On the use and misuse of the most common measures of association in clinical and epidemiological studies

Rino Bellocco, Sc.D.

Department of Statistics and Quantitative Methods, University of Milano-Bicocca
&
Department of Epidemiology and Biostatistics, Karolinska Institutet

November 3, 2022

Outline

How to deal with Measures of Associations: M. Knol et al, Cerebrovasc Dis, 2012

- In reading medical literature doctors, clinician look at measures of association
- Can the results be applied to clinical practice?
- Therefore we need to make sure that these measures of association are clearly understood and properly interpreted
- Of course, we need also to deal with the study design in order to understand what we can estimate

Relative Risk (RR) and Odds Ratio (OR) in a RCT

- 500 patients are treated with drug A
- 500 patients are treated with placebo
- The outcome is survival for 30 days.

Relative Risk (RR) and Odds Ratio (OR) in a RCT

```
           |         surv
      drug |         1          0 |     Total
-----------+----------------------+----------
   Placebo |       350        150 |       500
           |     70.00      30.00 |    100.00
-----------+----------------------+----------
    Drug A |       425         75 |       500
           |     85.00      15.00 |    100.00
-----------+----------------------+----------
     Total |       775        225 |     1,000
           |     77.50      22.50 |    100.00
```

Relative Risk (RR) and Odds Ratio (OR) in a RCT

```
cs  surv drug, or

                 |         drug            |
                 |   Exposed   Unexposed   |      Total
-----------------+-------------------------+-----------
           Cases |       425         350   |        775
        Noncases |        75         150   |        225
-----------------+-------------------------+-----------
           Total |       500         500   |       1000
                 |                         |
            Risk |       .85          .7   |       .775
                 |                         |
                 |      Point estimate     |    [95% conf. interval]
                 |-------------------------+-----------------------
      Risk ratio |         1.214286        |    1.134255    1.299963
      Odds ratio |         2.428571        |    1.780251    3.312825
                 +-------------------------------------------------
                               chi2(1) =    32.26  Pr>chi2 = 0.000
```

Examples: Schulman at al (NEJM, 1999)

The effect of race and sex on physicians' recommendations for cardiac catheterization

TABLE 1. RATE OF REFERRAL FOR CARDIAC CATHETERIZATION, ODDS OF REFERRAL, ODDS RATIO, AND RISK RATIO ACCORDING TO SEX AND RACE.*

PATIENTS	MEAN REFERRAL RATE	ODDS OF REFERRAL	ODDS RATIO (95% CI)	RISK RATIO (95% CI)
	%			
Four strata				
White men†	90.6	9.6 to 1	1.0	
Black men	90.6	9.6 to 1	1.0 (0.5–2.1)	
White women	90.6	9.6 to 1	1.0 (0.5–2.1)	
Black women	78.8	3.7 to 1	0.4 (0.2–0.7)	0.87 (0.80–0.95)
Aggregate data				
White†	90.6	9.6 to 1	1.0	
Black	84.7	5.5 to 1	0.6 (0.4–0.9)	0.93 (0.89–0.99)
Men†	90.6	9.6 to 1	1.0	
Women	84.7	5.5 to 1	0.6 (0.4–0.9)	0.93 (0.89–0.99)
Overall	87.7	7.1 to 1		

*Referral rates for the four strata were inferred from aggregate rates and odds ratios reported by Schulman et al.[1] The odds of referral were calculated according to the following formula: referral rate÷(100%−referral rate). The risk ratio was calculated as the referral rate for the group in question divided by the referral rate for the reference group. CI denotes confidence interval.

†This was the reference group.

