```
// Code in stata to evaluate the ability of SucCoA and Cobalamin to predict HSIL
// Date: June 1st, 2022. Author: Sergio Serrano-Villar
* Outcome variables: bHSIL (biopsy-proven HSIL), cHSIL (biopsy-proven + cytologic HSIL)
* Predictors: cyto_abnormal (abnormal anal cytology), succ (Succinyl CoA), cob (Cobalamin)
*Dataset: "db_metab"

use "db_metab", clear

////////////////////////////////////////////////////////////////////////////////
/// Diagnostic accuraccy of anal cytology in the Discovery cohort (filter: "deriv")
////////////////////////////////////////////////////////////////////////////////

*Discriminative ability of the reference test ("cyto_abnormal")
tab cyto_abnormal bHSIL if deriv == 1 , col
/*
  Abnormal |
  cytology |  Biopsy-proven HSIL
    yes/no |  No bHSIL      bHSIL |     Total
-----------+----------------------+----------
    Normal |        29          5 |        34
           |     34.12       8.77 |     23.94
-----------+----------------------+----------
  Abnormal |        56         52 |       108
           |     65.88      91.23 |     76.06
-----------+----------------------+----------
     Total |        85         57 |       142
           |    100.00     100.00 |    100.00
 */

diagt bHSIL cyto_abnormal  if deriv == 1
/*
           |   Abnormal cytology
Biopsy-pro |        yes/no
  ven HSIL |      Pos.       Neg. |     Total
-----------+----------------------+----------
  Abnormal |        52          5 |        57
    Normal |        56         29 |        85
-----------+----------------------+----------
     Total |       108         34 |       142
True abnormal diagnosis defined as bHSIL = 1 (labelled bHSIL)


                                                  [95% Confidence Interval]
-------------------------------------------------------------------------
Prevalence                          Pr(A)       40%       32%      48.7%
-------------------------------------------------------------------------
Sensitivity                         Pr(+|A)     91.2%     80.7%     97.1%
Specificity                         Pr(-|N)     34.1%     24.2%     45.2%
ROC area                (Sens. +5 Spec.)/2       .627      .564       .69
-------------------------------------------------------------------------
Likelihood ratio (+)     Pr(+|A)/Pr(+|N)       1.38      1.16      1.65
Likelihood ratio (-)     Pr(-|A)/Pr(-|N)       .257      .106      .625
Odds ratio                   LR(+)/LR(-)       5.39      1.99      14.4
Positive predictive value        Pr(A|+)     48.1%     38.4%       58%
Negative predictive value        Pr(N|-)     85.3%     68.9%       95%
-------------------------------------------------------------------------
*/

logistic bHSIL cyto_abnormal if deriv == 1

/*
Logistic regression                                     Number of obs =    142
                                                        LR chi2(1)    =  13.33
                                                        Prob > chi2   = 0.0003
Log likelihood = -88.983294                             Pseudo R2     = 0.0697

------------------------------------------------------------------------------
```

```
        bHSIL | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
--------------+----------------------------------------------------------------
cyto_abnormal |   5.385714   2.806627     3.23   0.001     1.939373    14.95634
        _cons |   .1724138   .0834887    -3.63   0.000     .0667408    .4454028
------------------------------------------------------------------------------
*/

estat classification //Correctly classified  59.86%


////////////////////////////////////////////////////////////////////////////////
/// Diagnostic accuraccy of SuCCoaA and Cobalamin in the Discovery cohort (filter:
"deriv")
////////////////////////////////////////////////////////////////////////////////

**Metabolite mean and median values according to the dysplasia degree
tabstat succoa_ugml cobal_ugml if deriv == 1, statistics( mean median ) by(bAINcat)

**Metabolite mean and median values according to the presence of bHSIL
tabstat succoa_ugml cobal_ugml if deriv == 1, statistics( mean median N) by(bHSIL)

**Discriminative ability as log2 transformed variables
gen log2suc = ln(succ)/ln(2)
gen log2cob = ln(cob)/ln(2)

**For biopsy-proven HSIL
logistic bHSIL log2suc if deriv == 1
lroc, all msize(vtiny) legend(region(fcolor(white)))

logistic bHSIL log2cob if deriv == 1
lroc, all msize(vtiny) legend(region(fcolor(white)))

**For biopsy-proven + cytologic HSIL
logistic cHSIL log2suc if deriv == 1
lroc, all msize(vtiny) legend(region(fcolor(white)))

logistic cHSIL log2cob if deriv == 1
lroc, all msize(vtiny) legend(region(fcolor(white)))

**Selection of most discriminative cutoffs
*SucCoA
logistic bHSIL  succoa_ugml if deriv == 1
lroc, all msize(vtiny) legend(region(fcolor(white)))
roctab bHSIL suc, binomial detail graph specificity summary
cutpt bHSIL succoa_ugml if deriv == 1, noadjust
gen suc_cat2 = suc >=50  //~50 the best cutoff

*Modelo con el succinilCoA categorizado por este punto de corte
logistic bHSIL suc_cat2 if deriv == 1

/*
Logistic regression                                     Number of obs =    146
