define older_lbl 0 "age<=56" 1 "age>56"

. label values older older_lbl
```

Now, we run `reri stcox`:

```
. reri stcox drug older

Fitting stcox ...

Interaction of drug and older on an additive scale

Model: Cox proportional hazards                    Number of observations = 48

   drug#older
      - +     0#age>56
      + -     1#age<=56
      + +     1#age>56
```

	ERR	Std. err.	z	P>|z|	[95% conf.	interval]
drug#older						
- +	.6268459	.7688877	1.03	0.303	-.3557603	3.108141
+ -	-.9158604	.0532904	-3.91	0.000	-.9756843	-.7088514
+ +	-.7313977	.1392912	-2.53	0.011	-.9027935	-.2577947
RERI	-.4423833	.7436483	-0.59	0.552	-1.899907	1.015141
Attr. prop.	-1.646983	2.711232	-0.61	0.544	-6.960901	3.666935
Synergy ind.	2.530662	6.713001	0.35	0.726	.0139727	458.3401

```
Notes: P>|z| for synergy index (SI) is for test H0: SI = 1.
       Some estimates of excess relative risk are not positive.
```

The ERR is negative for the +- (drug = 1 and older = 0) and ++ (drug = 1 and older = 1) interactions. This means the relative risk is less than one.

The RERI, AP, and SI statistics were designed for models in which the −+, +−, and ++ interactions each have positive ERRs. That is, values of 1 for the predictors are hypothesized to be associated with increased risk. Clearly, this is not true here. drug = 1 is hypothesized to be associated with decreased risk. We incorrectly specified the model for `reri`. drug = 0 should be the category hypothesized for a positive ERR.

We can correct this by specifying the option `baselevel1(1)`, which tells `reri` to use drug = 1 as the base level for the first predictor.

```
. reri stcox drug older, baselevel1(1)

Fitting stcox ...

Interaction of drug and older on an additive scale

Model: Cox proportional hazards                    Number of observations = 48

   drug#older
      - +     1#age>56
      + -     0#age<=56
      + +     0#age>56
```

	ERR	Std. err.	z	P>|z|	[95% conf.	interval]
drug#older						
- +	2.192339	1.962714	1.89	0.059	-.0433052	9.652331
+ -	10.885	7.527446	3.91	0.000	2.434672	40.12572
+ +	18.33507	12.94706	4.42	0.000	4.20437	70.83291
RERI	5.257728	8.484262	0.62	0.535	-11.37112	21.88658
Attr. prop.	.271927	.3309285	0.82	0.411	-.376681	.920535
Synergy ind.	1.402049	.677738	0.70	0.484	.5436253	3.615984

