MetaDAS: A SAS macro for meta-analysis of diagnostic accuracy studies

User guide

Yemisi Takwoingi and Jon Deeks

Version 1.3 Released 30 July 2010

Please cite this version as: Takwoingi Y, Deeks JJ. MetaDAS: A SAS macro for meta-analysis of diagnostic accuracy studies. User Guide Version 1.3. 2010 July. Available from: http://srdta.cochrane.org/.

Contents

1 Advanced meta-analysis methods

Meta-analysis of diagnostic accuracy studies requires the use of more advanced methods than meta-analysis of intervention studies. This is due to the fact that test performance is typically reported in terms of the two measures sensitivity and specificity and there often exists a negative correlation between them as the threshold for test positivity changes. Also, there is usually considerable heterogeneity between the studies in the meta-analysis. The hierarchical or multilevel modelling approach takes account of both within and between study variations. Two models have been proposed namely the hierarchical summary receiver operating characteristic (HSROC) model (1) and the bivariate random-effects model (2). Both models have been shown to be closely related and are identical in common situations (3).

1.1 HSROC model

The hierarchical summary ROC model is a non linear mixed model in which the implicit threshold θ and diagnostic accuracy α for each study are specified as random effects. This model includes a shape or scale parameter β which enables asymmetry in the SROC by allowing accuracy (lnDOR) to vary with implicit threshold. Estimation of β requires information from more than one study and is therefore assumed to be constant across studies (1), i.e., a fixed effect. Covariates can be included in the model to assess whether threshold, accuracy, and the shape of the SROC (singly or in combination) vary with patient or study characteristics. Each covariate is generally fitted as a fixed effect (4).

1.2 Bivariate model

The bivariate model is a linear mixed model that enables analysis of the effects of sensitivity and specificity. The model assumes a bivariate normal distribution for the logit transforms of the sensitivities and specificities from individual studies within a meta-analysis. Thus, the possibility of a correlation between sensitivity and specificity within studies is explicitly taken into account in the analysis. The bivariate model also incorporates the precision by which sensitivity and specificity have been measured in each study (2).

2 The SAS NLMIXED procedure

This procedure fits nonlinear and generalized linear mixed models using likelihood based methods. It requires a regression equation and declaration of parameters with their initial estimates (starting values). These starting values are required for the iterative process and it is essential to select good ones in order to avoid excessively long computing time and also to facilitate convergence. Proc NLMIXED has the capability to search for initial values over a grid of parameter values. Empirical Bayes predictions of the random effects and estimates of functions of the parameters can be obtained and the delta method is used to estimate their standard error (5). NLMIXED cannot accommodate a large number of random effects (<5) and is limited to only 2 levels (6).

3 MetaDAS

3.1 Overview

MetaDAS is a SAS macro developed to automate the fitting of bivariate and HSROC models for meta-analysis of diagnostic accuracy studies using Proc NLMIXED. In ***MetaDAS***, NLMIXED uses maximum likelihood estimation via adaptive Gaussian quadrature and a dual quasi-Newton optimization algorithm as the default optimizer.

Explanatory variables (covariates) can be added to models to produce separate effects on the summary measures of test accuracy. Also, distributional assumptions of the random effects can be checked and predicted values of sensitivity and specificity, based on empirical Bayes estimates of the random effects, can be obtained for each study in the meta-analysis. The output from the analysis is presented in a Word document.

3.2 Basic Proc NLMIXED syntax for the models

HSROC model:

```
proc nlmixed data=diag ;
      parms alpha=4 theta=0 beta=0 s2ua=1 s2ut=1;
      logitp = (theta + ut + (alpha + ua) * dis) * exp(-(beta)*dis);
      p = exp(logitp)/(1+exp(logitp));
      model  pos ~ binomial(n,p);
      random ua ut ~ normal([0 , 0], [s2ua,0,s2ut]) subject=study_id;
```

where

diag is the name of the input data set

parms statement defines the model parameters with their starting values:

alpha = accuracy
theta=threshold
beta=shape
s2ua = variance of accuracy
s2ut = variance of threshold

logitp specifies the regression equation, *ut* and *ua* are the random effects for threshold and accuracy respectively. The variable *dis* is the disease indicator which takes the value 0.5 if diseased and -0.5 if not diseased.

p is probability, the inverse transformation of *logitp*

model statement indicates that the number of positives (variable = *pos*) has a binomial distribution with parameters *n* (number in the diseased/non diseased group so *n* = *tp* + *fn* if diseased or *fp* + *tn* if non-diseased) and probability, *p*.

random statement specifies the random effects and their distribution. They are assumed to be independent and normally distributed. *subject* = *study*_id indicates when new realisations of the random effects are to be obtained.

Bivariate model:

```
proc nlmixed data=diag ;
      parms msens=1 mspec=2 s2usens=0.2 s2uspec=0.6 covsesp=0;
      logitp = (msens + usens)*sens + (mspec + uspec)*spec;
      p = exp(logitp)/(1+exp(logitp));
      model  true ~ binomial(n,p);
      random usens uspec ~ normal([0 , 0],  [s2usens,
            covsesp,s2uspec]) subject=study_id;
```

This syntax is similar to that for the HSROC model and fits the model based on the generalized linear mixed model approach proposed by Chu and Cole (7).The only notable changes to the HSROC syntax previously described are:

parms statement which defines the model parameters with their starting values:

msens = mean logit sensitivity

mspec = mean logit specificity

s2usens = variance of logit sensitivity

s2uspec = variance of logit specificity

covsesp = covariance of logit sensitivity and specficity

logitp specifies the regression equation, *usens* and *uspec* are the random effects for logit sensitivity and specificity respectively. *sens* is an indicator variable which takes the value 1 if diseased and 0 if not diseased while *spec* takes the value 1 if not diseased and 0 if diseased.

model statement indicates that the number of true positives or negatives (variable = *true*) has a binomial distribution with parameters *n* (number in the diseased/non diseased group) and probability, *p*.

3.3 Program structure

MetaDAS is modular in structure as illustrated in figure 3.1, comprising of 14 other macros each with distinct functionality.

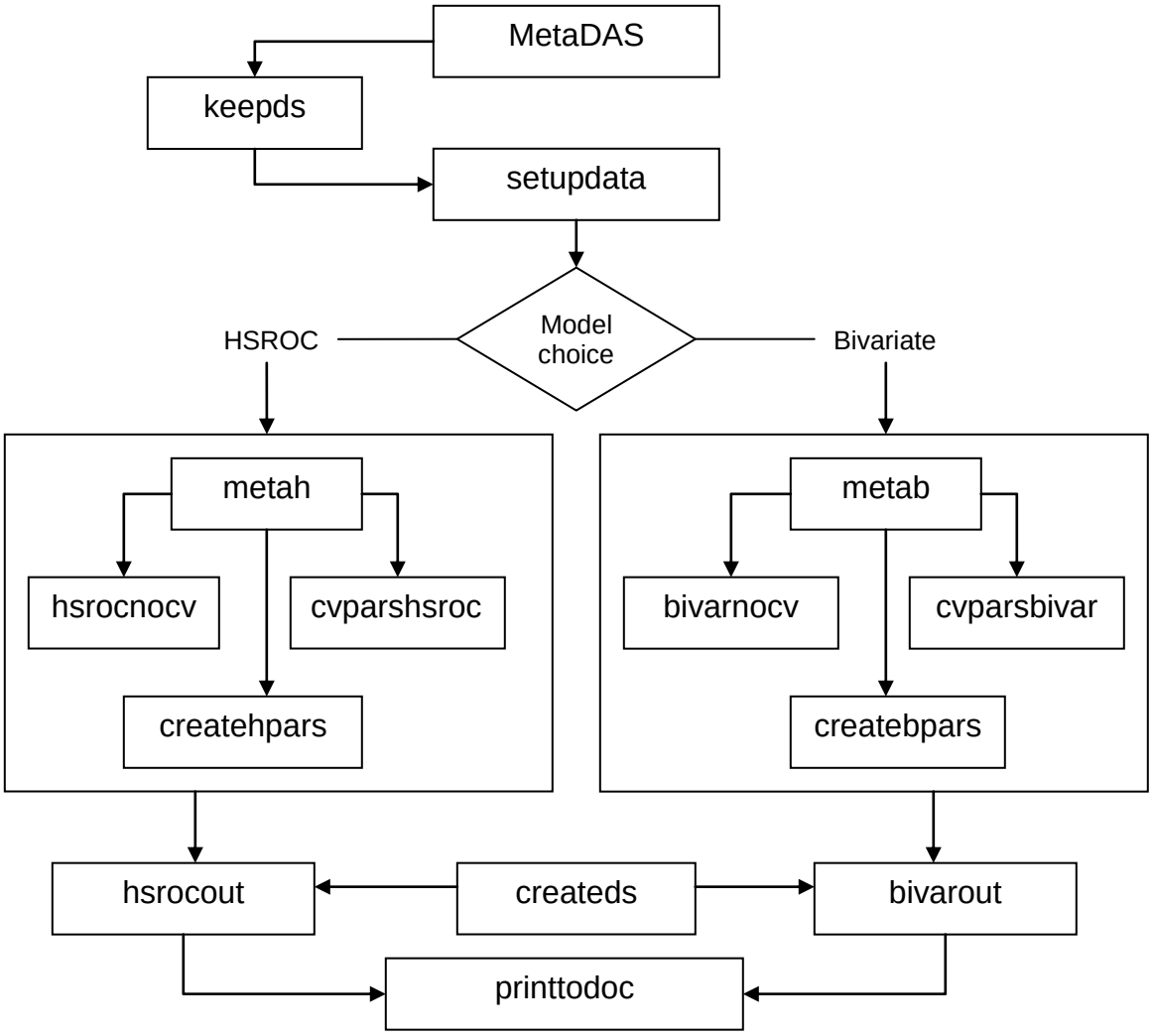


Figure 3.1 Structure of MetaDAS

keepds
The output data sets to keep are selected depending on the option specified for the input parameter *keepds*. This macro is also executed prior to each run of ***MetaDAS*** in order to delete any data sets prefixed with *_metadas_* that may exist in the current library from a previous run of the macro to avoid incorrect output if any of the components of ***MetaDAS*** fails.

setupdata
This macro imports data from the specified Excel .csv or .xls file and modifies the resulting data set into the structure required for the analysis. If a covariate is specified, it creates the required dummy variables. If a *BY* variable is specified, a data set is created for each level of the variable thus enabling the production of separate models.

metah
Performs the HSROC method and calls the macros ***hsrocnocv***, ***createhpars*** and ***cvparshsroc***. If no parameter starting values are specified then it uses the macro specified ones. If the model does not converge, modelling is repeated without the random effects. If this model converges, an attempt is made to fit a model with random effects using a range of values based on the parameter estimates as starting values. When a covariate is specified, modelling with the covariate is only performed if the model without covariate was successfully fitted because its parameter estimates are used as starting values for the covariate model. If predictions are requested, a *PREDICT* statement for *logitp* is included in the specification for Proc NLMIXED. This statement enables predictions of *logitp* (logits of sensitivity and 1-specificity) across all of the observations in the input data set. The predicted values are computed using the parameter estimates and empirical Bayes estimates of the random effects.

hsrocnocv
This macro is called by ***metah*** to perform the HSROC method without a covariate.

createhpars
HSROC parameters and their starting values are created for a model with or without random effects based on the parameter estimates from a previous model. A range is created for each parameter (lower bound is -1 and upper bound is +1 of the value, except for variance terms where the lower bound is 0) in order to produce a grid of values.

