Statistical Methods in Epidemiology
Lab 8 - Solutions.
Matched Case-Control Studies

I. Matched studies

i) Endometrial cancer, 1:1 matched sets

1.

```
. mhodds d est set

Mantel-Haenszel estimate of the odds ratio
Comparing est==2 vs. est==1, controlling for set

note: only 32 of the 63 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       9.666667      21.13        0.0000         2.944702  31.733072
    ------------------------------------------------------------------

. sort set d est

. by set: gen new1=est[_n+1]-est
(63 missing values generated)

. li set est  d new1 in 1/40

     +----------------------+
     | set   est   d   new1 |
     |----------------------|
  1. |   1     1   0      1 |
  2. |   1     2   1      . |
  3. |   2     2   0      0 |
  4. |   2     2   1      . |
  5. |   3     2   0      0 |
     |----------------------|
  6. |   3     2   1      . |
  7. |   4     2   0      0 |
  8. |   4     2   1      . |
  9. |   5     2   0      0 |
 10. |   5     2   1      . |
     |----------------------|
 11. |   6     1   0      1 |
 12. |   6     2   1      . |
 13. |   7     1   0      1 |
 14. |   7     2   1      . |
 15. |   8     1   0      1 |
     |----------------------|
 16. |   8     2   1      . |
 17. |   9     1   0      0 |
 18. |   9     1   1      . |
 19. |  10     1   0      1 |
 20. |  10     2   1      . |
     |----------------------|
 21. |  11     1   0      1 |
 22. |  11     2   1      . |
 23. |  12     2   0      0 |
 24. |  12     2   1      . |
 25. |  13     2   0      0 |
```

```
     |----------------------|
 26. |  13     2   1      . |
 27. |  14     1   0      1 |
 28. |  14     2   1      . |
 29. |  15     1   0      1 |
 30. |  15     2   1      . |
     |----------------------|
 31. |  16     1   0      1 |
 32. |  16     2   1      . |
 33. |  17     2   0     -1 |
 34. |  17     1   1      . |
 35. |  18     2   0      0 |
     |----------------------|
 36. |  18     2   1      . |
 37. |  19     2   0      0 |
 38. |  19     2   1      . |
 39. |  20     2   0      0 |
 40. |  20     2   1      . |
     +----------------------+

. count if new1==1
   29
. count if new1==-1
    3
. di 29/3
9.6666667
```

2.

```
. mhodds d age3 set, compare(2,1)

Mantel-Haenszel estimate of the odds ratio
Comparing age3==2 vs. age3==1, controlling for set

note: only 0 of the 49 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
              .          .            .                 .          .
    ------------------------------------------------------------------

. mhodds d age3 set, compare(3,1)

Mantel-Haenszel estimate of the odds ratio
Comparing age3==3 vs. age3==1, controlling for set

note: only 0 of the 27 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
              .          .            .                 .          .
    ------------------------------------------------------------------
```

3.

```
. mhodds d est set, by(age3)

Mantel-Haenszel estimate of the odds ratio
Comparing est==2 vs. est==1, controlling for set
by age3

note: only 32 of the 63 strata formed in this analysis contribute
      information about the effect of the explanatory variable
```

```
------------------------------------------------------------------------------
  age3 | Odds Ratio        chi2(1)         P>chi2       [95% Conf. Interval]
-------+----------------------------------------------------------------------
     1 |   6.000000           3.57         0.0588         0.72235   49.83724
     2 |  15.000000          12.25         0.0005         1.98141  113.55557
     3 |   8.000000           5.44         0.0196         1.00059   63.96252
------------------------------------------------------------------------------

    Mantel-Haenszel estimate controlling for set and age3
    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       9.666667      21.13        0.0000         2.944702  31.733072
    ------------------------------------------------------------------

Test of homogeneity of ORs (approx): chi2(2)   =    0.41
                                     Pr>chi2   =  0.8128
```