Examples: Schulman at al (NEJM, 1999)

The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization

- Misunderstandings About The Effects of Race and Sex on Physicians' Referrals For Cardiac Catheterization
- New York Times: "Doctor Bias May Affect Heart Care, Study Finds"
- Doctors are only 60% as likely to order cardiac catheterization for women and blacks as for men and whites
- Women and blacks complaining of chest pain are less likely than men and whites to receive the best cardiac testing
- Unconscious prejudices among doctors may help explain the findings
- Interpreting OR's as RRs can thus yield to wrong conclusions which could seriously impact on treatment decision effect

Example: Mansi et al (JAMA, 2013)

Statins and Musculoskeletal Conditions, Arthropathies, and Injuries

- Statins Can Weaken Muscles and Joints: Cholesterol Drug Raises Risk of Problems by up to 20 per cent, Mail Online, 3 June 2013
- A study published in 2013 found that 87 % of people taking statins reported muscle pains, compared to 85% in those who did not take statins
- We might report either a 2% increase in absolute risk, or a relative risk of 0.87/0.85 = 1.02 (a 2% relative increase in risk)

Example: Mansi et al (JAMA, 2013) - con't

- The odds in the two groups are given by 0.87/0.13 = 6.7 and 0.85/0.15 = 5.7, and so the odds ratio is therefore 6.7/5.7 = 1.18
- The Daily Mail misinterpreted this odds ratio of 1.18 as a relative risk, and produced a headline claiming statins 'raises risk by up to 20 per cent', which is a serious misrepresentation of what the study actually found.
- the abstract of the paper mentioned only the odds ratio without mentioning that this corresponded to a difference between absolute risks of 85% vs 87%

Definitions

- Measures of Risk - Absolute Versus Relative
  - Risk Difference (RD)
  - Relative Risk (RR)
  - Odds Ratio (OR)
- Estimating RD and OR
- Estimating Adjusted RR
  - Logistic Regression with Transformation
  - Binomial Regression
  - Modified Poisson Regression
- Summary

Outline

- Study relationship between two binary variables E and D
- Binary variables: 0/1 or No/Yes
- Usually expressed as

$$RD = ER = P_1 - P_0$$

- The RD looks at the absolute, rather than relative, difference in risk levels.
- It can be estimated in both RCT and cohort studies
- If E means treatment, like in a RCT, we can calculate NNT as the inverse and it tells us how many patients need to be treated with the drug to prevent 1 outcome
- Absolute risk are important in Public Health, but are important also to understand Relative Risk.

Outline

- Study relationship between two binary variables
  - Binary variables: 0/1 or No/Yes
- Usually expressed as
  - At how much greater risk of D is one group of patients than another?
- Example
  - At how much greater risk of osteoarthritis (OA) are women than men?
  - Patients having an anterior infarct are 50% more likely to die within 48 hours of hospital admission than are patients having just sustained an infarct at another primary site

- Aim is to take into account for differences between groups in other variables
  - Remove the effects of these variables from the group difference
- Example
  - At how much higher risk of OA are women than men after controlling for age and BMI?
  - Anterior infarct patients history are still more likely to die compared with other primary site infarct patients, all with similar co-morbidity history?

- Relative Risk (RR)
  - Ratio of the probabilities of the occurrence of the outcome of interest in group 1 (usually exposed) to group 0 (unexposed)

$$RR = \frac{P_1}{P_0}$$

- P_1 is the probability of the outcome in group 1
- P_0 is the probability of the outcome in group 0
- If exposure (E) and outcome (D) are independent $RR = 1$

Advantages

- Easy to communicate and interpret (RR)

Disadvantages

- RR must be greater than 0
- Given a measure of baseline risk then

$$RR \times P_0 = P_1$$

- But P_1 is a probability, therefore must be always less than 1
- Therefore

$$(RR \times P_0) < 1 \rightrightarrows RR < \frac{1}{P_0}$$

- This restriction on the range of RR only becomes an issue with common disease outcomes.