cvparshsroc
If a covariate is included in the model, this will create additional parameters for the regression equation depending on the effect required as specified by the input parameter *cveffect*, estimate statements for functions of the model parameters in order to obtain summary measures such as the DOR and likelihood ratios etc. as well as contrast statements for testing the effect of the covariate on accuracy, threshold or shape. The contrast statement(s) enable statistical testing that several expressions simultaneously equal zero. PROC NLMIXED constructs approximate *F* tests for each statement using the delta method.

hsrocout
Manipulates the output data sets obtained from Proc NLMIXED for the HSROC method such that the data required is extracted into data sets for output by ***printtodoc*** to the Word document.

createds
Creates a blank data set that is replicated as needed in ***hsrocout***/***bivarout.*** This is used for storing the values needed for constructing confidence and prediction regions for the SROC curve in RevMan.

printtodoc
Enables output of data to the Word document using the SAS output delivery system. Data sets are checked for missing standard errors and for the bivariate model, whether or not the between study correlation is +1 or -1. Standard errors may be missing where the model converges but the final Hessian matrix is not positive definite. Riley et al. (8) found that a between study correlation of +1 or -1 is likely to be associated with non-convergence and unstable pooled estimates. There may be little information to estimate the correlation and they suggest the HSROC method in such circumstances if possible. This macro also transforms predictions into sensitivities and specificities and calculates their observed values.

metab
This macro calls the macros ***bivarnocv***, ***createbpars*** and ***cvparsbivar*** for the bivariate method. It is similar in function to ***metah***, the major difference being that if the first attempt at fitting the model using Proc NLMIXED fails, ***metab*** fits the bivariate model using Proc MIXED to obtain the new starting values for Proc NLMIXED instead of using Proc NLMIXED without random effects.

bivarnocv
This macro is called by ***metab*** to perform the bivariate method without a covariate.

createbpars
Bivariate parameters and their starting values are created based on the parameter estimates from a previous model.

cvparsbivar
If a covariate is included in the model, this will create additional parameters for the regression equation and estimate statements. as well as contrast statements for testing the effect of the covariate on sensitivity, specificity or both.

bivarout
Manipulates the output data sets obtained from Proc NLMIXED for the bivariate method such that the data required is extracted into data sets for output by ***printtodoc*** to the Word document.

3.4 Syntax

%macro metadas(dtfile=, import=, dsname=, tech=, ident=,
tp=, fp=, fn=, tn=,
subject=, cialpha=, byvar=, covariate=, cvref=,
sortcv=, cvtype=, cveffect=, cvsummorder=,
formatlr=, test=, method=, mtitle=, tbpe=,
p1=, p2=, p3=, p4=, p5=,
cspa1=, cspa2=, cspa3=, cspa4=,
cset1=, cset2=, cset3=, cset4=,
cpb1=, cpb2=, cpb3=, cpb4=,
randeffs=, predict=, checkmod=, debug=,
logfile=, outfile=, keepds=, revman=,
info=, bothmodels=, incbasic=,
rfile=);

There are 52 input parameters available with ***MetaDAS*** as outlined in table 1 below.

Input parameter	Description and parameter values
dtfile='text'	The path and name of the Excel file to import e.g. 'C:\Documents\DTA\Revman Test Data.xls'. The file extension (.xls or .csv) must be included.
import=y/n	If =*n*, a data set must be provided with the *dsname*= option. The default is *y*.
dsname=data set	The input data set if no data import is required.
tech=quanew/newrap/trureg/ nrridg/dbldog/congra/ nmsimp	There are several optimization techniques available with Proc NLMIXED. No algorithm for optimizing general nonlinear functions exists that always finds the global optimum for a general nonlinear minimization problem in a reasonable amount of time (9). This parameter enables the user to select a technique as they would do if they were running NLMIXED directly. The default is *tech*=QUANEW. With the exception of options *START*, *DF*, *ALPHA*, *HESS*, *COV* and *ECOV* (they are already in use), you can also specify other Proc NLMIXED options by tagging them on to this parameter e.g. *tech* = *newrap gconv*=1e-9 *qtol*=1e-5. For more information and algorithm descriptions, see the SAS user documentation for NLMIXED.
ident=y/n	A potential problem with numerical maximization of the likelihood function is identifiability of model parameters. When this occurs, the likelihood will equal its maximum value at a set of parameter values instead of at a single point. To detect if there is a problem, you could try different initial values of the parameters and check for changes in parameter estimates or by examining the Hessian matrix at convergence (10).

	If *ident=y* the Hessian matrix after optimization is produced and the eigenvalues of the Hessian are calculated (with values saved in _metadas_a_eigenvals_ / _metadas_cv_eigenvals_). At a true minimum, the eigenvalues will all be positive, i.e., positive definite. The default is *y*. The starting Hessian matrix is also produced because Proc NLMIXED option *START* is always used by ***MetaDAS*** to output the gradient at the starting values.
tp=variable	The number of true positives. The default variable name is *tp* so that RevMan users or those who have named their variables accordingly do not need to specify this input parameter.
fp=variable	The number of false positives. The default variable name is *fp*.
fn=variable	The number of false negatives. The default variable name is *fn*.
tn=variable	The number of true negatives. The default variable name is *tn*.
subject=variable	This determines when new realizations of the random effects are assumed to occur. Proc NLMIXED assumes that a new realization occurs whenever the *subject*= variable changes from the previous observation, so the input data set is clustered according to this variable. The default variable name is *study*_id (as named in the RevMan 5 data export file).
cialpha=numeric	Specifies the alpha level for computing *z* statistics and confidence limits. The default is 0.05.
byvar=variable	This enables multiple analyses, i.e., consecutive calls to Proc NLMIXED for each test or group of studies in the data file. This may also be used to produce separate models using subsets of the data (subgroup analyses as in traditional meta-analysis) but be aware this is not recommended because you cannot formally test for a difference. A better approach is to use all the data and include the variable as a covariate in the model.
covariate=variable	Specifies a covariate for inclusion in the model (meta-regression). Covariates can be included in the model to determine the effect of patient or study characteristics on threshold, accuracy, and the shape of the SROC (individually or in any combination) for the HSROC model or on sensitivity and/or specificity for the bivariate model. For example, to compare multiple tests use test type as a covariate in the model.
cvref='text'/numeric	This specifies the reference level of the covariate. If it is not specified, the reference level is selected based on the sort order. Sorting is done in ascending order by default and for descending specify *sortcv=d*.
sortcv=d/a	The sort order for the covariate. *sortcv=d* specifies descending order and *a* specifies ascending. The default is to sort in ascending order.
cvtype=cat/con	Type of covariate. Options are *cat* for categorical or *con* for continuous. If the parameter is not specified, the covariate is assumed to be categorical.
cveffect= a/t/b/at/ab/bt/	For the HSROC model *t* specifies that the effect of the covariate be assessed only on theta, *a* on alpha only, *b* on only beta, *ab* on alpha

abt/se/sp/sesp	and beta, *at* on alpha and theta, *bt* on beta and theta, and *abt* on all three parameters. Default is *abt*. For the bivariate model, *se* specifies that the effect be assessed only on sensitivity while *sp* on specificity and *sesp* specifies effect on both sensitivity and specificity. Default is *sesp*.
cvsummorder= stat/level	Specifies the ordering of items in the table of summary estimates for a model with covariate. If *level* is specified, items are listed in the table according to covariate level. If *stat* is specified, items are listed according to summary statistic such that all levels of the covariate are grouped together for each statistic. The default is *stat*.
formatlr=y/n	For formatting the log likelihood difference and *p*-value obtained for the likelihood ratio test. If =*y*, then -2logL difference is formatted to 3 decimal places if it is greater than or equal to 0.001 otherwise the exact value is reported. The *p*-value is formatted to 3 d.p. if less than or equal to 0.001 and as <0.001 if less than 0.001. The default is *y*.
test='text'/numeric	The name of the test to analyse if the data file contains more than one test on which we wish to perform a variety of analyses. No need to specify a test if there is only one.
method=h/b	Specifies the type of model to fit. Options are *b* for bivariate or *h* for HSROC method. The default is *h*.
mtitle=text	Title of the meta-analysis that is placed in the Word document. Default is Meta-analysis of diagnostic test accuracy studies. **Note: no quotation marks allowed unlike some of the other text options.**
tbpe=data set	Use parameters and starting values stored in the named table. The data set can be in either a narrow or wide form. The narrow-form data set contains the variables *PARAMETER* and *ESTIMATE*, with parameters and values listed as distinct observations. The wide-form data set has the parameters themselves as variables, and each observation provides a different set of starting values. **Note:** In this version of *MetaDAS*, the data set should only contain the 5 basic parameters for either the HSROC (alpha, theta, beta, s2ua and s2ut) or bivariate model (msens, mspec, s2usens, s2uspec, covsesp). If there is a covariate, the starting values for additional parameters can be specified using *cspa1 – cspa5*, *cset1 – cset5* and/or *cpb1 – cpb5*.
p1 – p5	These are the basic parameters and their starting values. There are five such parameters for either model. You can either specify a single number e.g. *p1*= 2.5 or you can use the TO and BY keywords to specify a number list for a grid search e.g. *p1* = -2 to 2 by 0.5. If you specify a grid of points, the objective function value at each grid point is calculated and the best (feasible) grid point is chosen as an initial point for the optimization process. ***For HSROC model:*** *p1* = alpha (accuracy parameter), *p2* = theta (threshold parameter), *p3*

	= beta (shape parameter), *p4* = variance of accuracy, *p5* = variance of threshold. The default values are: *p1*= -4 to 4 by 1 *p2*= -2 to 2 by 1 *p3*= -3 to 1 by 0.5 *p4*= 0 to 1 by 0.2 *p5*= 0 to 1 by 0.2 ***For bivariate model:*** *p1* = mean logit sensitivity, *p2* = mean logit specificity, *p3*= variance of logit sensitivity, *p4* = variance of logit specificity, *p5* = covariance of logit sensitivity and specificity. The default values are: *p1*= -2 to 4 by 1 *p2*= -2 to 4 by 1 *p3*= 0 to 2 by 0.25 *p4*= 0 to 2 by 0.25 *p5*= -1 to 1 by 0.2
cspa1 – cspa5	If the HSROC model is required, these specify starting values for additional alpha parameters or if it is the bivariate model then they are for additional specificity parameters e.g. *cspa1* = 0 to 2 by 1. *cspa1 – cspa5* indicates a maximum of 5 parameters, i.e., a covariate with 6 levels. The default is 0 for any of the 5 parameters, i.e., *cspa1* = 0, *cspa2* = 0 for a covariate with 3 levels.
cset1 – cset5	Starting values for additional theta or sensitivity parameters. The maximum is 5, i.e., a covariate with 6 levels,. The default is 0 for any of the 5 parameters, i.e., cset1 = 0, cset2 = 0 for a covariate with 3 levels.
cpb1 – cpb5	Starting values for additional beta parameters. Applies to each level of the covariate except the reference level, therefore a maximum of 5 parameters, i.e., a covariate with 6 levels.
randeffs=y/n	Produce table of empirical Bayes estimates of the random effects if = *y*. The default is *n*.
predict=y/n	If =*y*, predictions are obtained using the estimated model, parameter estimates and empirical Bayes estimates of the random effects. Standard errors of prediction are computed using the delta method and the predicted values of logit p (stored in data sets prefixed with _logitp_ and _logitp_cv_) are transformed to obtain predictions of sensitivity and specificity (stored in data sets prefixed with _predsesp_ and _predsesp_cv_). The default is *n*.
checkmod=y/n	If =*y*, produce histograms and normal probability plots of the empirical Bayes estimates of the random effects to check assumption of normality. The default is *n.*
debug=y/n	Debugging tool. If =*y*, displays the SAS statements that are generated by macro execution. The default is *n.*
logfile='text'	Path and file name to save the contents of the SAS log. Must add the .log extension. Contents of the log file are scanned and any errors