The χ^2 test for heterogeneity of the age specific odds ratios does not provide evidence for a modifying effect of age. That test is in fact a Pearson $\chi 2$ performed on a 2x3 table of the age group vs. exposure. Each cell includes either the number of exposed cases and unexposed controls, or the number of unexposed cases and exposed controls in a specific age group. We can construct such a table and perform the chi-squared test of the null hypothesis of no association. Observe that it is exactly the same test as the one produced above.

```
. tab est age3 if new1==1|new1==-1, chi2

           |               age3
       est |         1          2          3 |     Total
-----------+---------------------------------+----------
         1 |         6         15          8 |        29
         2 |         1          1          1 |         3
-----------+---------------------------------+----------
     Total |         7         16          9 |        32

          Pearson chi2(2) =   0.4145   Pr = 0.813
```

4.

```
. mhodds d est set hyp

Mantel-Haenszel estimate of the odds ratio
Comparing est==2 vs. est==1, controlling for set and hyp

note: only 18 of the 94 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
      17.000000      14.22        0.0002         2.262402 127.740345
    ------------------------------------------------------------------
```

ii) Endometrial cancer, 1:4 matched sets

1.

```
. mhodds d est set

Mantel-Haenszel estimate of the odds ratio
Comparing est==1 vs. est==0, controlling for set

note: only 58 of the 63 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       8.461538      31.16        0.0000         3.437773  20.826746
    ------------------------------------------------------------------
```

2.

```
. mhodds d age3 set, compare(2,1)

Mantel-Haenszel estimate of the odds ratio
Comparing age3==2 vs. age3==1, controlling for set

note: only 0 of the 49 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
              .          .            .                 .          .
    ------------------------------------------------------------------

. mhodds d age3 set, compare(3,1)

Mantel-Haenszel estimate of the odds ratio
Comparing age3==3 vs. age3==1, controlling for set

note: only 0 of the 27 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
              .          .            .                 .          .
    ------------------------------------------------------------------
```

3.

```
. mhodds d est set, by(age3)

Mantel-Haenszel estimate of the odds ratio
Comparing est==1 vs. est==0, controlling for set
by age3

note: only 58 of the 63 strata formed in this analysis contribute
      information about the effect of the explanatory variable

-------------------------------------------------------------------------
  age3 | Odds Ratio        chi2(1)         P>chi2       [95% Conf. Interval]
-------+-----------------------------------------------------------------
     1 |   3.800000           3.38         0.0660         0.82165   17.57438
     2 |  10.666667          18.69         0.0000         2.78773   40.81376
     3 |  13.500000           9.77         0.0018         1.59832  114.02620
-------------------------------------------------------------------------
```

```
   Mantel-Haenszel estimate controlling for set and age3
    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       8.461538      31.16        0.0000         3.437773  20.826746
    ------------------------------------------------------------------

Test of homogeneity of ORs (approx): chi2(2)   =    1.41
                                     Pr>chi2   =  0.4943
```

4.

```
. mhodds d est set hyp

Mantel-Haenszel estimate of the odds ratio
Comparing est==1 vs. est==0, controlling for set and hyp

note: only 39 of the 114 strata formed in this analysis contribute
      information about the effect of the explanatory variable

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
      29.222222      33.47        0.0000         4.984146 171.330903
    ------------------------------------------------------------------
```

II. Conditional logistic regression

1.

```
. xi: clogit d i.est, group(set) or
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)

Iteration 0:   log likelihood = -31.455999
Iteration 1:   log likelihood = -31.443712
Iteration 2:   log likelihood = -31.443696
Iteration 3:   log likelihood = -31.443696

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(1)      =   24.45
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.443696                       Pseudo R2       =  0.2799


------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |   9.666667   5.862625     3.74   0.000     2.944702    31.73307
------------------------------------------------------------------------------
```

2.

```
. xi: clogit d i.age3, group(set) or
i.age3            _Iage3_1-3          (naturally coded; _Iage3_1 omitted)
note: _Iage3_2 omitted due to no within-group variance.
note: _Iage3_3 omitted due to no within-group variance.

Iteration 0:   log likelihood = -43.668272

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(0)      =    0.00
                                                  Prob > chi2     =       .
Log likelihood = -43.668272                       Pseudo R2       =  0.0000
```

```
------------------------------------------------------------------------------
        d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
------------------------------------------------------------------------------
```