A final important comment on the Relative Risk is that it is not symmetric in the role of the two factors D and E. The Relative Risk for E associated with D is a different measure of association

$$\frac{P(D|E)}{P(D|not \ \ E)} \neq \frac{P(E|D)}{P(E|not \ \ D)}$$

Outline

- Odds
  - Odds are the probability of occurrence of disease/death divided by the probability of non-occurrence (disease free, surviving)
  - Odds among exposed

$$\text{Odds}_1 = \frac{P_1}{1 - P_1}$$

  - Odds among unexposed

$$\text{Odds}_0 = \frac{P_0}{1 - P_0}$$

- Odds are used a lot in gambling
    - The odds are two to one for Manchester City to win
    - 2:1 $\Rightarrow$ odds = 2 $\Rightarrow$Pr = 0.67
- Translating odds to probabilities

Odds	=	3.0	$\Leftrightarrow$	P	=	0.75
Odds	=	2.0	$\Leftrightarrow$	P	=	0.67
Odds	=	1.0	$\Leftrightarrow$	P	=	0.50
Odds	=	0.5	$\Leftrightarrow$	P	=	0.33

- Odds ratio (OR)
- **Ratio** of the *odds* of the occurrence of the event of interest in group 1 to group 0

$$OR = \frac{\text{Odds}_1}{\text{Odds}_0} = \frac{\frac{P_1}{1-P_1}}{\frac{P_0}{1-P_0}}$$

- As with the Relative Risk, the null value of the Odds Ratio is OR=1, again equivalent to independence of D and E
- In addition, $OR > 1$ when there is a greater risk of D with E present, and $OR < 1$ when there is a lower risk of D if E is present.
- The Odds Ratio is also the basis of a multiplicative model for the risk of D. Like RR, OR must be nonnegative, but unlike RR, OR has no upper limit whatever the baseline risk $P(D|not\ \ E)$ for the unexposed.
- Thus, the Odds Ratio can be effectively used as a scale for association even when $P(D|not\ \ E)$ is large.

Outline

RR and OR are *ratio* measures

- 1.0 is the point of no difference between groups (the null value)
- Are greater than 1 if group 1 is at higher risk relative to group 0
- Are less than 1 if group 1 is at lower risk relative to group 0
- Reciprocals are the same distance from the null value
  - E.g. 2 and 1/2 are equivalent group differences

Hypothetical Data for a Trial of Drug X

Table 2. Hypothetical Data for a Trial of Drug *X*

Treatment	Outcome, No. Died	Outcome, No. Survived	Risk of Death	Odds of Death
Drug *X*	25	75	25/(25 + 75) = .25	25/75 = 0.33
Placebo	50	50	50/(50 + 50) = .50	50/50 = 1.00

- There is symmetry for both the odds and risk ratios with regard to the definition of the exposure: both ratio estimates for treatment with X compared with no X are the inverse of the ratio estimates for no X compared with treatment with X
- However, if we change the definition of the outcome from the occurrence of Y to no occurrence of Y, only the odds ratio is symmetrical

- The odds ratio for Y among those treated with X compared with those who did not get X is =(25/75)/(50/50)=(1/3)/(1)=0.33
- The odds ratio for no occurrence of Y among those treated with drug X compared with those who did not get X is =(75/25)/(50/50)=3/1=3
- These odds ratios are simply the reciprocal of each other.

- The corresponding risk ratios are = (25/100)/(50/100) = 0.5 and
- (75/100)/(50/100)=1.5
- These risk ratios are not reciprocal
- The symmetry property of the odds ratio is attractive because 1 odds ratio can summarize the association of X with Y, and the choice between outcome Y and outcome not Y is unimportant

- The RR is more understandable for clinicians
- When the RR=2 then the probability of the outcome in group 1 is twice that of group 0
- This is not true for the odds ratio
- Most people are more comfortable with probabilities or percentages that with odds

- The OR has some advantages
- In case-control studies the OR can be estimated but not the RR
- The OR is symmetric to which outcome level is chosen as being of interest, the RR is not