	found are stored in _metadas_errors, warnings in _metadas_warnings, and model failure messages generated by ***MetaDAS*** in _metadas_modfail. The data set for the log contents is _metadas_log.
outfile='text'	Path and filename to save the contents of the SAS output window. The file name must have the .lst extension. This is especially useful if the analysis is expected to run for awhile because the output window will fill up and user input is required before SAS can proceed. However, this is not the case if the output is being saved to a file.
keepds=all/some/log/none	Selectively keeps the data sets produced as output from the analyses. Option *some* is the default. With this option, data sets containing data from the Excel file are kept, including any data sets generated from the log file if a log file was specified. For the option *log*, only the data sets generated from the log file are kept. If option *none* is specified, all data sets prefixed with _metadas_ are deleted. Option *all* keeps all output data sets from NLMIXED as well as two summary ones for covariate summary and relative measures of test accuracy. Data sets for predictions, random effects, the Hessian matrix and eigenvalues are also kept with options *all* and *some* if parameters have been specified for them. ***MetaDAS* output data sets** All data from Excel file =_metadas_meta Unique values of the BY variable = _metadas_*variablename* Data set for level i of the BY variable = _metadas_dsi Unique values of the covariate = _metadas_*variablename* Predicted logitp for model without covariate = _metadas_logitp_i Predicted logitp for model with covariate = _metadas_cv_logitp_i Predicted sensitivities and specificities for model without covariate = _metadas_predict_i Predicted sensitivities and specificities for model with covariate = _metadas_cv_predict_i Relative estimates of accuracy measures for covariate = _metadas_cv_relsummary_i Summary estimates of accuracy measures for covariate = _metadas_cv_statsummary_i Eigenvalues for model without covariate =_metadas_a_eigenvals_ Eigenvalues for model with covariate =_metadas_cv_eigenvals_ SAS NLMIXED output data sets are prefixed by ***metadas*** as follows: **Model without covariate** Starting values =_metadas _a_sv_ Parameters=_metadas_a_parms_ Parameter estimates=_metadas_a_pe_ Fit statistics=_metadas_a_fit_ Additional estimates=_metadas_a_addest_

	Covariance matrix of additional estimates =_metadas_a_covaddest_ Convergence status=_metadas_a_convgstat_ Final Hessian matrix=_metadas_a_hessian_ **Model with covariate** Starting values = _metadas_cv_sv_ Parameters=_metadas_cv_parms_ Parameter estimates=_metadas_cv_pe_ Fit statistics=_metadas_cv_fit_ Additional estimates=_metadas_cv_addest_ Covariance matrix of additional estimates =_metadas_cv_covaddest_ Convergence status=_metadas_cv_convgstat_ Contrasts=_metadas_cv_contrasts_ Final Hessian matrix=_metadas_cv_hessian_ For the bivariate model there are 2 additonal tables, metadas_cv_covparmest_ and metadas_cv_covparmest_, for the covariance marix of parameter estimates.
revman='text'	Launch the specified RevMan 5 file at the end of analysis so that parameters can be copied and pasted into the appropriate cells for the analysis in the external analyses section.
info=y/n	If =y, include details of some of the input parameters specified for the macro. The default is *y*.
bothmodels=y/n	If = *y* both models are included in the output. For instance, if the method is HSROC then bivariate parameters are obtained as functions of the HSROC parameters and included in the output. The default is *n*.
incbasic=y/n	If = *n* then the output for the model with no covariate is suppressed. This may be useful where the model with no covariate has already been investigated and the parameters are no longer of interest for extraction to RevMan or in test comparisons where the covariate is test type. The default is *y*.
rfile='text'	Path and name of the Word document to save the result of the analyses. The file name must have the .rtf extension (rich text file).

Table 1 Input parameters for *MetaDAS*

Note: Options are not case sensitive. The macro requires a minimum of 2 or 3 options depending on whether data import is required or not. These are the path and name of the Excel data input file or SAS data set if data import is not required (*import*=n), and the Word file for the analysis output.

3.5 Examples

3.5.1 Only the 2 required options: file to import and file to output

```
%metadas(dtfile= 'C:\Documents and Settings\username\My
                  Documents\DTA\Revman Test Data.xls',
                  rfile ='c:\hsroc test.rtf');
run;
```

3.5.2 Some more options included

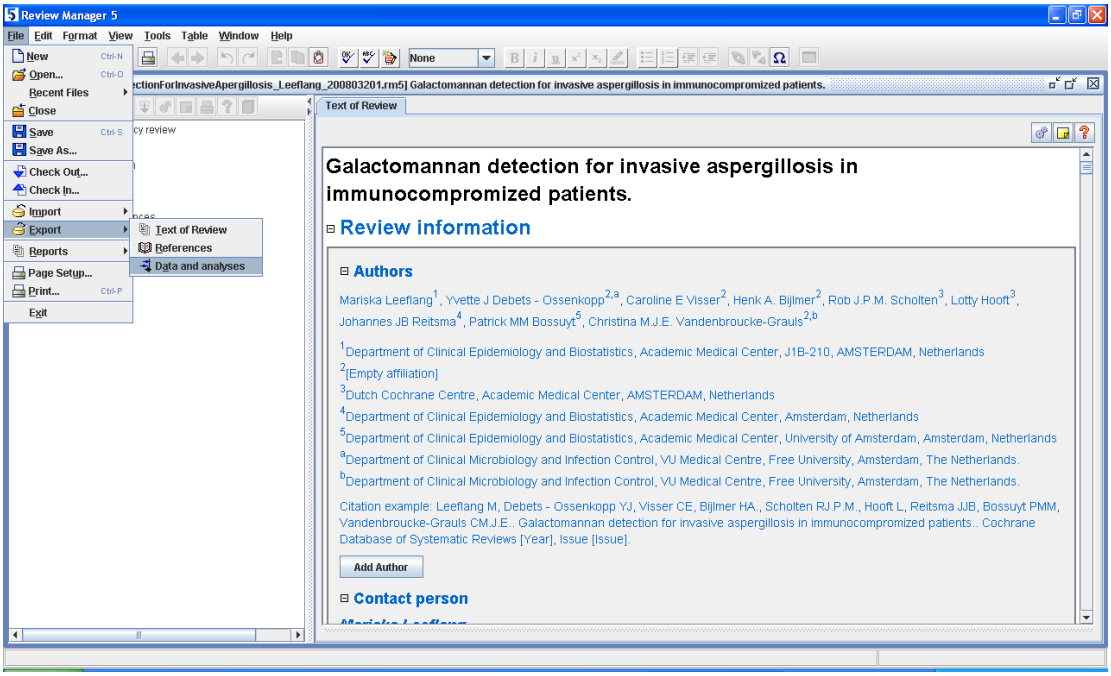
```
%metadas(dtfile= 'C:\Revman Test Data.xls', tech=newrap,
                  covariate=stage, byvar=test_type,
                  cveffect=a, test='HPV', predict=y,
                  debug=n, rfile ='c:\hsroc test.rtf');
run;
```

```
%metadas(dtfile= 'C:\DTA\Galactomannan detection.xls',
                  test='Platelia - cutoff 0.5',
                  debug=y, method=b,
                  covariate=Pat_base, checkmod=y,
                  tech = newrap gconv=1e-9 qtol=1e-5,
                  rfile ='c:\DTA\GD basic hsroc model.rtf',
                  cvref='Patient-based data', cvsummorder=stat,
                  bothmodels=y, keepds=some,
                  logfile='C:\DTA\GD logtest.log',
                  outfile='C:\DTA\GD outtest.lst');
run;
```

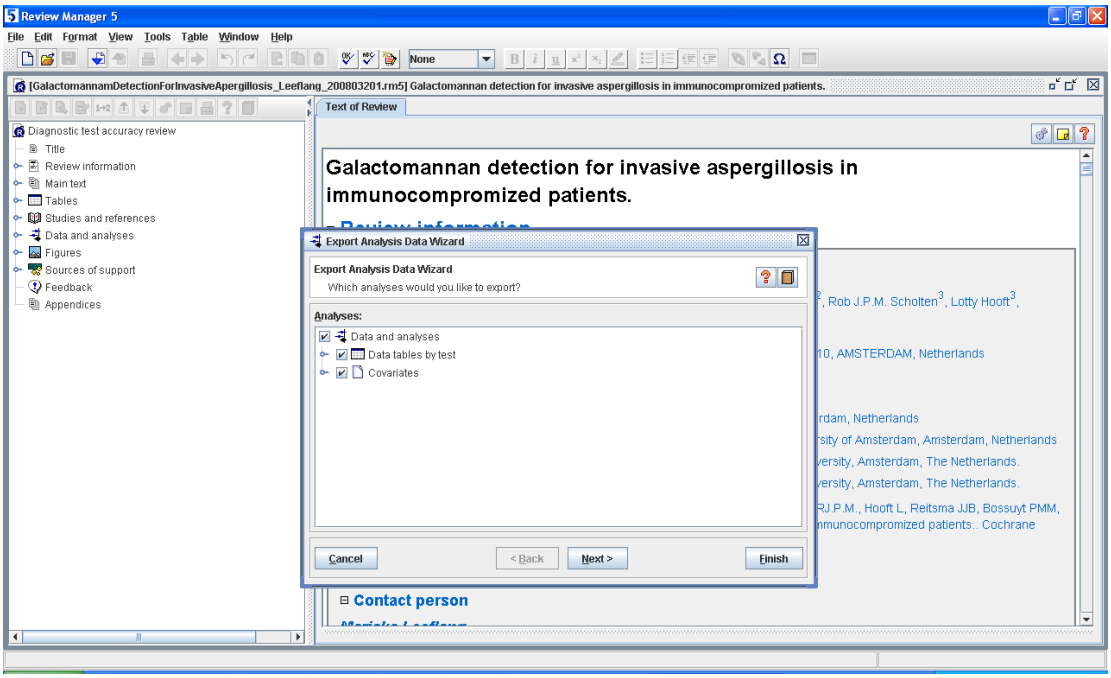
4 Worked example

4.1 File export from RevMan

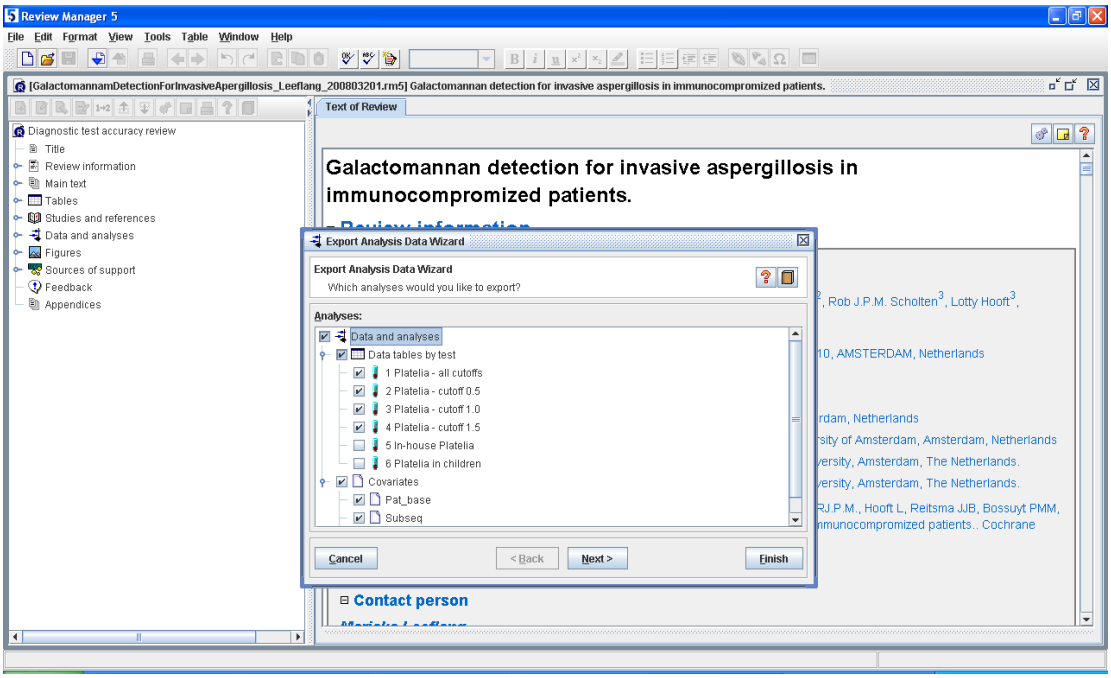
Open your RevMan file. On the menu bar, click on file and then click on export and select data and analyses.