3.

```
. xi: clogit d i.est*i.age3, group(set) or
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)
i.age3            _Iage3_1-3          (naturally coded; _Iage3_1 omitted)
i.est*i.age3      _IestXage_#_#       (coded as above)
note: _Iage3_2 omitted due to no within-group variance.
note: _Iage3_3 omitted due to no within-group variance.

Iteration 0:   log likelihood = -31.545327
Iteration 1:   log likelihood = -31.249157
Iteration 2:   log likelihood = -31.238555
Iteration 3:   log likelihood = -31.238532
Iteration 4:   log likelihood = -31.238532

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(3)      =   24.86
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.238532                       Pseudo R2       =  0.2846

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |          6   6.480741     1.66   0.097     .7223514    49.83724
_IestXage2|        2.5   3.736085     0.61   0.540     .1336212    46.77401
_IestXage3|   1.333333   2.018434     0.19   0.849      .068607    25.91248
------------------------------------------------------------------------------

. est store A

. xi: clogit d i.est i.age3, group(set) or
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)
i.age3            _Iage3_1-3          (naturally coded; _Iage3_1 omitted)
note: _Iage3_2 omitted due to no within-group variance.
note: _Iage3_3 omitted due to no within-group variance.

Iteration 0:   log likelihood = -31.455999
Iteration 1:   log likelihood = -31.443712
Iteration 2:   log likelihood = -31.443696
Iteration 3:   log likelihood = -31.443696

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(1)      =   24.45
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.443696                       Pseudo R2       =  0.2799

------------------------------------------------------------------------------
       d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------+--------------------------------------------------------------------
 _Iest_2 |   9.666667   5.862625     3.74   0.000     2.944702    31.73307
------------------------------------------------------------------------------

. lrtest A
likelihood-ratio test                                 LR chi2(2)  =   0.41
(Assumption: . nested in A)                           Prob > chi2 = 0.8145
```

The effect of estrogen does not seem to differ statistically significantly across age groups as indicated by the p-value of 0.82.

OR using a Wald test,

```
. xi: clogit d i.est*i.age3, group(set) or nolog

i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)
i.age3            _Iage3_1-3          (naturally coded; _Iage3_1 omitted)
i.est*i.age3      _IestXage_#_#       (coded as above)
note: _Iage3_2 omitted due to no within-group variance.
note: _Iage3_3 omitted due to no within-group variance.

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(3)      =   24.86
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.238532                       Pseudo R2       =  0.2846

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |          6   6.480741     1.66   0.097     .7223514    49.83724
_IestXage2|        2.5   3.736085     0.61   0.540     .1336212    46.77401
_IestXage3|   1.333333   2.018434     0.19   0.849      .068607    25.91248
------------------------------------------------------------------------------

. testparm _IestXage_2_*

 ( 1)  [d]_IestXage_2_2 = 0
 ( 2)  [d]_IestXage_2_3 = 0

           chi2(  2) =    0.40
         Prob > chi2 =    0.8197
```

4.

```
. xi: clogit d i.est*age, group(set) or nolog
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)
i.est*age         _IestXage_#         (coded as above)

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(3)      =   25.05
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.143046                       Pseudo R2       =  0.2868

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |   2.347916   15.05843     0.13   0.894     8.16e-06    675908.9
      age |   .7123318   .3134671    -0.77   0.441       .30068    1.687564
 _IestXa_2|   1.020938   .0921052     0.23   0.818     .8554743    1.218404
------------------------------------------------------------------------------

. est store A

. xi: clogit d i.est age, group(set) or nolog
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(2)      =   25.00
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.169601                       Pseudo R2       =  0.2862

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |   10.24897   6.391759     3.73   0.000     3.018766    34.79614
      age |   .7406473   .2988585    -0.74   0.457     .3358483    1.633352
------------------------------------------------------------------------------

. lrtest A

likelihood-ratio test                                 LR chi2(1)  =   0.05
(Assumption: . nested in A)                           Prob > chi2 = 0.8177
```

5.

```
. xi: clogit d i.est i.hyp, group(set) or nolog
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)
i.hyp             _Ihyp_1-2           (naturally coded; _Ihyp_1 omitted)

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(2)      =   24.56
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.390231                       Pseudo R2       =  0.2812


------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |   10.04889   6.231033     3.72   0.000     2.980659     33.8785
  _Ihyp_2 |   .8669705   .3790934    -0.33   0.744     .3679647     2.04269
------------------------------------------------------------------------------
```

The effect is increased after adjusting for hypertension. However, this change is not statistically significant as indicated by the Wald test on the hypertension term because as you can see the SE of the estrogen estimate has increased also.