                                                        LR chi2(1)    =  90.04
                                                        Prob > chi2   = 0.0000
Log likelihood = -53.076599                             Pseudo R2     = 0.4589

------------------------------------------------------------------------------
       bHSIL | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    suc_cat2 |   50.99999   26.32119     7.62   0.000     18.54656    140.2416
       _cons |   .0909091   .0358883    -6.07   0.000     .0419351    .1970775
------------------------------------------------------------------------------
*/

*Discrimination
diagt bHSIL suc_cat2 if deriv == 1
/*
Biopsy-pro |       suc_cat2
```

```
  ven HSIL |      Pos.       Neg. |     Total
-----------+----------------------+----------
  Abnormal |        51          7 |        58
    Normal |        11         77 |        88
-----------+----------------------+----------
     Total |        62         84 |       146
True abnormal diagnosis defined as bHSIL = 1 (labelled bHSIL)


                                                      [95% Confidence Interval]
-------------------------------------------------------------------------------
Prevalence                          Pr(A)       40%       32%      48.1%
-------------------------------------------------------------------------------
Sensitivity                         Pr(+|A)     87.9%     76.7%       95%
Specificity                         Pr(-|N)     87.5%     78.7%     93.6%
ROC area                (Sens. + Spec.)/2       .877      .822      .932
-------------------------------------------------------------------------------
Likelihood ratio (+)      Pr(+|A)/Pr(+|N)       7.03      4.01      12.3
Likelihood ratio (-)      Pr(-|A)/Pr(-|N)       .138     .0686      .278
Odds ratio                    LR(+)/LR(-)         51      18.8       138
Positive predictive value         Pr(A|+)     82.3%     70.5%     90.8%
Negative predictive value         Pr(N|-)     91.7%     83.6%     96.6%
-------------------------------------------------------------------------------


*/

logistic bHSIL suc_cat2 if deriv == 1
estat classification  ///Correctly classified  87.67%

*Cobalamina
logistic bHSIL  cob if deriv == 1
lroc, all msize(vtiny) legend(region(fcolor(white)))
roctab bHSIL cob, binomial detail graph specificity summary  //listado de puntos de corte
cutpt bHSIL cob if deriv == 1, noadjust
bootstrap e(cutpoint) if deriv == 1, rep(100): cutpt bHSIL cob, noadjust
gen cob_cat2 = cob >=150 //~50 the best cutoff

logistic bHSIL cob_cat2 if deriv == 1

/*
Logistic regression                                     Number of obs =    146
                                                        LR chi2(1)    =  83.29
                                                        Prob > chi2   = 0.0000
Log likelihood = -56.449088                             Pseudo R2     = 0.4245

------------------------------------------------------------------------------
       bHSIL | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    cob_cat2 |   44.35714   22.62976     7.43   0.000     16.31941    120.5654
       _cons |   .1481481   .0458252    -6.17   0.000     .0807983    .2716377
------------------------------------------------------------------------------
*/

*Discrimination
diagt bHSIL cob_cat if deriv == 1
/*
Biopsy-pro |       cob_cat2
  ven HSIL |      Pos.       Neg. |     Total
-----------+----------------------+----------
  Abnormal |        46         12 |        58
    Normal |         7         81 |        88
-----------+----------------------+----------
     Total |        53         93 |       146
True abnormal diagnosis defined as bHSIL = 1 (labelled bHSIL)


                                                      [95% Confidence Interval]
```

```
------------------------------------------------------------------------------
Prevalence                        Pr(A)       40%       32%      48.1%
------------------------------------------------------------------------------
Sensitivity                      Pr(+|A)     79.3%     66.6%     88.8%
Specificity                      Pr(-|N)       92%     84.3%     96.7%
ROC area               (Sens. + Spec.)/2      .857      .797      .917
------------------------------------------------------------------------------
Likelihood ratio (+)     Pr(+|A)/Pr(+|N)      9.97      4.84      20.5
Likelihood ratio (-)     Pr(-|A)/Pr(-|N)      .225      .135      .373
Odds ratio                   LR(+)/LR(-)      44.4      16.5       119
Positive predictive value        Pr(A|+)     86.8%     74.7%     94.5%
Negative predictive value        Pr(N|-)     87.1%     78.5%     93.2%
------------------------------------------------------------------------------
*/

logistic bHSIL cob_cat2 if deriv == 1
estat classification  // Correctly classified  86.99%

**Does each metabolite independently predict bHSIL? Multivariate models
*Outcome: biopsy-proven HSIL
logistic bHSIL suc_cat2 cob_cat2 if deriv == 1
*Outcome: biopsy-proven + cytologic HSIL
logistic cHSIL suc_cat2 cob_cat2 if deriv == 1

*Generate composite variable with 2 categories: any metabolite positive vs. both
metabolites negatives
gen supertest_2cat = .