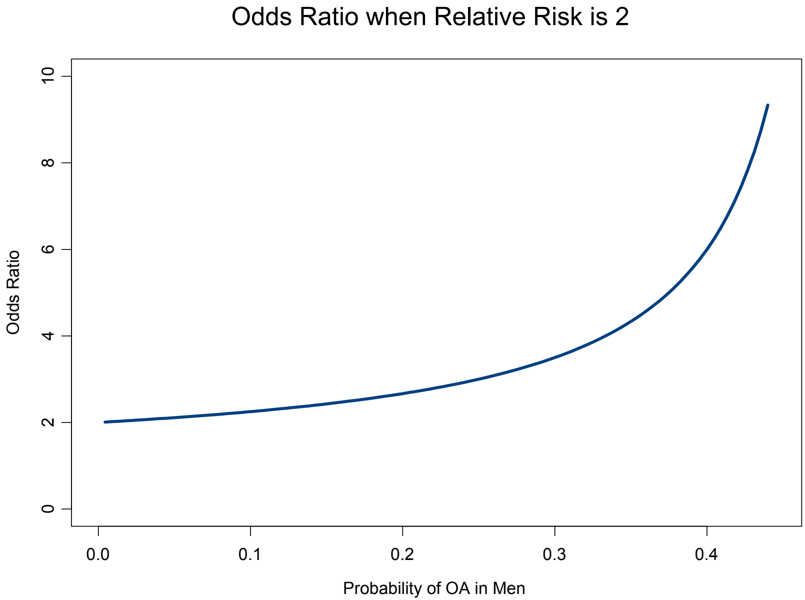
When are the RR and OR Similar?

- If the probability of the event is small, the odds and the probability are close

$$\text{Odds}_1 = \frac{P_1}{1 - P_1} \approx P_1$$

- When the probability of the event is small in both groups the OR is a good approximation to the RR
- Rule of thumb for small: P < 0.1

- The OR is always more extreme (farther from 1) than the RR
- When the events of interest are common, the OR can be much larger than the RR


Odds Ratio when Relative Risk is 2
Odds Ratio
0
2
4
6
8
10
0.0
0.1
0.2
0.3
0.4
Probability of OA in Men

Which is better to report?

- For case-control studies need to present the OR
- For cohort studies and clinical trials the RR is better to report
    - Reduces the chance of incorrect interpretation
    - Becoming preferred to report RR in medical journals

Osteoarthritis in Framingham

- In the Framingham Osteoarthritis study, prevalence of osteoarthritis (OA) was measured in 1992-93
- Female sex is an established risk factor for OA
- At how much greater risk of osteoarthritis are women than men in this study?

Osteoarthritis in Framingham

- Subset of 840 subjects to evaluate the prevalence of OA in women versus men
- 538 women, 302 men, 513 (61%) no OA, 327 (39%) with OA

```
           |         osteo
    gender |         0          1 |     Total
-----------+----------------------+----------
         0 |       316        222 |       538
           |     58.74      41.26 |    100.00
-----------+----------------------+----------
         1 |       197        105 |       302
           |     65.23      34.77 |    100.00
-----------+----------------------+----------
     Total |       513        327 |       840
           |     61.07      38.93 |    100.00
```

Osteoarthritis in Framingham

```
cs  osteo gender [fw=count ], or

           |       gender          |
           |     FEMALE      MALES |    Total
-----------+-----------------------+----------
     Cases |     222          105  |     327
  Noncases |     316          197  |     513
-----------+-----------------------+----------
     Total |     538          302  |     840
           |                       |
      Risk |.4126394     .3476821  |.3892857
           |                       |
           |    Point estimate     | [95% Conf. Interval]
           |-----------------------+----------------------
Risk ratio |        1.18683        | .9869029    1.427257
Odds ratio |       1.318083        | .9841915    1.765194
          +-----------------------------------------------
chi2(1) =     3.43  Pr>chi2 = 0.0639
```

Osteoarthritis in Framingham

- Women have 1.19 times the **risk** of OA compared to men
- Women have 1.30 times the **odds** of OA compared to men
- If we interpret OR as an RR, we would mistakenly conclude women are at 1.3 times the risk of OA

Osteoarthritis in Framingham

- Suppose now we consider as outcome **NOT HAVING** developed OA
- RR for No OA is 0.59/0.65 = 0.91
- But RR for OA is 1.19 and 1/1.19 = 0.84
- The RR implies that sex plays a larger role for OA than for No OA

Osteoarthritis in Framingham

- RR is not symmetric around the null value for both outcome levels
  - RR for No OA $\neq$ 1/RR for OA
- OR is symmetric
  - OR for No OA $\neq$ 1/OR for OA
- Usually the outcome to choose is clear. But some situations are not clear (E.g. use 'lived' or 'died')?