6. If we want to check for an interaction this is done with a LR-test as follows:

```
. xi: clogit d i.est*i.hyp, group(set) or nolog
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)
i.hyp             _Ihyp_1-2           (naturally coded; _Ihyp_1 omitted)
i.est*i.hyp       _IestXhyp_#_#       (coded as above)

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(3)      =   24.73
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.304703                       Pseudo R2       =  0.2831

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |   11.53888   8.259096     3.42   0.001     2.837301    46.92689
  _Ihyp_2 |   1.249836   1.218767     0.23   0.819     .1848463    8.450749
_IestXhyp2|   .6406665   .6842618    -0.42   0.677     .0789785    5.197028
------------------------------------------------------------------------------

. est store A

. xi: clogit d i.est i.hyp, group(set) or nolog
i.est             _Iest_1-2           (naturally coded; _Iest_1 omitted)
i.hyp             _Ihyp_1-2           (naturally coded; _Ihyp_1 omitted)

Conditional (fixed-effects) logistic regression   Number of obs   =     126
                                                  LR chi2(2)      =   24.56
                                                  Prob > chi2     =  0.0000
Log likelihood = -31.390231                       Pseudo R2       =  0.2812

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_2 |   10.04889   6.231033     3.72   0.000     2.980659     33.8785
  _Ihyp_2 |   .8669705   .3790934    -0.33   0.744     .3679647     2.04269
------------------------------------------------------------------------------
```

```
. lrtest A

likelihood-ratio test                                 LR chi2(1)  =   0.17
(Assumption: . nested in A)                           Prob > chi2 = 0.6792
```

As you can see, the effect of abnormal level of estrogen on the odds of having endometrial cancer is not statistically significantly different (p-value = 0.68) across levels of hypertension, compared to similar people with the same level of hypertension but with normal level of estrogen.

7.

```
. use "C:\SME_2006-2007\PRACTICALS\Lab5\bdendo.dta"
(1:4 Endometrial Cancer Study)

. li set d est age in 1/10

     +----------------------+
     | set   d   est   age |
     |----------------------|
  1. |   1   1     1    74 |
  2. |   1   0     0    75 |
  3. |   1   0     0    74 |
  4. |   1   0     1    75 |
  5. |   1   0     0    74 |
     |----------------------|
  6. |   2   1     1    67 |
  7. |   2   0     1    68 |
  8. |   2   0     0    67 |
  9. |   2   0     1    67 |
 10. |   2   0     1    67 |
     +----------------------+
```

7(1.)

```
. xi: clogit d est, group(set)

Iteration 0:   log likelihood = -83.721884
Iteration 1:   log likelihood =  -83.72159
Iteration 2:   log likelihood =  -83.72159

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(1)      =   35.35
                                                  Prob > chi2     =  0.0000
Log likelihood =  -83.72159                       Pseudo R2       =  0.1743

------------------------------------------------------------------------------
        d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
      est |   2.073761   .4208245     4.93   0.000      1.24896    2.898562
------------------------------------------------------------------------------
```

7(2.)

```
. xi: clogit d i.age3, group(set)
i.age3            _Iage3_1-3          (naturally coded; _Iage3_1 omitted)
note: _Iage3_2 omitted due to no within-group variance.
note: _Iage3_3 omitted due to no within-group variance.

Iteration 0:   log likelihood = -101.39459

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(0)      =    0.00
                                                  Prob > chi2     =       .
Log likelihood = -101.39459                       Pseudo R2       =  0.0000
```

```
------------------------------------------------------------------------------
        d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
------------------------------------------------------------------------------
```

