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

Quick reference and worked example

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. Quick reference and worked example Version 1.3. 2010 July. Available from: http://srdta.cochrane.org/.

Contents

1 MetaDAS

1.1 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. 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). 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.

1.2 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 and they 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 or Stata data input file or SAS data set if data import is not required (import=n), and the Word file (.rtf) for the analysis output.

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 (1). 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 (2). 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/ abt/se/sp/sesp	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 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 2 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*

1.3 Examples

1.3.1 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;
```

1.3.2 Some more options included

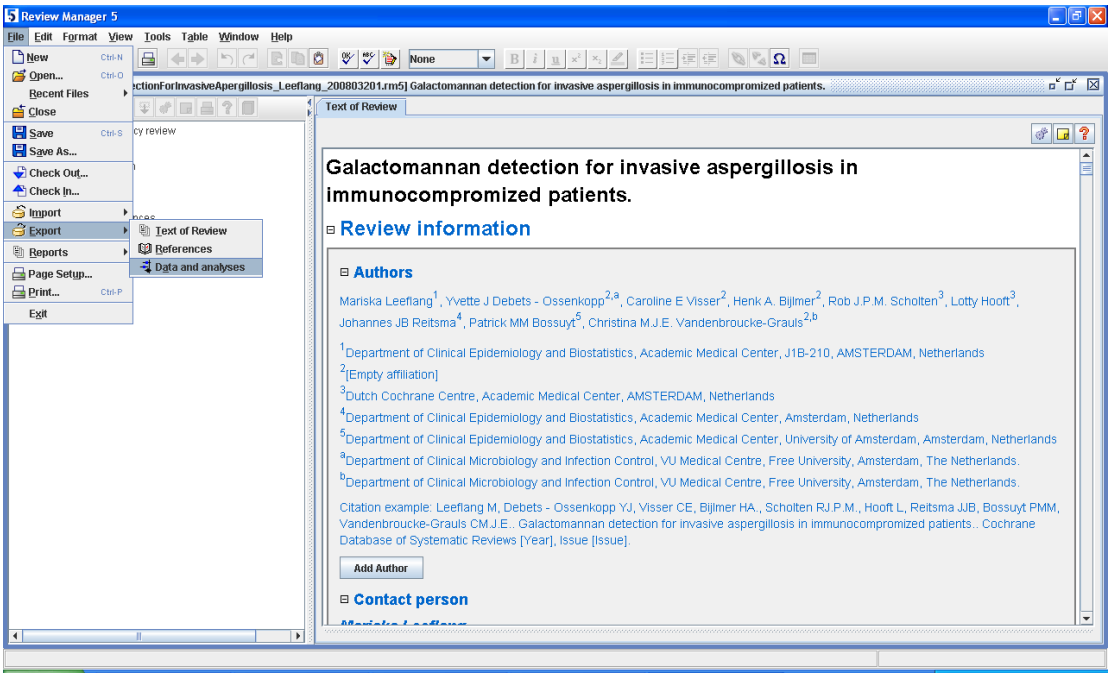
```
%metadas(dtfile= 'C:\Revman Test Data.xls',
      tech=newrap,
      covariate=stage,
      byvar=test_type,
      cveffect=a,
      test='HPV',
      predict=y,
      debug=n,
      rfile ='c:\hsroc test.rtf');
run;
```