Outline

- Risk ratios, but not odds ratios, have a mathematical property called collapsibility
- the size of the risk ratio will not change if adjustment is made for a variable that is not a confounder
- Because of collapsibility the risk ratio, assuming no confounding, has a useful interpretation as the ratio change in average risk due to exposure among the exposed
- Because odds ratios are not collapsible, they usually lack any interpretation either as the change in average odds or the average change in odds (the average odds ratio)

See

- Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. Am J Epidemiol. 1987;125(5):761-768.
- Newman SC. Biostatistical Methods in Epidemiology. New York, NY: John Wiley & Sons; 2001:33-67, 132-134, 148-149.

- Logistic regression provides adjusted OR
- But, until recently it has been difficult to obtain adjusted RR
- Modified Poisson regression

Logistic Regression

- Logistic regression is widely used regression method for binary outcomes
- Logistic regression coefficients are log(OR)
- Provides adjusted OR if adjustors are used as additional predictors

Logistic Regression

- If outcome probabilities are < 0.1 for all values of the predictors then the OR are good approximations to RR
- Otherwise Zhang and Yu proposed a formula to convert OR to RR

$$RR = \frac{OR}{(1 - P_0) + (P_0 \times OR)}$$

Logistic Regression

- However the conversion formula has been criticized (McNutt et al.)
- Leads to confidence intervals for RR that are too small
- Gives biased estimate if some regression predictors are confounders
- Does not work if there are interactions in the regression model

Binomial Regression

- Binomial regression is a rarely used regression method for binary outcomes
- Binomial regression coefficients are log(RR)
- Provides adjusted RR if adjustors are used as additional predictors

Binomial Regression

- This model often fails due to numerical problems
- Especially failure prone if
  - Correlated predictors
  - One or more continuous predictors

Modified Poisson Regression

- Poisson regression is a method for count outcomes
- Poisson regression coefficients are log(RR)
- Provides adjusted RR if adjustors are used as additional predictors
- Poisson regression is conservative for binary outcomes
  - Less likely to be significant
  - Confidence intervals too wide

Modified Poisson Regression

- Modification due to Zou
  - Adjust variability with generalized estimating equations (GEE)
  - Uses variability in the data to adjust model
- This has been shown to work very well
- Software implementation
  - SAS in Lundquist
  - Stata in Barros and Hirakata (2003, Biomed CMRM)
  - Nice overview in Stata Journal, 2009 by P. Cunnings.

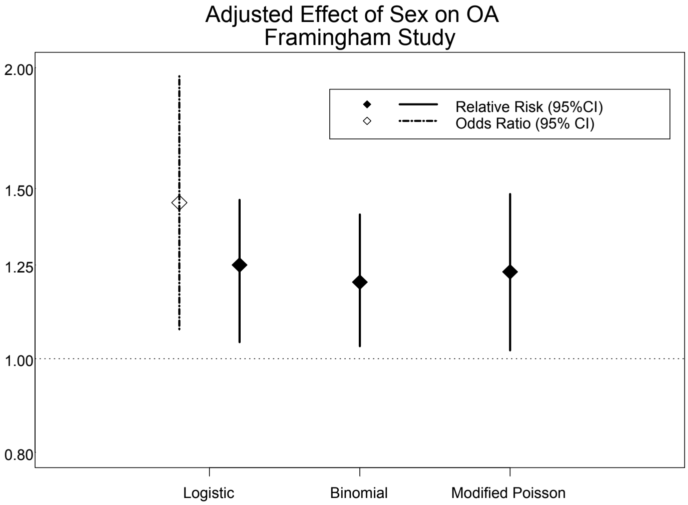
Osteoarthritis in Framingham

- Greater risk of OA in women than men was found
- Could this be due to age differences between women and men?
  - Could this be due to age differences between women and men?
  - Could this be due to differences in body mass index between women and men?
  - Use regression models with sex, age, and body mass index