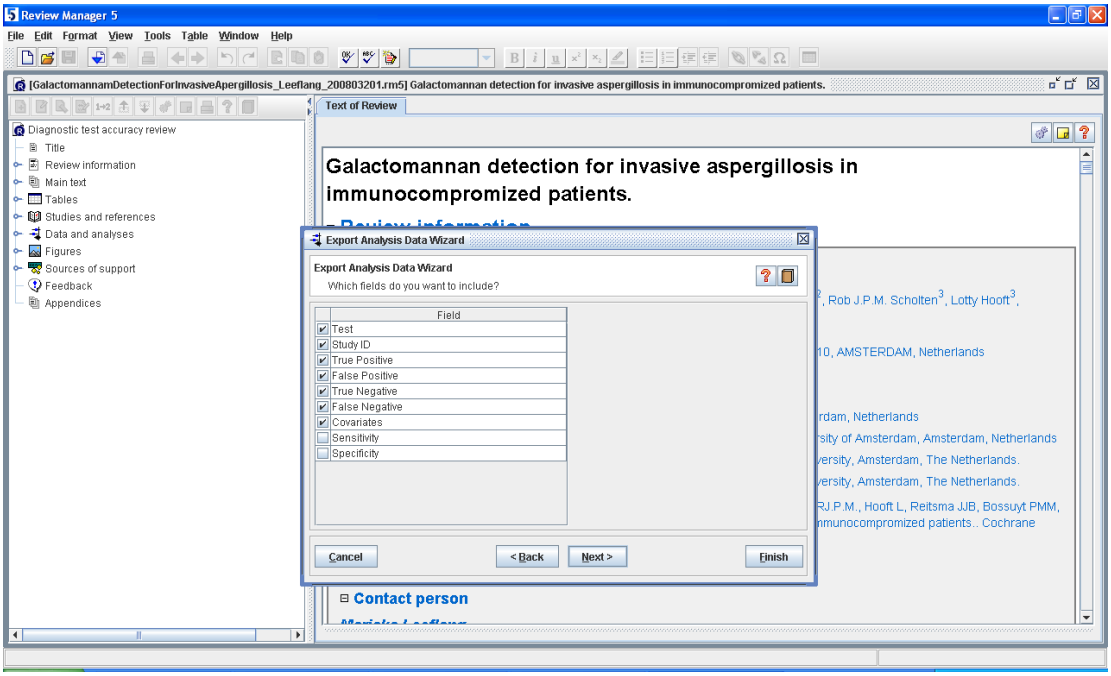
The export analysis data wizard is launched.



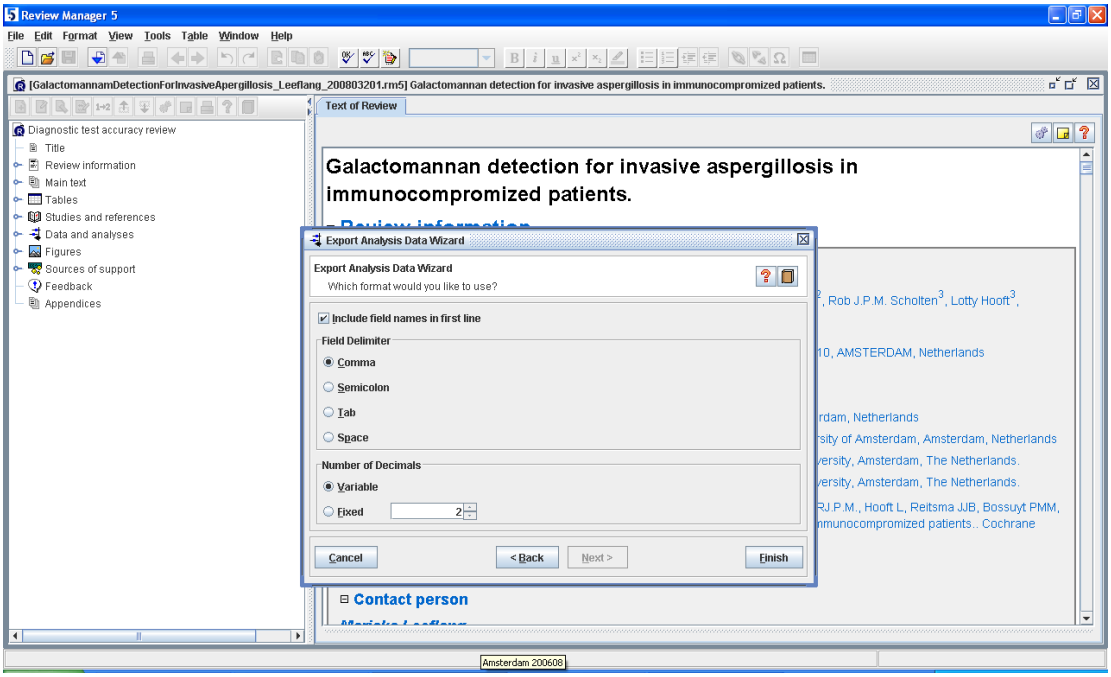
If you do not wish to select specific data tables by test or covariates then click next. If you wish to select then expand the tree and make your selection as shown below.



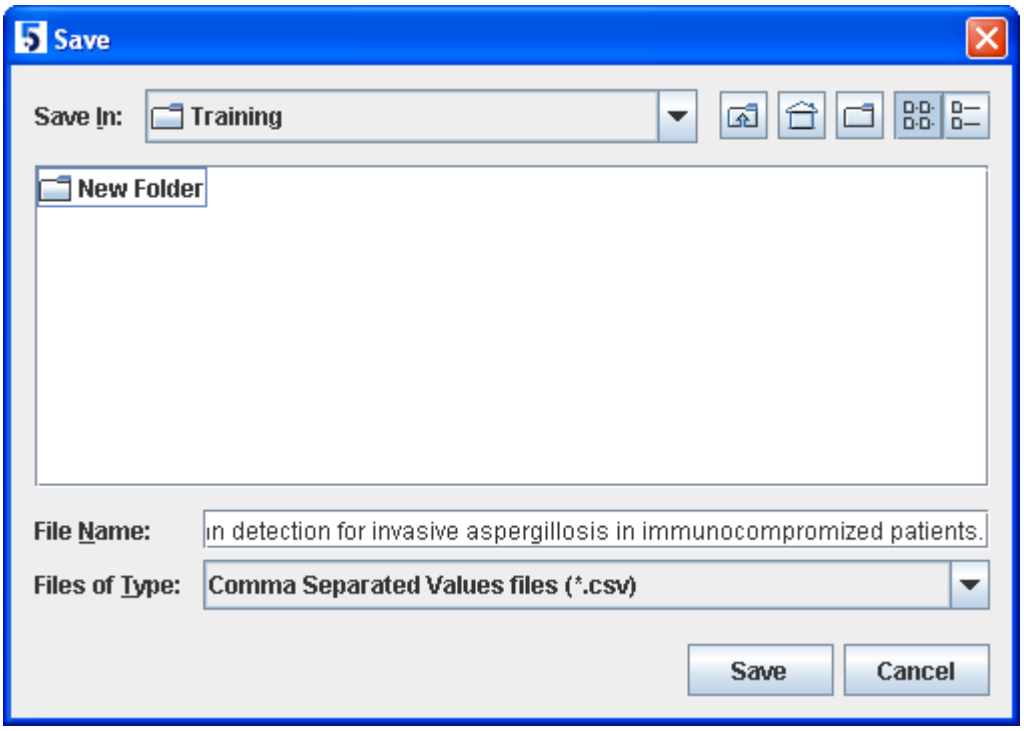
On the next page of the wizard select the fields you wish to export. Click on next.



Select the export format you require. Typically the field delimiter you require is comma and ensure that the box for *include field names in first line* is ticked.



Click on finish. This opens the save dialog box. Remember to add on the .csv extension to the file name as RevMan does not do that automatically as of RevMan 5.12. Click save.



A sample of the extracted data is shown in figure 4.1.

Test	Study ID	TP	FP	TN	FN	Pat_base
Platelia - cutoff 0.5	Allan 2005	0	11	113	1	Episode-based
Platelia - cutoff 0.5	Florent 2006	8	39	116	4	Patient-based data
Platelia - cutoff 0.5	Foy 2007	6	7	102	6	Patient-based data
Platelia - cutoff 0.5	Kawazu 2004	11	23	115	0	Episode-based
Platelia - cutoff 0.5	Suankratay 2006	16	13	20	1	Patient-based data
Platelia - cutoff 0.5	Weisser 2005	16	41	100	4	Episode-based
Platelia - cutoff 0.5	Yoo 2005	12	25	89	2	Patient-based
Platelia - cutoff 1.0	Allan 2005	0	1	123	1	Episode-based
Platelia - cutoff 1.0	Becker 2003	6	12	62	7	Patient-based data
Platelia - cutoff 1.0	Bretagne 1998	14	5	18	4	Patient-based
Platelia - cutoff 1.0	Busca 2006	2	12	60	0	Patient-based
Platelia - cutoff 1.0	Challier 2004	20	9	35	6	Patient-based data
Platelia - cutoff 1.0	Kawazu 2004	7	4	134	4	Episode-based
Platelia - cutoff 1.0	Maertens 2002	11	7	80	2	Episode-based
Platelia - cutoff 1.0	Marr 2004	13	11	32	11	Patient-based
Platelia - cutoff 1.0	Pereira 2005	1	9	29	0	Patient-based
Platelia - cutoff 1.0	Pinel 2003	17	17	756	17	Patient-based
Platelia - cutoff 1.0	Suankratay 2006	16	2	31	1	Patient-based data
Platelia - cutoff 1.0	Ulusakarya 2000	16	11	108	0	Patient-based
Platelia - cutoff 1.5	Adam 2004	1	41	175	1	Patient-based data
Platelia - cutoff 1.5	Allan 2005	0	1	123	1	Episode-based
Platelia - cutoff 1.5	Bialek 2002	1	8	8	0	Patient-based data
Platelia - cutoff 1.5	Buchheidt 2004	3	1	167	6	Episode-based
Platelia - cutoff 1.5	Doermann 2002	10	4	407	2	Patient-based

Figure 4.1 Sample of data from the Excel .csv file

4.2 Run *MetaDAS*

4.2.1 Sample SAS statements to run the macro

```
%include 'C:\My SAS programs\Metadas version 1.0.sas';

%metadas(dtfile= 'C:\Training\Galactomannan detection for invasive
aspergillosis in immunocompromized patients.csv', test='Platelia - cutoff
1.0', covariate=Pat_base, logfile='C:\Training\GD Log.log', debug=y,
keepds=all, predict=y, bothmodels=y, checkmod=y,
rfile ='C:\Training\GD hsroc model with covariate 1.0.rtf');

run;
```

The *include* statement specifies the path and name of the SAS file containing the macro. This is followed by the macro statement with options.