7(3.)

```
. xi: clogit d i.est*i.age3, group(set) nolog
i.est             _Iest_0-1           (naturally coded; _Iest_0 omitted)
i.age3            _Iage3_1-3          (naturally coded; _Iage3_1 omitted)
i.est*i.age3      _IestXage_#_#       (coded as above)
note: _Iage3_2 omitted due to no within-group variance.
note: _Iage3_3 omitted due to no within-group variance.

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(3)      =   36.03
                                                  Prob > chi2     =  0.0000
Log likelihood = -83.380155                       Pseudo R2       =  0.1777

------------------------------------------------------------------------------
        d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_1 |   1.430827   .8256893     1.73   0.083    -.1874939    3.049149
_IestXage2|    .847401    1.03377     0.82   0.412    -1.178751    2.873553
_IestXage3|   .7801409   1.154229     0.68   0.499    -1.482107    3.042389
------------------------------------------------------------------------------

. est store A

. xi: clogit d i.est i.age3, group(set) nolog
i.est             _Iest_0-1           (naturally coded; _Iest_0 omitted)
i.age3            _Iage3_1-3          (naturally coded; _Iage3_1 omitted)
note: _Iage3_2 omitted due to no within-group variance.
note: _Iage3_3 omitted due to no within-group variance.

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(1)      =   35.35
                                                  Prob > chi2     =  0.0000
Log likelihood =  -83.72159                       Pseudo R2       =  0.1743

------------------------------------------------------------------------------
        d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_1 |   2.073761   .4208245     4.93   0.000      1.24896    2.898562
------------------------------------------------------------------------------

. lrtest A

likelihood-ratio test                                 LR chi2(2)  =   0.68
(Assumption: . nested in A)                           Prob > chi2 = 0.7107
```

7(4.)

```
. xi: clogit d i.est*age, group(set) or nolog
i.est             _Iest_0-1           (naturally coded; _Iest_0 omitted)
i.est*age         _IestXage_#         (coded as above)
Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(3)      =   36.86
                                                  Prob > chi2     =  0.0000
Log likelihood = -82.963318                       Pseudo R2       =  0.1818

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_1 |   2.076744   9.452452     0.16   0.872     .0002774    15548.54
      age |   .7343274   .1852424    -1.22   0.221      .447883    1.203968
 _IestXag1|    1.01918   .0645293     0.30   0.764     .9002377    1.153838
------------------------------------------------------------------------------
```

```
. est store A

. xi: clogit d i.est age, group(set) or nolog
i.est             _Iest_0-1           (naturally coded; _Iest_0 omitted)

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(2)      =   36.77
                                                  Prob > chi2     =  0.0000
Log likelihood = -83.008369                       Pseudo R2       =  0.1813

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_1 |   8.127874   3.462022     4.92   0.000     3.527057    18.73016
      age |   .7502839   .1813111    -1.19   0.234     .4672261    1.204825
------------------------------------------------------------------------------

. lrtest A

likelihood-ratio test                                 LR chi2(1)  =   0.09
(Assumption: . nested in A)                           Prob > chi2 = 0.7640
```

7(5.)

```
. xi: clogit d i.est i.hyp, group(set) or nolog
i.est             _Iest_0-1           (naturally coded; _Iest_0 omitted)
i.hyp             _Ihyp_1-2           (naturally coded; _Ihyp_1 omitted)

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(2)      =   35.35
                                                  Prob > chi2     =  0.0000
Log likelihood = -83.721527                       Pseudo R2       =  0.1743

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_1 |   7.961153   3.399496     4.86   0.000     3.447498    18.38434
  _Ihyp_2 |   .9962937   .3294173    -0.01   0.991     .5211272    1.904719
------------------------------------------------------------------------------
```