Osteoarthritis in Framingham

- Logistic OR = 1.45
- Transformed Logistic RR = 1.25
- Binomial RR = 1.20 (convergence issues)
- Modified Poisson RR = 1.23

Osteoarthritis in Framingham



New Approach: Conditional and Marginal Standardization

Conditional Standardization : Relative Risks from Logistic regression

- logit(P) $= \log[\frac{P}{(1-P)}] = \beta_0 + \beta_1 \times E + \beta_2 \times x_2 + \beta_3 \times x3$
- $x_2 = 0$ represents a baseline value for x_2
- $x_3 = 0$ represents the reference value for x_3
- $P_1 = exp(\beta_0 + \beta_1)/(1 + exp(\beta_0 + \beta_1))$
- $P_0 = exp(\beta_0/(1 + exp(\beta_0))$

$$RR = \frac{1 + exp(-\beta_0)}{1 + exp(-\beta_0 - \beta_1)}$$

Conditional and Marginal Standardization

Marginal Standardization with Logistic regression

- Does not require fixing values of covariates
- $x_{i,p-1}\beta_{p-1}$
- $r_{i1} = \text{expit}(x_{i,p-1}\beta_{p-1} + \beta_1 \times \mathbf{1}$
- $r_{i0} = \text{expit}(x_{i,p-1}\beta_{p-1} + \beta_1 \times \mathbf{0}$
- Marginal Standardized Risk: Mean r_{i1} / Mean r_{i0}

DCCT trial (Diabetes Control and Complications Trial)

- Goal is determine the relative risk of standard therapy versus intensive treatments in terms of the prevalence of microalbuminuria at 6 years of follow-up.
- Covariates requiring adjustment are the percentage of total hemoglobin that has become glycosylated at baseline, the prior duration of diabetes in months, the level of systolic blood pressure (mmHg), and gender (female) (1 if female, 0 if male).

```
. univar  micro24 intens hbael duration sbp female
                                         ---------- Quantiles ------
Variable        n      Mean      S.D.         .25      Mdn       .75
--------------------------------------------------------------------
 micro24      172      0.24      0.43        0.00     0.00      0.00
  intens      172      0.52      0.50        0.00     1.00      1.00
   hbael      172      9.26      1.48        8.11     9.10     10.14
duration      172    113.16     40.07       84.00   116.00    144.00
     sbp      172    116.33     10.81      110.00   118.00    124.00
  female      172      0.45      0.50        0.00     0.00      1.00
--------------------------------------------------------------------
```

DCCT trial (Diabetes Control and Complications Trial)

```
. oddsrisk  micro24  intens hbael duration sbp female

-----------------------------------------------------------------
Incidence for unexposed risk group =     0.3735
-----------------------------------------------------------------
Predictor     Odds Ratio   Risk Ratio     [95% Conf. Interval]
-----------------------------------------------------------------
intens          0.2053       0.2920        0.1348       0.5898
hbael           1.7639       1.3723        1.1830       1.5604
duration        1.0008       1.0005        0.9940       1.0070
sbp             1.0236       1.0146        0.9891       1.0404
female          0.4104       0.5263        0.2474       0.9913
-----------------------------------------------------------------
```

```
. poisson  micro24  intens hbael duration sbp female, irr robust

Iteration 0:   log pseudolikelihood = -88.738164
Iteration 1:   log pseudolikelihood = -88.737623
Iteration 2:   log pseudolikelihood = -88.737623

Poisson regression                                Number of obs    =        172
                                                  Wald chi2(5)     =      30.56
                                                  Prob > chi2      =     0.0000
Log pseudolikelihood = -88.737623                 Pseudo R2        =     0.1233

------------------------------------------------------------------------------
             |               Robust
     micro24 |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      intens |   .3395539   .1036305    -3.54   0.000     .1866919    .6175782
       hbael |   1.404157   .1177613     4.05   0.000      1.19132    1.655018
    duration |   .9997331   .0032366    -0.08   0.934     .9934095    1.006097
         sbp |   1.013085   .0146194     0.90   0.368     .9848329    1.042147
      female |   .5379037    .161796    -2.06   0.039     .2983127    .9699232
------------------------------------------------------------------------------
```