4.2.2 Results

a. Error checking

The input parameter *logfile* is considered to be very useful. If the content of the log window is saved to a file, the tables _metadas_errors, _metadas_warnings and _metadas_modfail are produced and can be used in identifying problems with the model or macro instead of trawling through the entire log. In the current example, the log as shown in figure 4.2 reveals zero observations in the respective tables, i.e., there were no errors, warnings or model failure messages. Whenever there are observations, examine the relevant table(s) and use the logline to further investigate the problem in the log table (_metadas_log). This is especially informative when the *debug* parameter has been specified as *y*.

```
*******************************************************
*                                                     *
*      META-ANALYSIS OF DIAGNOSTIC ACCURACY STUDIES   *
*                                                     *
*******************************************************
NOTE: PROCEDURE PRINTTO used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds



NOTE: The infile LOGFILE is:

      File Name=C:\Training\GD HSROC model with covariate 1.0.log,
      RECFM=V,LRECL=256

NOTE: 3199 records were read from the infile LOGFILE.
      The minimum record length was 0.
      The maximum record length was 135.
NOTE: The data set WORK._METADAS_LOG has 3199 observations and 1 variables.
NOTE: The data set WORK._METADAS_ERRORS has 0 observations and 4 variables.
NOTE: The data set WORK._METADAS_WARNINGS has 0 observations and 4
variables.
NOTE: The data set WORK._METADAS_MODFAIL has 0 observations and 4
variables.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.03 seconds
```

Figure 4.2 Log content with input parameter *logfile*

b. Data import

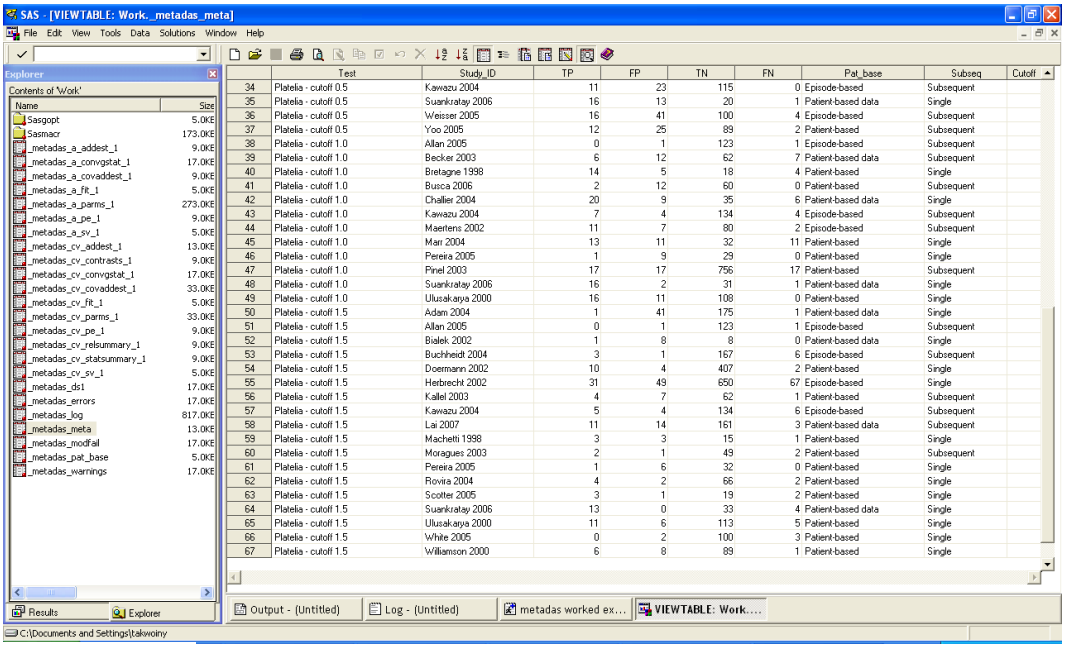
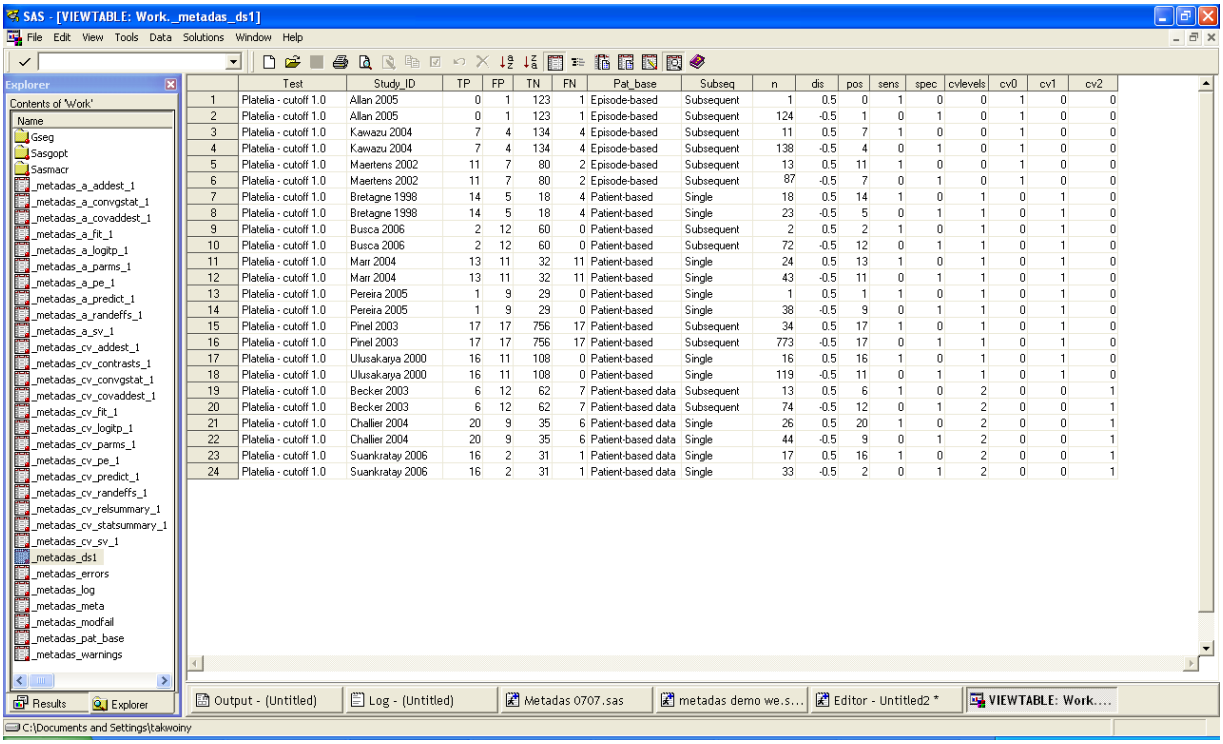


Figure 4.3 Sample of contents of table _metadas_meta

The data set _metadas_meta contains all the data from the input data file without any modification as shown in figure 4.3. Figure 4.4 shows the data set _metadas_ds1 which contains data for Platelia – cutoff 1.0 and this has been modified to include 2 records for each study as well as additional variables required for running the HSROC model with a covariate Pat_base.



	Test	Study_ID	TP	FP	TN	FN	Pat_base	Subseq	n	dis	pos	sens	spec	cvlevels	cv0	cv1	cv2
1	Platelia - cutoff 1.0	Allan 2005	0	1	123	1	Episode-based	Subsequent	1	0.5	0	1	0	0	1	0	0
2	Platelia - cutoff 1.0	Allan 2005	0	1	123	1	Episode-based	Subsequent	124	-0.5	1	0	1	0	1	0	0
3	Platelia - cutoff 1.0	Kawazu 2004	7	4	134	4	Episode-based	Subsequent	11	0.5	7	1	0	0	1	0	0
4	Platelia - cutoff 1.0	Kawazu 2004	7	4	134	4	Episode-based	Subsequent	138	-0.5	4	0	1	0	1	0	0
5	Platelia - cutoff 1.0	Maertens 2002	11	7	80	2	Episode-based	Subsequent	13	0.5	11	1	0	0	1	0	0
6	Platelia - cutoff 1.0	Maertens 2002	11	7	80	2	Episode-based	Subsequent	87	-0.5	7	0	1	0	1	0	0
7	Platelia - cutoff 1.0	Bretagne 1998	14	5	18	4	Patient-based	Single	18	0.5	14	1	0	1	0	1	0
8	Platelia - cutoff 1.0	Bretagne 1998	14	5	18	4	Patient-based	Single	23	-0.5	5	0	1	1	0	1	0
9	Platelia - cutoff 1.0	Busca 2006	2	12	60	0	Patient-based	Subsequent	2	0.5	2	1	0	1	0	1	0
10	Platelia - cutoff 1.0	Busca 2006	2	12	60	0	Patient-based	Subsequent	72	-0.5	12	0	1	1	0	1	0
11	Platelia - cutoff 1.0	Marr 2004	13	11	32	11	Patient-based	Single	24	0.5	13	1	0	1	0	1	0
12	Platelia - cutoff 1.0	Marr 2004	13	11	32	11	Patient-based	Single	43	-0.5	11	0	1	1	0	1	0
13	Platelia - cutoff 1.0	Pereira 2005	1	9	29	0	Patient-based	Single	1	0.5	1	1	0	1	0	1	0
14	Platelia - cutoff 1.0	Pereira 2005	1	9	29	0	Patient-based	Single	38	-0.5	9	0	1	1	0	1	0
15	Platelia - cutoff 1.0	Pinel 2003	17	17	756	17	Patient-based	Subsequent	34	0.5	17	1	0	1	0	1	0
16	Platelia - cutoff 1.0	Pinel 2003	17	17	756	17	Patient-based	Subsequent	773	-0.5	17	0	1	1	0	1	0
17	Platelia - cutoff 1.0	Ulusakarya 2000	16	11	108	0	Patient-based	Single	16	0.5	16	1	0	1	0	1	0
18	Platelia - cutoff 1.0	Ulusakarya 2000	16	11	108	0	Patient-based	Single	119	-0.5	11	0	1	1	0	1	0
19	Platelia - cutoff 1.0	Becker 2003	6	12	62	7	Patient-based data	Subsequent	13	0.5	6	1	0	2	0	0	1
20	Platelia - cutoff 1.0	Becker 2003	6	12	62	7	Patient-based data	Subsequent	74	-0.5	12	0	1	2	0	0	1
21	Platelia - cutoff 1.0	Challier 2004	20	9	35	6	Patient-based data	Single	26	0.5	20	1	0	2	0	0	1
22	Platelia - cutoff 1.0	Challier 2004	20	9	35	6	Patient-based data	Single	44	-0.5	9	0	1	2	0	0	1
23	Platelia - cutoff 1.0	Suankratay 2006	16	2	31	1	Patient-based data	Single	17	0.5	16	1	0	2	0	0	1
24	Platelia - cutoff 1.0	Suankratay 2006	16	2	31	1	Patient-based data	Single	33	-0.5	2	0	1	2	0	0	1