replace supertest_2cat = 0 if cob_cat2 == 0 & suc_cat2 == 0
replace supertest_2cat = 1 if cob_cat2 == 1 | suc_cat2 == 1

*Generate composite variables with 3 cats (metabolites --/-+/++)
gen supertest_3cat = .
replace supertest_3cat = 0 if cob_cat2 == 0 & suc_cat2 == 0
replace supertest_3cat = 1 if (cob_cat2 == 0 & suc_cat2 == 1) | (cob_cat2 == 1 & suc_cat2
== 0)
replace supertest_3cat = 2 if cob_cat2 == 1 & suc_cat2 == 1


**PREDICTIVE ABILITY OF CATEGORIZED SUC COA AND COBALAMINE
*Discovery cohort
logistic bHSIL supertest_2cat if deriv == 1
/*

Logistic regression                                     Number of obs =    146
                                                        LR chi2(1)    = 101.52
                                                        Prob > chi2   = 0.0000
Log likelihood = -47.333408                             Pseudo R2     = 0.5175

------------------------------------------------------------------------------
         bHSIL | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
---------------+--------------------------------------------------------------
supertest_2cat |        126   97.13007     6.27   0.000     27.81005    570.8727
         _cons |   .0277778   .0199128    -5.00   0.000     .0068156    .1132114
------------------------------------------------------------------------------
*/

*Internal validation and calibration
*Need the user-written package bsvalidation
bsvalidation, rseed(999) adjust (heuristic) group(2) graph

/*
Apparent performance
------------------------------------------------------------
                                  [95% Conf. Interval]
Overall:
        Brier scaled (%) =  58.8

Discrimination:
```

```
            C-Statistic =  0.892       0.845   0.939

Calibration:
               E:O ratio =  1.000
                    CITL =  0.000      -0.517   0.517
                   Slope =  1.000       0.688   1.312
-------------------------------------------------------------

Bootstrap performance (Optimism adjusted)
Number of replications:  50
-------------------------------------------------------------
                                    [Bootstrap 95% CI]
Overall:
        Brier scaled (%) =  58.2

Discrimination:
            C-Statistic =  0.893       0.846   0.934

Calibration:
               E:O ratio =  1.083       0.858   2.028
                    CITL = -0.047      -0.675   0.544
                   Slope =  0.935       0.000   1.316
-------------------------------------------------------------

Shrinkage factors
-------------------------------------------------------------
     Heuristic Shrinkage =  0.990
     Bootstrap shrinkage =  0.935
-------------------------------------------------------------
Model adjusted by heuristic shrinkage
------------------------------------------------------------------------------
         bHSIL | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
---------------+--------------------------------------------------------------
bHSIL          |
supertest_2cat |   4.787919   .7631648     6.27   0.000     3.292144    6.283695
         _cons |   -3.37432   .2347454   -14.37   0.000    -3.834412   -2.914227
------------------------------------------------------------------------------
*/

logistic cHSIL supertest_2cat if deriv == 1

*Discrimantion
diagt bHSIL supertest_2cat if deriv == 1
/*
Biopsy-pro |    supertest_2cat
  ven HSIL |      Pos.       Neg. |     Total
-----------+----------------------+----------
  Abnormal |        56          2 |        58
    Normal |        16         72 |        88
-----------+----------------------+----------
     Total |        72         74 |       146
True abnormal diagnosis defined as bHSIL = 1 (labelled bHSIL)


                                                  [95% Confidence Interval]
-------------------------------------------------------------------------------
Prevalence                          Pr(A)        40%       32%      48.1%
-------------------------------------------------------------------------------
Sensitivity                         Pr(+|A)     96.6%     88.1%     99.6%
Specificity                         Pr(-|N)     81.8%     72.2%     89.2%
ROC area                (Sens. + Spec.)/2       .892      .845      .939
-------------------------------------------------------------------------------
Likelihood ratio (+)     Pr(+|A)/Pr(+|N)       5.31       3.4      8.29
Likelihood ratio (-)     Pr(-|A)/Pr(-|N)      .0421     .0108      .165
Odds ratio                   LR(+)/LR(-)        126      30.3         .