7(6.)

```
. xi: clogit d i.est*i.hyp, group(set) nolog
i.est             _Iest_0-1           (naturally coded; _Iest_0 omitted)
i.hyp             _Ihyp_1-2           (naturally coded; _Ihyp_1 omitted)
i.est*i.hyp       _IestXhyp_#_#       (coded as above)

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(3)      =   35.87
                                                  Prob > chi2     =  0.0000
Log likelihood = -83.458347                       Pseudo R2       =  0.1769

------------------------------------------------------------------------------
        d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_1 |   2.250869   .5049042     4.46   0.000     1.261275    3.240463
  _Ihyp_2 |   .6194237   .8784346     0.71   0.481    -1.102276    2.341124
_IestXhyp2|  -.6888568   .9131123    -0.75   0.451    -2.478524     1.10081
------------------------------------------------------------------------------

. est store A

. xi: clogit d i.est i.hyp, group(set) nolog
i.est             _Iest_0-1           (naturally coded; _Iest_0 omitted)
i.hyp             _Ihyp_1-2           (naturally coded; _Ihyp_1 omitted)

Conditional (fixed-effects) logistic regression   Number of obs   =     315
                                                  LR chi2(2)      =   35.35
```

```
                                                  Prob > chi2     =  0.0000
Log likelihood = -83.721527                       Pseudo R2       =  0.1743

------------------------------------------------------------------------------
        d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Iest_1 |   2.074574   .4270105     4.86   0.000     1.237649    2.911499
  _Ihyp_2 |  -.0037132   .3306428    -0.01   0.991    -.6517611    .6443347
------------------------------------------------------------------------------


. lrtest A

likelihood-ratio test                                 LR chi2(1)  =   0.53
(Assumption: . nested in A)                           Prob > chi2 = 0.4681
```

III. Conditional vs. unconditional logistic

1.

```
. xi: logit d i.bcg i.age, or nolog
i.bcg             _Ibcg_0-1           (naturally coded; _Ibcg_0 omitted)
i.age             _Iage_0-6           (naturally coded; _Iage_0 omitted)

Logit estimates                                   Number of obs   =       1370
                                                  LR chi2(7)      =     106.83
                                                  Prob > chi2     =     0.0000
Log likelihood = -632.13736                       Pseudo R2       =     0.0779

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Ibcg_1 |   .3312718   .0638478    -5.73   0.000     .2270533    .4833272
  _Iage_1 |   1.339224   .3513803     1.11   0.266     .8007897    2.239692
  _Iage_2 |   1.237187   .3740553     0.70   0.481     .6840382    2.237641
  _Iage_3 |   1.624545   .5225785     1.51   0.131     .8648085    3.051712
  _Iage_4 |   3.401921   1.002921     4.15   0.000     1.908892    6.062713
  _Iage_5 |   2.422219   .5959147     3.60   0.000     1.495549     3.92307
  _Iage_6 |   2.084825   .4998738     3.06   0.002     1.303104    3.335495
------------------------------------------------------------------------------
```

2.

```
. gen grp=1

. xi: clogit d i.bcg i.age, group(grp) or
i.bcg             _Ibcg_0-1           (naturally coded; _Ibcg_0 omitted)
i.age             _Iage_0-6           (naturally coded; _Iage_0 omitted)
note: multiple positive outcomes within groups encountered.
Iteration 0:   log likelihood = -628.56209
Iteration 1:   log likelihood = -628.56207
Conditional (fixed-effects) logistic regression   Number of obs   =    1370
                                                  LR chi2(7)      =  106.75
                                                  Prob > chi2     =  0.0000
Log likelihood = -628.56207                       Pseudo R2       =  0.0783

------------------------------------------------------------------------------
        d | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
  _Ibcg_1 |   .3315068   .0638743    -5.73   0.000     .2272396    .4836163
  _Iage_1 |   1.338976   .3512043     1.11   0.266     .8007714    2.238913
  _Iage_2 |   1.237001   .3738914     0.70   0.482     .6840517    2.236924
  _Iage_3 |   1.624024   .5222335     1.51   0.132     .8647159     3.05008
  _Iage_4 |   3.398322   1.001408     4.15   0.000      1.90737    6.054721
  _Iage_5 |    2.42044   .5952393     3.59   0.000     1.494738    3.919435
  _Iage_6 |    2.08362   .4993951     3.06   0.002     1.302583    3.332972
------------------------------------------------------------------------------
```

Here grp is the variable declaring the matching set and it takes the value 1 for all subjects, meaning that the effect of matching is neutral. This could be the case if (a) the matching variable is associated neither with the disease nor with the exposure, or (b) the matching variable is associated only with the disease. Under such circumstances both approaches (conditional & unconditional) lead to approximately identical results and matching can be ignored in the analysis.