Figure 4.4 Content of table _metadas_ds1

c. Analysis

Only one test (Platelia – cutoff 1.0) is analysed as specified by the input parameter *test* although the data file contains a number of other tests. With parameter *keepds*=all, output data sets are not destroyed at the end of the analysis. The translated macro code relevant for fitting the model with a 3-level covariate (Pat_Base) with effect on accuracy(alpha), threshold(theta) and shape(beta) is as follows:

```
ods output StartingValues=_metadas_cv_sv_1
        Parameters=_metadas_cv_parms_1
        ParameterEstimates=_metadas_cv_pe_1
        FitStatistics=_metadas_cv_fit_1
        AdditionalEstimates=_metadas_cv_addest_1
        CovMatAddEst=_metadas_cv_covaddest_1
        ConvergenceStatus=_metadas_cv_convgstat_1
        Contrasts=_metadas_cv_contrasts_1;

proc nlmixed data=_metadas_ds1 cov ecov df=1000 start alpha=0.05 tech=quanew;
         title "HSROC analysis with covariate PAT_BASE";
        parms alpha = 2 to 4 by 1 theta = -2 to 0 by 1 beta = -3 to 2 by 1 s2ua = 0 to 2.33 by 1.165
                s2ut = 0 to 1.6 by 0.8 alpha_cv1=0 theta_cv1=0  beta_cv1=0 alpha_cv2=0 theta_cv2=0
                beta_cv2=0;
        bounds s2ua >= 0; /* set boundary constraint so variance to not negative */
        bounds s2ut >= 0;
        logitp= (theta + ut + ((theta_cv1 * cv1) + (theta_cv2 * cv2)) +(alpha + ua + (alpha_cv1 * cv1) +
                (alpha_cv2 * cv2)) * dis) * exp(-(beta+((beta_cv1 * cv1) + (beta_cv2 * cv2)))*dis);
        p = exp(logitp)/(1+exp(logitp));
        model pos ~ binomial(n,p);
        random ut ua ~ normal([0,0],[s2ut,0,s2ua]) subject=study_id ;
        predict logitp out=_metadas_cv_logitp_1
        estimate 'E(logitSe)' exp(-beta/2) * (theta + 0.5 * alpha);
        estimate 'E(logitSp)' -exp(beta/2) * (theta - 0.5 * alpha);
```

```
estimate 'Var(logitSe)' exp(-beta) * (s2ut + 0.25 * s2ua);
estimate 'Var(logitSp)' exp(beta) * (s2ut + 0.25 * s2ua);
estimate 'Cov(logits)' - (s2ut - 0.25 * s2ua);
estimate 'Corr(logits)'(- (s2ut - 0.25 * s2ua))/(sqrt(exp(-beta) * (s2ut + 0.25 * s2ua))*sqrt(exp(beta) *
                                    s2ut + 0.25 * s2ua)));
estimate "True positive log odds ratio  cv level 1 vs 0" exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1
          + 0.5 * (alpha + alpha_cv1))-(exp(-(beta+beta_cv1)*0.5) * (theta + 0.5 * alpha)) ;
estimate "True positive log odds ratio  cv level 2 vs 0" exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2
          + 0.5 * (alpha + alpha_cv2))-(exp(-(beta+beta_cv2)*0.5) * (theta + 0.5 * alpha));
estimate "True negative log odds ratio  cv level 1 vs 0"  -exp((beta+beta_cv1)*0.5) * (theta + heta_cv1
          - 0.5 * (alpha + alpha_cv1))+ (exp((beta+beta_cv1)*0.5) * (theta - 0.5 * alpha)) ;
estimate "True negative log odds ratio  cv level 2 vs 0" -exp((beta+beta_cv2)*0.5) * (theta + theta_cv2
          - 0.5 * (alpha + alpha_cv2))+ (exp((beta+beta_cv2)*0.5) * (theta - 0.5 * alpha));
estimate "E(logitSe)_1" exp(-(beta+beta_cv1)*0.5) * (theta + 0.5 * alpha)+exp(-(beta+beta_cv1)*0.5) *
          (theta + theta_cv1 + 0.5 * (alpha + alpha_cv1)) - (exp(-(beta+beta_cv1)*0.5) * (theta + 0.5 *
          alpha)) ;
estimate "E(logitSe)_2" exp(-(beta+beta_cv2)*0.5) * (theta + 0.5 * alpha)+exp(-(beta+beta_cv2)*0.5) *
          (theta + theta_cv2 + 0.5 * (alpha + alpha_cv2)) - (exp(-(beta+beta_cv2)*0.5) * (theta + 0.5 *
          alpha));
estimate "E(logitSp)_1" -exp((beta+beta_cv1)*0.5) * (theta - 0.5 * alpha)+-exp((beta+beta_cv1)*0.5) *
          (theta + theta_cv1 - 0.5 * (alpha + alpha_cv1)) + (exp((beta+beta_cv1)*0.5) * (theta - 0.5 * alpha))
          ;
estimate "E(logitSp)_2" -exp((beta+beta_cv2)*0.5) * (theta - 0.5 * alpha)+-exp((beta+beta_cv2)*0.5) *
          (theta + theta_cv2 - 0.5 * (alpha + alpha_cv2)) + (exp((beta+beta_cv2)*0.5) * (theta - 0.5 * alpha));
estimate 'logDOR_0' log(((exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) * (theta + 0.5 *
          alpha)))) /(1-(exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) * (theta + 0.5 *
          alpha))))) /((1-(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5 *
          alpha)))) /(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5 * alpha))))));
estimate "logDOR_1" log(((exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
          alpha_cv1)))/(1+exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha
          + alpha_cv1))))) /(1-(exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
          alpha_cv1)))/(1+exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
          alpha_cv1)))))) /((1-(exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1)))/(1+exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1))))) /(exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1)))/(1+exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1)))))) ;
estimate "logDOR_2" log(((exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha +
          alpha_cv2)))/(1+exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha
          + alpha_cv2))))) /(1-(exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha +
          alpha_cv2)))/(1+exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha +
          alpha_cv2)))))) /((1-(exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
          alpha_cv2)))/(1+exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
          alpha_cv2))))) /(exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
          alpha_cv2)))/(1+exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
          alpha_cv2)))))));
estimate "logRelative sensitivity cv level 1 vs 0" log(exp(exp(-(beta+beta_cv1)/2) * (theta + theta_cv1
          + 0.5 * (alpha+alpha_cv1)))/(1+exp(exp(-(beta+beta_cv1)/2) * (theta + theta_cv1 + 0.5 *
          (alpha+alpha_cv1))))) - log(exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) *
          (theta + 0.5 * alpha))));
estimate "logRelative sensitivity cv level 2 vs 0" log(exp(exp(-(beta+beta_cv2)/2) * (theta +
          theta_cv2 + 0.5 * (alpha+alpha_cv2)))/(1+exp(exp(-(beta+beta_cv2)/2) * (theta + theta_cv2 + 0.5 *
          (alpha+alpha_cv2))))) - log(exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) * (theta +
          0.5 * alpha))));
estimate "logRelative specificity cv level 1 vs 0" log(exp(-exp((beta+beta_cv1)/2) * (theta + theta_cv1
          - 0.5 * (alpha +alpha_cv1)))/(1+exp(-exp((beta+beta_cv1)/2) * (theta + theta_cv1 - 0.5 *
          (alpha+alpha_cv1))))) - log(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta -
          0.5 * alpha))));
estimate "logRelative specificity cv level 2 vs 0" log(exp(-exp((beta+beta_cv2)/2) * (theta + theta_cv2
          - 0.5 * (alpha +alpha_cv2)))/(1+exp(-exp((beta+beta_cv2)/2) * (theta + theta_cv2 - 0.5 *
          (alpha+alpha_cv2))))) - log(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta -
          0.5 * alpha))));
estimate "logRDOR cv level 1 vs 0" log(((exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 *
          (alpha + alpha_cv1)))/(1+exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
          alpha_cv1))))) /(1-(exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
          alpha_cv1)))/(1+exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
          alpha_cv1)))))) /((1-(exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1)))/(1+exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1))))) /(exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1)))/(1+exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
          alpha_cv1)))))) - log(((exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) *
```

```
                (theta + 0.5 * alpha)))) /(1-(exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) * (theta +
                0.5 * alpha)))))) /((1-(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5 *
                alpha))))) /(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5 * alpha))))));
        estimate "logRDOR cv level 2 vs 0" log(((exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 *
                (alpha + alpha_cv2)))/(1+exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 +
                0.5 * (alpha + alpha_cv2))))) /(1-(exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 *
                (alpha + alpha_cv2)))/(1+exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha +
                alpha_cv2)))))) /((1-(exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
                alpha_cv2)))/(1+exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
                alpha_cv2))))) /(exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
                alpha_cv2)))/(1+exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
                alpha_cv2)))))) - log(((exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) * (theta + 0.5 *
                alpha)))) /(1-(exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) * (theta + 0.5 *
                alpha))))) /((1-(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5 *
                alpha))))) /(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5 *
                alpha))))));
        estimate 'logLR+_0' log((exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) * (theta + 0.5 *
                alpha)))) /(1-(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5 *
                alpha))))));
        estimate "logLR+_1" log((exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
                alpha_cv1)))/(1+exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha + alpha_cv1)))))
                /(1-(exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
                alpha_cv1)))/(1+exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
                alpha_cv1))))))) ;
        estimate "logLR+_2" log((exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha +
                alpha_cv2)))/(1+exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha + alpha_cv2)))))
                /(1-(exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
                alpha_cv2)))/(1+exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
                alpha_cv2))))));
        estimate 'logLR-_0' log((1-(exp(exp(-beta/2) * (theta + 0.5 * alpha))/(1+exp(exp(-beta/2) *
                (theta + 0.5 * alpha))))) /(exp(-exp(beta/2) * (theta - 0.5 * alpha))/(1+exp(-exp(beta/2) * (theta - 0.5
                * alpha)))));
        estimate "logLR-_1" log((1-(exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha +
                alpha_cv1)))/(1+exp(exp(-(beta+beta_cv1)*0.5) * (theta + theta_cv1 + 0.5 * (alpha
                + alpha_cv1)))))) /(exp(-exp((beta+beta_cv1)*0.5) * (theta + theta_cv1 - 0.5 * (alpha +
                alpha_cv1)))/(1+exp(-exp(beta/2) * (theta + theta_cv1 - 0.5 * (alpha + alpha_cv1)))))) ;
        estimate "logLR-_2" log((1-(exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha +
                alpha_cv2)))/(1+exp(exp(-(beta+beta_cv2)*0.5) * (theta + theta_cv2 + 0.5 * (alpha
                + alpha_cv2)))))) /(exp(-exp((beta+beta_cv2)*0.5) * (theta + theta_cv2 - 0.5 * (alpha +
                alpha_cv2)))/(1+exp(-exp(beta/2) * (theta + theta_cv2 - 0.5 * (alpha + alpha_cv2)))))) ;
        estimate "alpha_1" alpha+alpha_cv1 ;
        estimate "alpha_2" alpha+alpha_cv2;
        estimate "theta_1" theta+theta_cv1 ;
        estimate "theta_2" theta+theta_cv2;
        estimate "beta_1" beta+beta_cv1 ;
        estimate "beta_2" beta+beta_cv2;
        contrast "Pooled test for alpha" alpha, alpha+alpha_cv1, alpha+alpha_cv2;
        contrast "Pooled test for theta" theta, theta+theta_cv1, theta+theta_cv2;
        contrast "Pooled test for beta" beta, beta+beta_cv1, beta+beta_cv2;
run;
```

d. Word Output

The .rtf document contains tables for model starting values, convergence status, fit and estimates for parameters and summary measures of test accuracy. Parameters for both the HSROC and bivariate models are included in this example because the input parameter *bothmodels=y*. The distributional assumptions for the random effects can be checked using the histograms and normal probability plots of the empirical Bayes estimates of the random effects that are produced with parameter *checkmod=y*. You can create your own plots if you choose to save the random effects to a data set with parameter *randeffs=y*. The Word document output is as follows:

META-ANALYSIS OF DIAGNOSTIC ACCURACY STUDIES

Analysis Information
Data: 'C:\Training\Galactomannan detection for invasive aspergillosis in immunocompromized patients.csv'
Test: 'Platelia - cutoff 1.0'
Confidence Interval: 95%

Covariate Information

Pat_base	Level
Episode-based	0
Patient-based	1
Patient-based data	2

HSROC model basic analysis for 'Platelia - cutoff 1.0'

Starting values

Parameter	Estimate	Gradient	LowerBC	UpperBC
alpha	3.0000	-1.48935	.	.
theta	0	4.640987	.	.
beta	0.5000	-1.7937	.	.
s2ua	1.0000	-0.3609	0	.
s2ut	0.5000	-0.44104	0	.