```
Note: P>|z| for synergy index (SI) is for test H0: SI = 1.
```

The ERRs are now all positive, and the model is easy to interpret. The RERI statistic is 5.26, so the effects are superadditive. But the RERI statistic has a p-value of 0.535, so we cannot reject the additive model.

The message of this example is clear. If you have negative ERRs, first check that you correctly specified the hypothesized direction of effects. If the model is correctly specified and there are negative ERRs (not close to 0), then the RERI, AP, and SI statistics may not be useful descriptors for the observed effects.

◁

Stored results

reri stores the following in `r()`:

Scalars
r(N)	number of observations
r(reri)	RERI
r(ap)	AP
r(si)	SI
r(level)	confidence level

Matrices
r(reri_table)	table of results

The `e()` and `r()` stored results from *cmd* are returned as well.

Methods and formulas

The RERI statistic is

$$\text{RERI} = \text{ERR}_{A+B+} - \text{ERR}_{A-B+} - \text{ERR}_{A+B-}$$

where $\text{ERR} = \text{RR} - 1$ is the ERR and RR denotes relative risk. Here $A+B+$ refers to subjects positive for exposure A and positive for exposure B, with $A-B+$ and $A+B-$ defined similarly.

The AP is

$$\text{AP} = \frac{\text{RERI}}{\text{RR}_{A+B+}}$$

The SI is

$$\text{SI} = \frac{\text{ERR}_{A+B+}}{\text{ERR}_{A-B+} + \text{ERR}_{A+B-}}$$

When incidence-rate ratios, odds ratios, or hazard ratios are estimated rather than risk ratios, they are used in the calculation in place of relative risk.

Standard errors are calculated by `nlcom` using the "delta method". See *Methods and formulas* in [R] **nlcom** for details. The standard error for the SI is calculated by first calculating the standard error for the logarithm of SI and then transforming the result back to the SI scale.

Acknowledgment

References

Andersson, T., L. Alfredsson, H. Källberg, S. Zdravkovic, and A. Ahlbom. 2005. Calculating measures of biological interaction. *European Journal of Epidemiology* 20: 575–579. https://doi.org/10.1007/s10654-005-7835-x.

Brender, J. D., P. J. Weyer, P. A. Romitti, B. P. Mohanty, M. U. Shinde, A. M. Vuong, J. R. Sharkey, D. Dwivedi, S. A. Horel, J. Kantamneni, J. C. Huber, Jr., Q. Zheng, M. M. Werler, K. E. Kelley, J. S. Griesenbeck, F. B. Zhan, P. H. Langlois, L. Suarez, M. A. Canfield, and the National Birth Defects Prevention Study. 2013. Prenatal nitrate intake from drinking water and selected birth defects in offspring of participants in the National Birth Defects Prevention Study. *Environmental Health Perspectives* 121: 1083–1089. https://doi.org/10.1289/ehp.1206249.

Bruun, N. H., M. Fenger-Gron, and A. Prior. 2017. ic: Module to compute measures of interaction contrast (biological interaction). Boston College Department of Economics, Statistical Software Components S457975. https://ideas.repec.org/c/boc/bocode/s457975.html.

Cummings, P. 2009. Methods for estimating adjusted risk ratios. *Stata Journal* 9: 175–196.

Greenland, S., and D. C. Thomas. 1982. On the need for the rare disease assumption in case–control studies. *American Journal of Epidemiology* 116: 547–553. https://doi.org/10.1093/oxfordjournals.aje.a113439.

Kalbfleisch, J. D., and R. L. Prentice. 2002. *The Statistical Analysis of Failure Time Data*. 2nd ed. New York: Wiley.

Knol, M. J., T. J. VanderWeele, R. H. H. Groenwold, O. H. Klungel, M. M. Rovers, and D. E. Grobbee. 2011. Estimating measures of interaction on an additive scale for preventive exposures. *European Journal of Epidemiology* 26: 433–438. https://doi.org/10.1007/s10654-011-9554-9.

Lash, T. L., T. J. VanderWeele, S. Haneuse, and K. J. Rothman. 2021. *Modern Epidemiology*. 4th ed. Philadelphia: Wolters Kluwer.

Rothman, K. J., S. Greenland, and A. M. Walker. 1980. Concepts of interaction. *American Journal of Epidemiology* 112: 467–470. https://doi.org/10.1093/oxfordjournals.aje.a113015.

Skrondal, A. 2003. Interaction as departure from additivity in case–control studies: A cautionary note. *American Journal of Epidemiology* 158: 251–258. https://doi.org/10.1093/aje/kwg113.

Also see

[R] **binreg** — Generalized linear models: Extensions to the binomial family

[R] **Epitab** — Tables for epidemiologists

[R] **logistic** — Logistic regression, reporting odds ratios

[R] **nbreg** — Negative binomial regression

[R] **poisson** — Poisson regression

[ST] **stcox** — Cox proportional hazards model

[ST] **stintcox** — Cox proportional hazards model for interval-censored survival-time data

[ST] **stintreg** — Parametric models for interval-censored survival-time data

[ST] **streg** — Parametric survival models

For suggested citations, see the FAQ on citing Stata documentation.