Positive predictive value        Pr(A|+)     77.8%     66.4%     86.7%
Negative predictive value        Pr(N|-)     97.3%     90.6%     99.7%
-------------------------------------------------------------------------------
```

```
*/

logistic bHSIL supertest_2cat if deriv == 1
estat classification
*Correctly classified                        87.67%


**We repeat the previous tests using the composite variables with 3 cats, supertest_3cat
(metabolites --/-+/++)

logistic bHSIL i.supertest_3cat if deriv == 1
/*
Logistic regression                                     Number of obs =    146
                                                        LR chi2(2)    = 121.46
                                                        Prob > chi2   = 0.0000
Log likelihood = -37.367441                             Pseudo R2     = 0.6191

------------------------------------------------------------------------------
         bHSIL | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
---------------+--------------------------------------------------------------
supertest_3cat |
            1  |   38.57143   31.14471     4.52   0.000     7.924256     187.747
            2  |   737.9995   751.9917     6.48   0.000      100.165     5437.46
               |
         _cons |   .0277778   .0199128    -5.00   0.000     .0068156    .1132114
------------------------------------------------------------------------------
*/
estat classification
logistic bHSIL i.supertest_3cat if deriv == 1
bsvalidation, rseed(111) adjust (heuristic) group(2) graph //Discrimination:  C-Statistic
=  0.948       0.870   1.012

logistic bHSIL i.supertest_3cat if deriv == 1
logistic bHSIL supertest_3cat if deriv == 1
estat classification
/*
Classified + if predicted Pr(D) >= .5
True D defined as bHSIL != 0
--------------------------------------------------
Sensitivity                     Pr( +| D)   70.69%
Specificity                     Pr( -|~D)   97.73%
Positive predictive value       Pr( D| +)   95.35%
Negative predictive value       Pr(~D| -)   83.50%
--------------------------------------------------
False + rate for true ~D        Pr( +|~D)    2.27%
False - rate for true D         Pr( -| D)   29.31%
False + rate for classified +   Pr(~D| +)    4.65%
False - rate for classified -   Pr( D| -)   16.50%
--------------------------------------------------
Correctly classified                        86.99%
--------------------------------------------------
*/

diagt bHSIL supertest_3cat if deriv == 1


**We repeat the previous tests including the continuous variables log2suc and log2cob in a
multivariate model
logistic bHSIL log2cob log2suc if deriv == 1

estat classification
/*
Classified + if predicted Pr(D) >= .5
True D defined as bHSIL != 0
--------------------------------------------------
Sensitivity                     Pr( +| D)   87.93%
Specificity                     Pr( -|~D)   92.05%
Positive predictive value       Pr( D| +)   87.93%
```

```
Negative predictive value       Pr(~D| -)   92.05%
-------------------------------------------------
False + rate for true ~D        Pr( +|~D)    7.95%
False - rate for true D         Pr( -| D)   12.07%
False + rate for classified +   Pr(~D| +)   12.07%
False - rate for classified -   Pr( D| -)    7.95%
-------------------------------------------------
Correctly classified                        90.41%
-------------------------------------------------
*/
lroc

logistic bHSIL log2cob log2suc if deriv == 1
bsvalidation, rseed(111) adjust (heuristic) group(2) graph


////////////////////////////////////////////////////////////////////////////////////////
/// Diagnostic accuraccy of SuCCoaA and Cobalamin in the VALIDATION cohort (filter:
"deriv")
////////////////////////////////////////////////////////////////////////////////////////

**We estimate the bHSIL predicted probabilities in the discovery cohort and compare the
predicted vs. observed bHSIL in the validation cohort.