DCCT trial (Diabetes Control and Complications Trial)

```
. glm micro24  intens hbael duration sbp female, f(binomial) link(log) eform difficult

Iteration 0:   log likelihood = -133.80429  (not concave)
Iteration 1:   log likelihood = -82.246511
Iteration 2:   log likelihood = -79.550896
Iteration 3:   log likelihood = -79.221869
Iteration 4:   log likelihood = -79.218877
Iteration 5:   log likelihood = -79.218875
```

DCCT trial (Diabetes Control and Complications Trial)

```
. glm micro24  intens hbael duration sbp female, f(binomial) link(log) eform difficult
Generalized linear models                         Number of obs     =        172
Optimization     : ML                             Residual df       =        166
                                                  Scale parameter =          1
Deviance         =  158.4377504                   (1/df) Deviance =   .9544443
Pearson          =  169.7421233                   (1/df) Pearson  =   1.022543

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

                                                  AIC               =   .9909172
Log likelihood   = -79.21887521                   BIC               =  -696.0463

------------------------------------------------------------------------------
             |                 OIM
     micro24 | Risk ratio   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
      intens |   .3504934   .1083385    -3.39   0.001     .1912367    .6423749
       hbael |    1.30492   .0837918     4.14   0.000     1.150605     1.47993
    duration |   .9988549   .0027826    -0.41   0.681      .993416    1.004324
         sbp |   1.005441   .0130563     0.42   0.676     .9801742     1.03136
      female |   .4823159   .1513856    -2.32   0.020     .2607131     .892278
       _cons |   .0228976   .0343245    -2.52   0.012     .0012128     .432291
------------------------------------------------------------------------------
```

DCCT trial (Diabetes Control and Complications Trial)

DCCT trial (Diabetes Control and Complications Trial)

```
. oddsrisk  micro24  intens female

-----------------------------------------------------------------------
Incidence for unexposed risk group =     0.3735
-----------------------------------------------------------------------
Predictor     Odds Ratio   Risk Ratio     [95% Conf. Interval]
-----------------------------------------------------------------------
intens          0.2306       0.3236       0.1589       0.6173
female          0.5556       0.6662       0.3615       1.1050
-----------------------------------------------------------------------
```

DCCT trial (Diabetes Control and Complications Trial)

```
. poisson  micro24  intens  female, irr robust
------------------------------------------------------------------------------
             |               Robust
     micro24 |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      intens |   .3298168   .1040912    -3.51   0.000     .1776768    .6122301
      female |   .6637346    .181328    -1.50   0.134     .3885543    1.133802
------------------------------------------------------------------------------

. binreg micro24  intens  female, rr ml

------------------------------------------------------------------------------
             |                 OIM
     micro24 | Risk Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      intens |   .3293066     .10346    -3.54   0.000     .1778992    .6095747
      female |   .6611327   .1782996    -1.53   0.125     .3896994    1.121625
------------------------------------------------------------------------------
```

DCCT trial (Diabetes Control and Complications Trial)

```
#delimit cr; /* Cancels end of line by ; */
logistic micro24  intens female
#delimit ; /* Allows for long lines terminated by ; */
bootstrap rrintens = ( ( 1+ exp(- _b[_cons]) ) / (1+exp(- _b[_cons] - _b[intens]) ))
, reps(999):logit micro24  intens female;
estat bootstrap;

Logistic regression                               Number of obs      =       172
                                                  Replications       =       999

      command:  logit micro24 intens female
     rrintens:  ( 1+ exp(- _b[_cons]) ) / (1+exp(- _b[_cons] - _b[intens]) )
------------------------------------------------------------------------------
             |    Observed               Bootstrap
             |       Coef.       Bias    Std. Err.  [95% Conf. Interval]
-------------+----------------------------------------------------------------
    rrintens |   .34683146  -.0024999   .10859388    .1671485    .615394  (BC)
------------------------------------------------------------------------------
(BC)   bias-corrected confidence interval
```

