Models In Epidemiology And Biostatistics
Gordon Hilton Fick

Example 2 with Annotation

```
. use Session_4_Example_2.dta

. cs dis expo,by(age gender) or

      age gender |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
             0 0 |    2.181717     1.256808   3.785856        10.35 (Cornfield)
             0 1 |    1.962957     1.327944   2.901236     18.43281 (Cornfield)
             1 0 |    3.305753     1.953334   5.592873     10.37538 (Cornfield)
             1 1 |    3.379872      1.92368   5.934744     7.374564 (Cornfield)
-----------------+-------------------------------------------------
           Crude |    .4575415     .3940612   .5312483
    M-H combined |    2.535571     1.981625   3.244369
-------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(3) =    3.851  Pr>chi2 = 0.2780

                   Test that combined OR = 1:
                                Mantel-Haenszel chi2(1) =     57.36
                                                Pr>chi2 =    0.0000
```

For the young, male and female OR estimates are close and about 2. For the old, male and female OR estimates are close and about 3.3. Modification by age ? Crude and adjusted are very different so there is some kind of confounding.

```
. cs dis expo,by(gender) or

          gender |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |    2.742677     1.871612    4.01882     20.74615 (Cornfield)
               1 |     2.34375     1.707093   3.217633     26.35914 (Cornfield)
-----------------+-------------------------------------------------
           Crude |    .4575415     .3940612   .5312483
    M-H combined |    2.519446     1.971614   3.219498
-------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =    0.386  Pr>chi2 = 0.5343

                   Test that combined OR = 1:
                                Mantel-Haenszel chi2(1) =     57.06
                                                Pr>chi2 =    0.0000
```

This 'one-at-a-time' analysis does not display the age difference seen in the joint analysis but does reveal a confounding issue.

```
. cs dis expo,by(age) or

             age |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |    .4168153     .3369634   .5155913     129.4351 (Cornfield)
               1 |    .5023695     .4072843   .6196537     121.6115 (Cornfield)
-----------------+-------------------------------------------------
           Crude |    .4575415     .3940612   .5312483
    M-H combined |    .4582593     .3946601   .5321074
-------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =    1.498  Pr>chi2 = 0.2210

                   Test that combined OR = 1:
                                Mantel-Haenszel chi2(1) =    107.15
                                                Pr>chi2 =    0.0000
```

This 'one-at-a-time' analysis misses the confounding and misses the modification.

```
. cs dis expo if age==0,by(gender) or

          gender |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |   2.181717      1.256808   3.785856        10.35 (Cornfield)
               1 |   1.962957      1.327944   2.901236     18.43281 (Cornfield)
-----------------+-------------------------------------------------
           Crude |   .4168153      .3369634   .5155913
    M-H combined |   2.041621      1.481015   2.814431
-------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =    0.093  Pr>chi2 = 0.7605

                   Test that combined OR = 1:
                                Mantel-Haenszel chi2(1) =     19.42
                                                Pr>chi2 =    0.0000
```

For the young, adjusting for gender looks promising.

```
. cs dis expo if age==1,by(gender) or

          gender |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |   3.305753      1.953334   5.592873     10.37538 (Cornfield)
               1 |   3.379872       1.92368   5.934744     7.374564 (Cornfield)
-----------------+-------------------------------------------------
           Crude |   .5023695      .4072843   .6196537
    M-H combined |   3.336548      2.257874   4.930545
-------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =    0.003  Pr>chi2 = 0.9553

                   Test that combined OR = 1:
                                Mantel-Haenszel chi2(1) =     40.27
                                                Pr>chi2 =    0.0000
```

For the old, adjusting for gender also seems reasonable.
There appears to be age modification but there is no test available here. The confidence intervals for the assumed common ORs overlap.

Now we can try a model based approach

```
. logit dis expo age gender ga ge ae gae

------------------------------------------------------------------------------
         dis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        expo |   .7801121   .2846032     2.74   0.006        .2223    1.337924
         age |   .0680535   .3690482     0.18   0.854    -.6552677    .7913746
      gender |   2.952439   .2795989    10.56   0.000     2.404435    3.500443
          ga |  -.0840535   .3901241    -0.22   0.829    -.8486828    .6805757
          ge |  -.1056603   .3478914    -0.30   0.761    -.7875148    .5761943
          ae |   .4155523   .3931613     1.06   0.291    -.3550297    1.186134
         gae |   .1278339   .5281787     0.24   0.809    -.9073773    1.163045
       _cons |  -2.944439   .2649065   -11.12   0.000    -3.463646   -2.425232
------------------------------------------------------------------------------
```