*For supertest_2cat, Predicted probabilities in the discovery cohort
logistic bHSIL supertest_2cat if deriv == 1
predict p1  //gen predicted probabilities
estat classification
/*
Classified + if predicted Pr(D) >= .5
True D defined as bHSIL != 0
-------------------------------------------------
Sensitivity                     Pr( +| D)   96.55%
Specificity                     Pr( -|~D)   81.82%
Positive predictive value       Pr( D| +)   77.78%
Negative predictive value       Pr(~D| -)   97.30%
-------------------------------------------------
False + rate for true ~D        Pr( +|~D)   18.18%
False - rate for true D         Pr( -| D)    3.45%
False + rate for classified +   Pr(~D| +)   22.22%
False - rate for classified -   Pr( D| -)    2.70%
-------------------------------------------------
Correctly classified                        87.67%
-------------------------------------------------*/

* Probabilities predicted in the model fitted in the discovery cohort vs. observed bHSIL
in the validation cohort
diagt bHSIL p1 if valid == 1
/*

                                                      [95% Confidence Interval]
-------------------------------------------------------------------------------
Prevalence                          Pr(A)       56%       40%      71.5%
-------------------------------------------------------------------------------
Sensitivity                        Pr(+|A)     95.7%     78.1%     99.9%
Specificity                        Pr(-|N)     83.3%     58.6%     96.4%
ROC area                (Sens. + Spec.)/2      .895      .797      .993
-------------------------------------------------------------------------------
Likelihood ratio (+)     Pr(+|A)/Pr(+|N)      5.74      2.04      16.2
Likelihood ratio (-)     Pr(-|A)/Pr(-|N)     .0522     .0076      .359
Odds ratio                   LR(+)/LR(-)       110      12.3         .
Positive predictive value        Pr(A|+)       88%     68.8%     97.5%
Negative predictive value        Pr(N|-)     93.8%     69.8%     99.8%
-------------------------------------------------------------------------------
*/

*Predicted probabilities in the validation cohort
logistic bHSIL supertest_2cat if valid == 1
```

```
estat classification
/*
Classified + if predicted Pr(D) >= .5
True D defined as bHSIL != 0
--------------------------------------------------
Sensitivity                     Pr( +| D)   95.65%
Specificity                     Pr( -|~D)   83.33%
Positive predictive value       Pr( D| +)   88.00%
Negative predictive value       Pr(~D| -)   93.75%
--------------------------------------------------
False + rate for true ~D        Pr( +|~D)   16.67%
False - rate for true D         Pr( -| D)    4.35%
False + rate for classified +   Pr(~D| +)   12.00%
False - rate for classified -   Pr( D| -)    6.25%
--------------------------------------------------
Correctly classified                        90.24%
--------------------------------------------------*/


*For supertest_3cat, Predicted probabilities in the discovery cohort
logistic bHSIL supertest_3cat if deriv == 1
predict p2  //gen predicted probabilities
estat classification
/*
Classified + if predicted Pr(D) >= .5
True D defined as bHSIL != 0
--------------------------------------------------
Sensitivity                     Pr( +| D)   70.69%
Specificity                     Pr( -|~D)   97.73%
Positive predictive value       Pr( D| +)   95.35%
Negative predictive value       Pr(~D| -)   83.50%
--------------------------------------------------
False + rate for true ~D        Pr( +|~D)    2.27%
False - rate for true D         Pr( -| D)   29.31%
False + rate for classified +   Pr(~D| +)    4.65%
False - rate for classified -   Pr( D| -)   16.50%
--------------------------------------------------
Correctly classified                        86.99%
--------------------------------------------------
*/

* Probabilities predicted in the model fitted in the discovery cohort vs. observed bHSIL
in the validation cohort
diagt bHSIL p2 if valid == 1
/*
                                                        [95% Confidence Interval]
-------------------------------------------------------------------------------
Prevalence                          Pr(A)       56%       40%      71.5%
-------------------------------------------------------------------------------
Sensitivity                       Pr(+|A)     65.2%     42.7%     83.6%
Specificity                       Pr(-|N)      100%     81.5%      100%
ROC area               (Sens. + Spec.)/2      .826      .727      .926
-------------------------------------------------------------------------------
Likelihood ratio (+)     Pr(+|A)/Pr(+|N)         .         .         .