Convergence status

Reason	Status
NOTE: GCONV convergence criterion satisfied.	0

Model fit

Description	Value
-2 Log Likelihood	129.9
AIC (smaller is better)	139.9
AICC (smaller is better)	143.3
BIC (smaller is better)	142.3

HSROC model parameter estimates

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper	Gradient	RM_Name
alpha	3.3683	0.5515	6.11	<.0001	2.2861	4.4505	-3.14E-7	Lambda
theta	-0.5605	0.4381	-1.28	0.2011	-1.4202	0.2992	-5.93E-6	Theta
beta	0.04399	0.4724	0.09	0.9258	-0.8830	0.9710	6.167E-6	beta
s2ua	1.3297	0.8640	1.54	0.1241	-0.3657	3.0251	1.547E-6	Var(accuracy)
s2ut	0.6003	0.3826	1.57	0.1170	-0.1505	1.3511	3.988E-7	Var(threshold)

Bivariate model parameter estimates

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper
E(logitSe)	1.0992	0.3722	2.95	0.0032	0.3688	1.8297
E(logitSp)	2.2946	0.3119	7.36	<.0001	1.6826	2.9066
Var(logitSe)	0.8926	0.7346	1.22	0.2246	-0.5490	2.3342
Var(logitSp)	0.9747	0.4920	1.98	0.0478	0.009278	1.9401
Cov(logits)	-0.2679	0.4181	-0.64	0.5218	-1.0883	0.5525
Corr(logits)	-0.2872	0.3938	-0.73	0.4660	-1.0600	0.4855

Confidence and prediction region parameters

Parameter	Estimate
SE(E(logitSe))	0.3722
SE(E(logitSp))	0.3119
Cov(Es)	-0.0223
Studies	12.0000

Summary estimates of test accuracy measures

Parameter	Estimate	Lower	Upper
Sensitivity	0.7501	0.5912	0.8617
Specificity	0.9084	0.8432	0.9482
DOR	29.7795	12.6252	70.2423
LR+	8.1915	4.7221	14.2099
LR-	0.2751	0.1603	0.4720

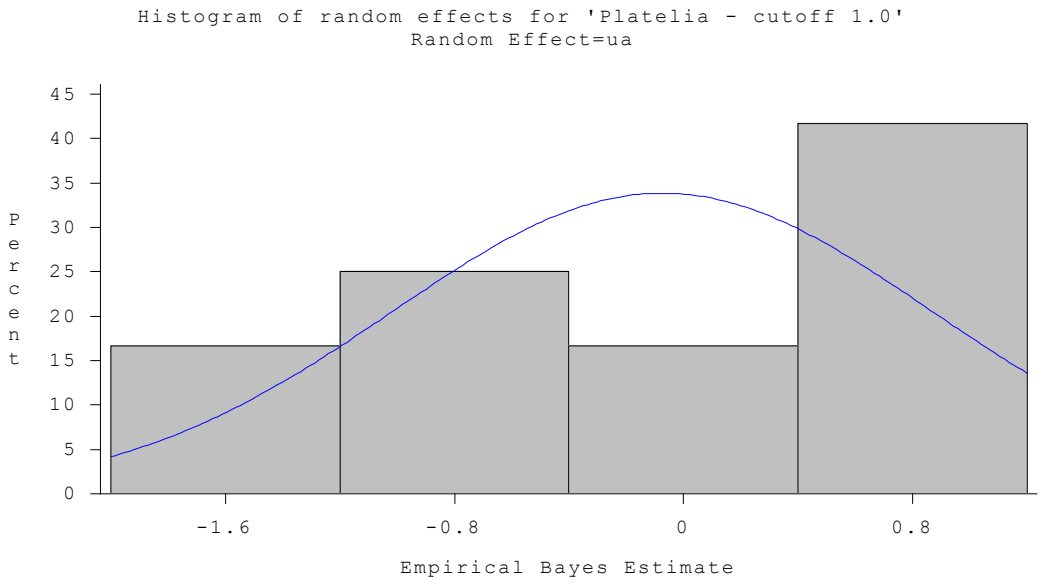
Predicted values of sensitivity and specificity based on parameter and empirical Bayes estimates

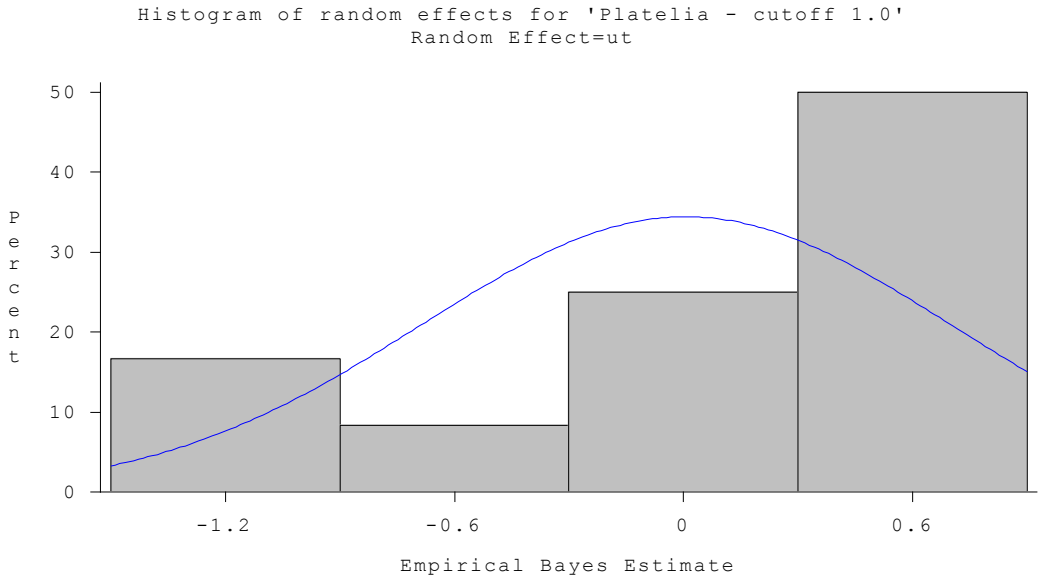
Study_ID	Pat_base	Observed sensitivity	Predicted sensitivity	Lower confidence limit for predicted sensitivity
Allan 2005	Episode-based	0.00000	0.55200	0.12668
Becker 2003	Patient-based data	0.46154	0.55750	0.31057
Bretagne 1998	Patient-based	0.77778	0.78085	0.57079
Busca 2006	Patient-based	1.00000	0.82596	0.42906
Challier 2004	Patient-based data	0.76923	0.77362	0.59567
Kawazu 2004	Episode-based	0.63636	0.65626	0.39277
Maertens 2002	Episode-based	0.84615	0.81030	0.57109
Marr 2004	Patient-based	0.54167	0.59306	0.39707
Pereira 2005	Patient-based	1.00000	0.81995	0.38423
Pinel 2003	Patient-based	0.50000	0.52220	0.36009
Suankratay 2006	Patient-based data	0.94118	0.87584	0.65279
Ulusakarya 2000	Patient-based	1.00000	0.90910	0.66330

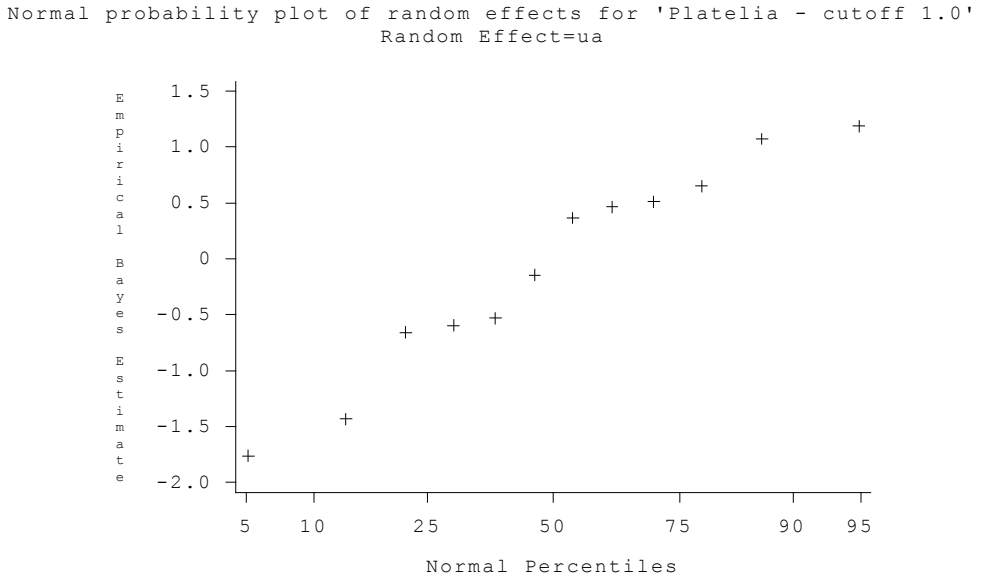
Upper confidence limit for predicted sensitivity	Observed specificity	Predicted specificity	Lower confidence limit for predicted specificity	Upper confidence limit for predicted specificity
0.91278	0.99194	0.97996	0.93807	0.99371
0.77894	0.83784	0.85021	0.75337	0.91340
0.90519	0.78261	0.81845	0.63506	0.92113
0.96771	0.83333	0.84097	0.74247	0.90654
0.88797	0.79545	0.81508	0.68228	0.90047
0.84929	0.97101	0.96406	0.92074	0.98411
0.93199	0.91954	0.91683	0.84373	0.95745
0.76331	0.74419	0.77703	0.63341	0.87545
0.97079	0.76316	0.78831	0.64007	0.88634
0.67976	0.97801	0.97638	0.96312	0.98495

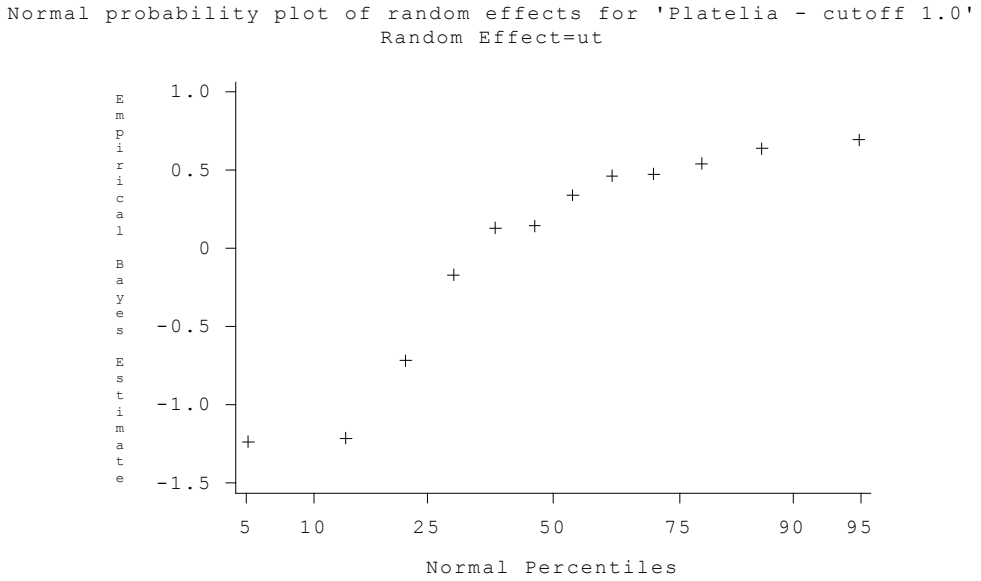
Model checking - distribution of random effects