There is no evidence that age modification is modified by gender [p=0.809].
From the classic analysis, for the women, we got 3.31 / 2.18 = 1.5 and for the men, we get 3.38 / 1.96 =1.7.
Similar numbers can be computed for the above model. exp(0.4155) = 1.5 exp(0.4155+0.127) = 1.7

```
. logit dis expo age gender ga ge ae

------------------------------------------------------------------------------
         dis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        expo |   .7433503   .2377981     3.13   0.002     .2772746    1.209426
         age |   .0055253   .2633948     0.02   0.983    -.5107191    .5217697
      gender |   2.916984   .2351269    12.41   0.000     2.456144    3.377824
          ga |  -.0141751   .2622424    -0.05   0.957    -.5281608    .4998107
          ge |  -.0504734    .260819    -0.19   0.847    -.5616693    .4607224
          ae |   .4865294   .2617646     1.86   0.063    -.0265198    .9995786
       _cons |  -2.912659   .2268167   -12.84   0.000    -3.357212   -2.468107
------------------------------------------------------------------------------
```

0.4865 is the estimate of the log of the ratio of assumed common ratio of ORs [old over young : adjusted for gender] . This is interesting and is supported by the classic analysis.
-0.0504 is this estimate [male over female adjusted for age]. It is near zero.

```
. logit dis expo age gender ga ae

------------------------------------------------------------------------------
         dis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        expo |   .7100288   .1628327     4.36   0.000     .3908825    1.029175
         age |  -.0017126   .2597452    -0.01   0.995    -.5108038    .5073787
      gender |   2.884892   .1654006    17.44   0.000     2.560713    3.209071
          ga |  -.0060814   .2583501    -0.02   0.981    -.5124383    .5002755
          ae |   .4959979   .2566618     1.93   0.053    -.0070499    .9990458
       _cons |  -2.883975   .1701056   -16.95   0.000    -3.217376   -2.550574
------------------------------------------------------------------------------
```

We will take the step to delete the ga term. It is instructive to interpret it though.

```
. logit dis expo age gender ae

Logistic regression                             Number of obs     =      3,814
                                                LR chi2(4)        =     840.37
                                                Prob > chi2       =     0.0000
Log likelihood = -1710.7169                     Pseudo R2         =     0.1972

------------------------------------------------------------------------------
         dis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        expo |   .7084344   .1480128     4.79   0.000     .4183347    .9985341
         age |  -.0071415    .119512    -0.06   0.952    -.2413807    .2270977
      gender |   2.882402   .1270758    22.68   0.000     2.633338    3.131466
          ae |   .5004742   .1724266     2.90   0.004     .1625242    .8384242
       _cons |  -2.881753   .1414169   -20.38   0.000    -3.158925   -2.604581
------------------------------------------------------------------------------
```

Now we have the model that provides for age modification adjusted for gender.

```
. logit dis expo age gender ae, or

------------------------------------------------------------------------------
         dis | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        expo |   2.030809   .3005857     4.79   0.000     1.519429      2.7143
         age |   .9928839   .1186615    -0.06   0.952     .7855425    1.254952
      gender |   17.85711   2.269207    22.68   0.000     13.92015    22.90753
          ae |   1.649503   .2844183     2.90   0.004     1.176477     2.31272
       _cons |   .0560365   .0079245   -20.38   0.000     .0424714    .0739341
------------------------------------------------------------------------------
```

```
. disp 2.030809*1.649503
3.3498255
```

Taking the exponent of the coefficient estimates gives us numbers very close to the classic analysis. Perhaps most crucially this model provides clear evidence [p=0.004] that age modifies after adjusted for gender.

```
. logit dis expo age ae

------------------------------------------------------------------------------
         dis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        expo |  -.8751117   .1085975    -8.06   0.000    -1.087959   -.6622645
         age |  -.0056374   .1061829    -0.05   0.958     -.213752    .2024772
          ae |   .1866923   .1525511     1.22   0.221    -.1123025     .485687
       _cons |  -.6968995    .075047    -9.29   0.000     -.843989     -.54981
------------------------------------------------------------------------------

. logit dis expo age ae,or

------------------------------------------------------------------------------
         dis | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        expo |   .4168155   .0452651    -8.06   0.000     .3369034    .5156822
         age |   .9943785    .105586    -0.05   0.958     .8075486    1.224432
          ae |   1.205256   .1838632     1.22   0.221     .8937739    1.625291
       _cons |   .4981273    .037383    -9.29   0.000     .4299919    .5770594
------------------------------------------------------------------------------
```

We can now compare 1.6495 [assessing age modification adjusted for gender] with 1.2052 [assessing age modification without adjustment for gender] 1.6495/1.2052 = 1.37. Relative to the 'crude', the adjusted is 37% more than the 'crude' here. We have evidence that age modification is confounded by gender. This confounding is identified by comparing the two estimates and is identified by noting there is statistical significance with the adjusted but there is no statistical significance with the 'crude'.