Likelihood ratio (-)     Pr(-|A)/Pr(-|N)      .348      .199      .609
Odds ratio                   LR(+)/LR(-)         .      7.73         .
Positive predictive value        Pr(A|+)      100%     78.2%      100%
Negative predictive value        Pr(N|-)     69.2%     48.2%     85.7%
-------------------------------------------------------------------------------
*/

*Predicted probabilities in the validation cohort
logistic bHSIL supertest_3cat if valid == 1
estat classification
/*
Classified + if predicted Pr(D) >= .5
True D defined as bHSIL != 0
--------------------------------------------------
```

```
Sensitivity                     Pr( +| D)   95.65%
Specificity                     Pr( -|~D)   83.33%
Positive predictive value       Pr( D| +)   88.00%
Negative predictive value       Pr(~D| -)   93.75%
-------------------------------------------------
False + rate for true ~D        Pr( +|~D)   16.67%
False - rate for true D         Pr( -| D)    4.35%
False + rate for classified +   Pr(~D| +)   12.00%
False - rate for classified -   Pr( D| -)    6.25%
-------------------------------------------------
Correctly classified                        90.24%
*/


////////////////////////////////////////////////////////////////////////////////////
/// NET RECLASSIFICATION INDEX AND MULTIVARIATE MODELS ADJUSTING FOR POTENTIAL CONFOUNDERS
////////////////////////////////////////////////////////////////////////////////////

**Whole cohort (discovery + validation)
*supertest_3cat
bysort bHSIL: tab cyto_abnormal supertest_3cat
logistic bHSIL supertest_3cat cyto_abnormal age cd4num  smoker hpv_simp
proctitis_chlm_ever
/*
Logistic regression                                     Number of obs =    147
                                                        LR chi2(7)    = 149.91
                                                        Prob > chi2   = 0.0000
Log likelihood = -25.950608                             Pseudo R2     = 0.7428

------------------------------------------------------------------------------
       bHSIL | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
superte~3cat |   107.5365   97.33773     5.17   0.000     18.24243    633.9125
cyto_abnor~l |   5.029541   4.428908     1.83   0.067     .8953198    28.25391
         age |   1.011491   .0304511     0.38   0.704     .9535351     1.07297
      cd4num |    1.00212   .0010035     2.11   0.034     1.000155    1.004089
      smoker |   .3855312   .2031454    -1.81   0.070     .1372592    1.082873
    hpv_simp |   .3266617   .2026331    -1.80   0.071     .0968468    1.101821
proct~m_ever |   .3635935   .3374972    -1.09   0.276     .0589533    2.242457
       _cons |   .0679814   .1555628    -1.17   0.240     .0007666    6.028462
------------------------------------------------------------------------------
*/

*supertest_2cat
bysort bHSIL: tab cyto_abnormal supertest_2cat
logistic bHSIL supertest_2cat cyto_abnormal age cd4num  smoker hpv_simp
proctitis_chlm_ever
/*
c regression                                    Number of obs =    147
                                                        LR chi2(7)    = 130.77
                                                        Prob > chi2   = 0.0000
Log likelihood = -35.520244                             Pseudo R2     = 0.6480

------------------------------------------------------------------------------
              bHSIL | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
--------------------+---------------------------------------------------------
     supertest_2cat |   424.6761    430.349     5.97   0.000     58.27546     3094.78
      cyto_abnormal |   3.798017    2.65189     1.91   0.056     .9665456    14.92422
                age |   1.030652   .0304471     1.02   0.307     .9726715    1.092089
             cd4num |   1.001071   .0009408     1.14   0.255     .9992291    1.002917
             smoker |    .577941   .2242165    -1.41   0.158     .2701802     1.23627
           hpv_simp |   .4524815   .2347818    -1.53   0.126     .1636563    1.251033
proctitis_chlm_ever |   .4499702   .3941761    -0.91   0.362     .0808215    2.505189
              _cons |   .0239087    .053406    -1.67   0.095        .0003    1.905144
------------------------------------------------------------------------------*/
```