Histograms and normal probability plots of the empirical Bayes estimates of the random effects (ua and ut, level two residuals)






Normal probability plot of random effects for 'Platelia - cutoff 1.0'
Random Effect=ua
Empirical Bayes Estimate
1.5
1.0
0.5
0
-0.5
-1.0
-1.5
-2.0
5
10
25
50
75
90
95
Normal Percentiles


Normal probability plot of random effects for 'Platelia - cutoff 1.0'
Random Effect=ut
Empirical Bayes Estimate
1.0
0.5
0
-0.5
-1.0
-1.5
5
10
25
50
75
90
95
Normal Percentiles

HSROC model analysis with covariate = Pat_base

Starting values

Parameter	Estimate	Gradient	LowerBC	UpperBC
alpha	3.0000	-1.89255	.	.
theta	-1.0000	-4.34687	.	.
beta	0	1.821256	.	.
s2ua	1.1650	-0.26655	0	.
s2ut	0.8000	0.200892	0	.
alpha_cv1	0	-0.13815	.	.
theta_cv1	0	-3.98331	.	.
beta_cv1	0	3.121569	.	.
alpha_cv2	0	0.041916	.	.
theta_cv2	0	-2.0781	.	.
beta_cv2	0	2.258011	.	.

Convergence status

Reason	Status
NOTE: GCONV convergence criterion satisfied.	0

Model fit

Description	Value
-2 Log Likelihood	123.5
AIC (smaller is better)	145.5
AICC (smaller is better)	167.5
BIC (smaller is better)	150.8

Likelihood ratio test for model with and without Pat_base:
-2 log likelihood difference is 6.444 on 6 degrees of freedom, p = 0.375

Predicted values of sensitivity and specificity based on parameter and empirical Bayes estimates

Study_ID	Pat_base	Observed sensitivity	Predicted sensitivity	Lower confidence limit for predicted sensitivity
Allan 2005	Episode-based	0.00000	0.57351	0.12242
Becker 2003	Patient-based data	0.46154	0.51628	0.26109
Bretagne 1998	Patient-based	0.77778	0.77945	0.57012
Busca 2006	Patient-based	1.00000	0.80786	0.41779
Challier 2004	Patient-based data	0.76923	0.77396	0.58461
Kawazu 2004	Episode-based	0.63636	0.65792	0.38076
Maertens 2002	Episode-based	0.84615	0.79587	0.52230
Marr 2004	Patient-based	0.54167	0.61094	0.40569
Pereira 2005	Patient-based	1.00000	0.80692	0.39372
Pinel 2003	Patient-based	0.50000	0.52944	0.36344
Suankratay 2006	Patient-based data	0.94118	0.89963	0.63945
Ulusakarya 2000	Patient-based	1.00000	0.88585	0.56692

Upper confidence limit for predicted sensitivity	Observed specificity	Predicted specificity	Lower confidence limit for predicted specificity	Upper confidence limit for predicted specificity
0.92838	0.99194	0.98512	0.94247	0.99627
0.76325	0.83784	0.84969	0.75990	0.90989
0.90401	0.78261	0.81443	0.63675	0.91659
0.96099	0.83333	0.83832	0.74108	0.90379
0.89283	0.79545	0.82395	0.70020	0.90365
0.85747	0.97101	0.97133	0.93181	0.98824
0.93290	0.91954	0.93238	0.85472	0.96999
0.78319	0.74419	0.77970	0.63553	0.87780
0.96415	0.76316	0.78751	0.64177	0.88462
0.68917	0.97801	0.97501	0.96102	0.98406

HSROC model parameter estimates for Pat_base = Episode-based

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper
alpha	4.1671	1.4313	2.91	0.0037	1.3583	6.9759
theta	-1.2665	1.3129	-0.96	0.3349	-3.8428	1.3098
beta	0.09235	1.1400	0.08	0.9355	-2.1446	2.3293
s2ua	0.8672	0.6936	1.25	0.2115	-0.4938	2.2282
s2ut	0.4048	0.2579	1.57	0.1169	-0.1014	0.9110

Bivariate model parameter estimates for Pat_base = Episode-based

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper
E(logitSe)	0.7802	0.7057	1.11	0.2692	-0.6046	2.1649
E(logitSp)	3.5083	0.5914	5.93	<.0001	2.3478	4.6688
Var(logitSe)	0.5668	0.7749	0.73	0.4647	-0.9538	2.0873
Var(logitSp)	0.6818	0.7900	0.86	0.3883	-0.8685	2.2320
Cov(logits)	-0.1880	0.2735	-0.69	0.4920	-0.7248	0.3487
Corr(logits)	-0.3025	0.4057	-0.75	0.4561	-1.0985	0.4936

Confidence and prediction region parameters for Pat_base = Episode-based

Parameter	Estimate
SE(E(logitSe))	0.70566
SE(E(logitSp))	0.59138
Cov(Es)	-0.05355
Studies	3.00000

HSROC model parameter estimates for Pat_base = Patient-based

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper
alpha_1	3.0755	0.7217	4.26	<.0001	1.6593	4.4916
theta_1	-0.3634	0.4870	-0.75	0.4557	-1.3192	0.5923
beta_1	0.05490	0.6258	0.09	0.9301	-1.1732	1.2830
s2ua	0.8672	0.6936	1.25	0.2115	-0.4938	2.2282
s2ut	0.4048	0.2579	1.57	0.1169	-0.1014	0.9110

Bivariate model parameter estimates for Pat_base = Patient-based

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper
E(logitSe)_1	1.1425	0.5170	2.21	0.0273	0.1280	2.1570
E(logitSp)_1	1.9541	0.3643	5.36	<.0001	1.2392	2.6689
Var(logitSe)	0.5668	0.7749	0.73	0.4647	-0.9538	2.0873
Var(logitSp)	0.6818	0.7900	0.86	0.3883	-0.8685	2.2320
Cov(logits)	-0.1880	0.2735	-0.69	0.4920	-0.7248	0.3487
Corr(logits)	-0.3025	0.4057	-0.75	0.4561	-1.0985	0.4936

Confidence and prediction region parameters for Pat_base = Patient-based

Parameter	Estimate
SE(E(logitSe))	0.51698
SE(E(logitSp))	0.36428
Cov(Es)	-0.02934
Studies	6.00000

HSROC model parameter estimates for Pat_base = Patient-based data

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper
alpha_2	3.7879	1.4819	2.56	0.0107	0.8799	6.6959
theta_2	-1.1864	1.0938	-1.08	0.2783	-3.3328	0.9600
beta_2	-1.0961	1.1459	-0.96	0.3390	-3.3448	1.1525
s2ua	0.8672	0.6936	1.25	0.2115	-0.4938	2.2282
s2ut	0.4048	0.2579	1.57	0.1169	-0.1014	0.9110

Bivariate model parameter estimates for Pat_base = Patient-based data

Parameter	Estimate	Standard Error	z	Pr > |z|	Lower	Upper
E(logitSe)_2	1.2240	0.8760	1.40	0.1626	-0.4949	2.9430
E(logitSp)_2	1.7807	0.3646	4.88	<.0001	1.0651	2.4962
Var(logitSe)	0.5668	0.7749	0.73	0.4647	-0.9538	2.0873
Var(logitSp)	0.6818	0.7900	0.86	0.3883	-0.8685	2.2320
Cov(logits)	-0.1880	0.2735	-0.69	0.4920	-0.7248	0.3487
Corr(logits)	-0.3025	0.4057	-0.75	0.4561	-1.0985	0.4936

Confidence and prediction region parameters for Pat_base = Patient-based data

Parameter	Estimate
SE(E(logitSe))	0.87597
SE(E(logitSp))	0.36464
Cov(Es)	-0.06665
Studies	3.00000

Summary estimates of test accuracy measures for covariate Pat_base

Parameter	Estimate	Lower	Upper
Sensitivity for cv level 0	0.6857	0.3533	0.8971
Sensitivity for cv level 1	0.7581	0.5320	0.8963
Sensitivity for cv level 2	0.7728	0.3787	0.9499
Specificity for cv level 0	0.9709	0.9128	0.9907
Specificity for cv level 1	0.8759	0.7754	0.9352
Specificity for cv level 2	0.8558	0.7437	0.9239
DOR_0	72.8589	13.4609	394.36
DOR_1	22.1219	7.0299	69.6138
DOR_2	20.1792	3.6185	112.53
LR+_0	23.5838	7.4332	74.8258
LR+_1	6.1086	3.2343	11.5373
LR+_2	5.3581	2.7860	10.3049
LR-_0	0.3237	0.1257	0.8335
LR-_1	0.2852	0.03290	2.4728
LR-_2	1.0025	0.006519	154.15

Estimates of relative measures of test accuracy for covariate Pat_base

Parameter	Estimate	Pr > |z|	Lower	Upper
True positive odds ratio cv level 1 vs 0	1.4156	0.7432	0.1766	11.3482
True positive odds ratio cv level 2 vs 0	0.8274	0.9251	0.01591	43.0255
True negative odds ratio cv level 1 vs 0	0.2256	0.4889	0.003311	15.3641
True negative odds ratio cv level 2 vs 0	0.8557	0.9207	0.03972	18.4328
Relative sensitivity cv level 1 vs 0	1.1056	0.7003	0.6627	1.8444
Relative sensitivity cv level 2 vs 0	1.1269	0.6885	0.6279	2.0226
Relative specificity cv level 1 vs 0	0.9021	0.0335	0.8204	0.9920
Relative specificity cv level 2 vs 0	0.8814	0.0224	0.7909	0.9822
RDOR cv level 1 vs 0	0.3036	0.2575	0.03853	2.3927
RDOR cv level 2 vs 0	0.2770	0.2948	0.02504	3.0633

Test of difference between levels of Pat_base

Test	Num DF	Den DF	F Value	Pr > F
Pooled test for alpha	3	1000	10.93	<.0001
Pooled test for theta	3	1000	0.91	0.4341
Pooled test for beta	3	1000	0.31	0.8204

References

1. Rutter CM, Gatsonis CA. hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001; 20:2865-84.

2. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005; 58:982-90.

3. Harbord RM, Deeks JJ, Whiting P, Sterne JAC. A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics. 2006; 1:1-21.

4. Macaskill P. Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis. J Clin Epidemiol. 2004; 57:925-32.

5. Wolfinger RD. Fitting Nonlinear Mixed Models with the New NLMIXED Procedure. SUGI Proceedings. 1999

6. Flom PL, McMahon JM, Pouget ER. Using PROC NLMIXED and PROC GLMMIX to analyze dyadic data with a dichotomous dependent variable. SAS Global Forum 2007. Paper 179.

7. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol. 2006; 59:1331-32.

8. Riley RD, Abrams KR, Sutton AJ, Lambert PC, Thompson JR. Bivariate random-effects meta-analysis and the estimation of between-study correlation. BMC Med Res Methodol. 2007; 7:3.

9. SAS Institute Inc. 2004. SAS OnlineDoc® 9.1.3. Cary, NC: SAS Institute Inc.

10. Patefield M. Fitting non-linear structural relationships using SAS procedure NLMIXED. The Statistician. 2002; 51:355-66.