DCCT trial (Diabetes Control and Complications Trial)

```
> marginal micro24  intens female
Bootstrap results                                 Number of obs      =       172
                                                  Replications       =      1000

      command:  marginal micro24 intens female
          pm1:  r(sp1)
          pm0:  r(sp0)
           rr:  r(pmrr)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         pm1 |   .1233481   .0350331     3.52   0.000     .0546845    .1920117
         pm0 |   .3740823    .054592     6.85   0.000      .267084    .4810807
          rr |   .3297351   .1074472     3.07   0.002     .1191424    .5403277
------------------------------------------------------------------------------
```

DCCT trial (Diabetes Control and Complications Trial)

```
. estat bootstrap, all
Bootstrap results                               Number of obs     =        172
                                                Replications      =       1000

      command:  marginal micro24 intens female
          pm1:  r(sp1)
          pm0:  r(sp0)
           rr:  r(pmrr)
------------------------------------------------------------------------------
             |    Observed               Bootstrap
             |       Coef.       Bias    Std. Err.  [95% Conf. Interval]
-------------+----------------------------------------------------------------
         pm1 |   .12334807    -.00018   .03503309    .0546845   .1920117   (N)
             |                                       .0583328   .1944489   (P)
             |                                       .0583457   .1949345  (BC)
         pm0 |   .37408235   .0031302   .05459202     .267084   .4810807   (N)
             |                                       .2657416   .4832112   (P)
             |                                       .2609552   .4779253  (BC)
          rr |   .32973508   .0037562    .1074472    .1191424   .5403277   (N)
             |                                       .1511442   .5622342   (P)
             |                                       .1572924   .6158956  (BC)
------------------------------------------------------------------------------
(N)    normal confidence interval
(P)    percentile confidence interval
(BC)   bias-corrected confidence interval
```

Outline

- Medical literature is moving toward reporting RR instead of OR whenever possible
- Need to keep in mind that the RR changes in non-intuitive ways when outcomes are switched
- When reporting OR make it clear that it is not the RR
- Modified Poisson regression has been now used for obtaining adjusted RR
- Conditional and Marginal Methods can be applied and seem to have best statistical properties

Outline

- Simon SD. Understanding the odds ratio and the relative risk. Journal of Andrology; 22:533-6, 2001
- Zhang J, Yu KF. What's the relative risk? Journal Of American Medical Association; 280: 1690-1, 1998
- McNutt L-A, Wu C, XueX, HafnerJP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. American Journal of Epidemiology; 157: 940-3, 2003
- Barros AJD, Hirakata VN. Alternatives for logistic regression in cross-sectional studies. BMC Medical Research Methodology 3:21, 2003
- ZouG. A modified poisson regression approach to prospective studies with binary data. American Journal of Epidemiology1; 59:702-6, 2004
- Lundquist K. SAS FAQ: How to estimate relative risk in SAS using PROC GENMOD for common outcomes in cohort studies. https://stats.oarc.ucla.edu/sas/faq/how-can-i-estimate-relative-risk-in-sas-using-proc-genmod-for-common-outcomes-in-cohort-studies/

- Knol M et al. How to deal with measures of association: a short guide for the clinician. Cerebrovasc Dis; 2012
- Knol M et al. Potential misinterpretation of treatment effects due to use of odds ratios and logistic regression in randomized controlled trials. Plos One; 2011.
- Knol M et al. Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression. CMAJ; 2012.
- Cummings P. Methods for estimating adjusted risk ratios. Stata Journal; 2009.
- Norton EC et al. Computing adjusted risk ratios and risk differences in Stata. Stata Journal; 2013.