```
%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;
```

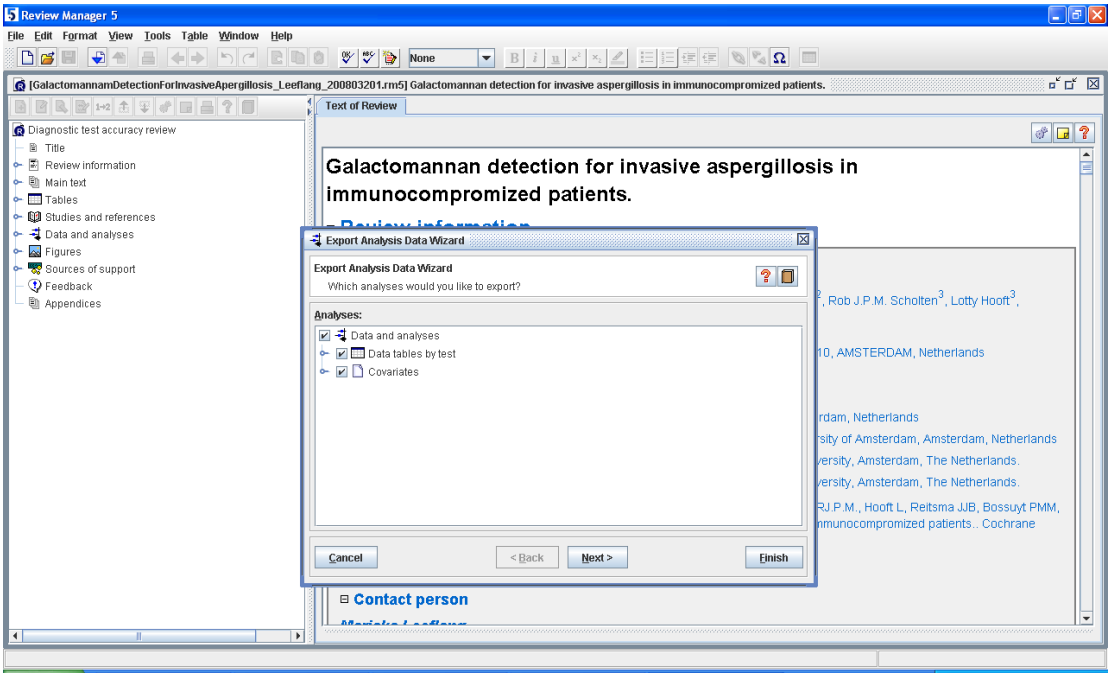
2 *Worked example*

2.1 File export from RevMan

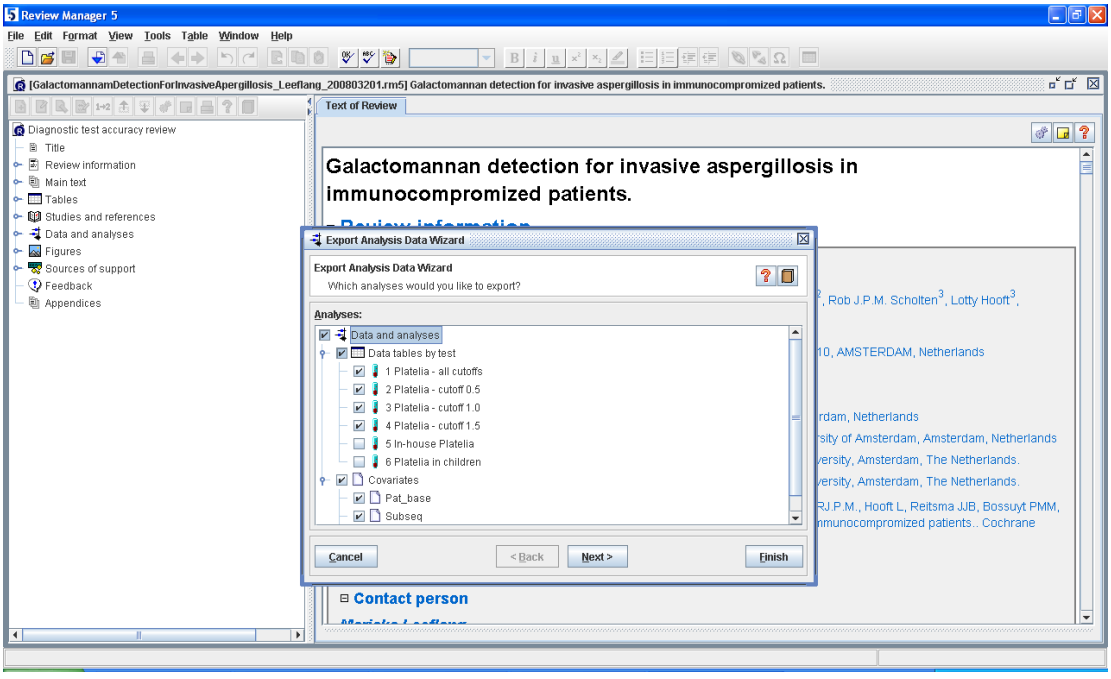
Open your RevMan file. On the menu bar, click on file and then click on export and select data and analyses.



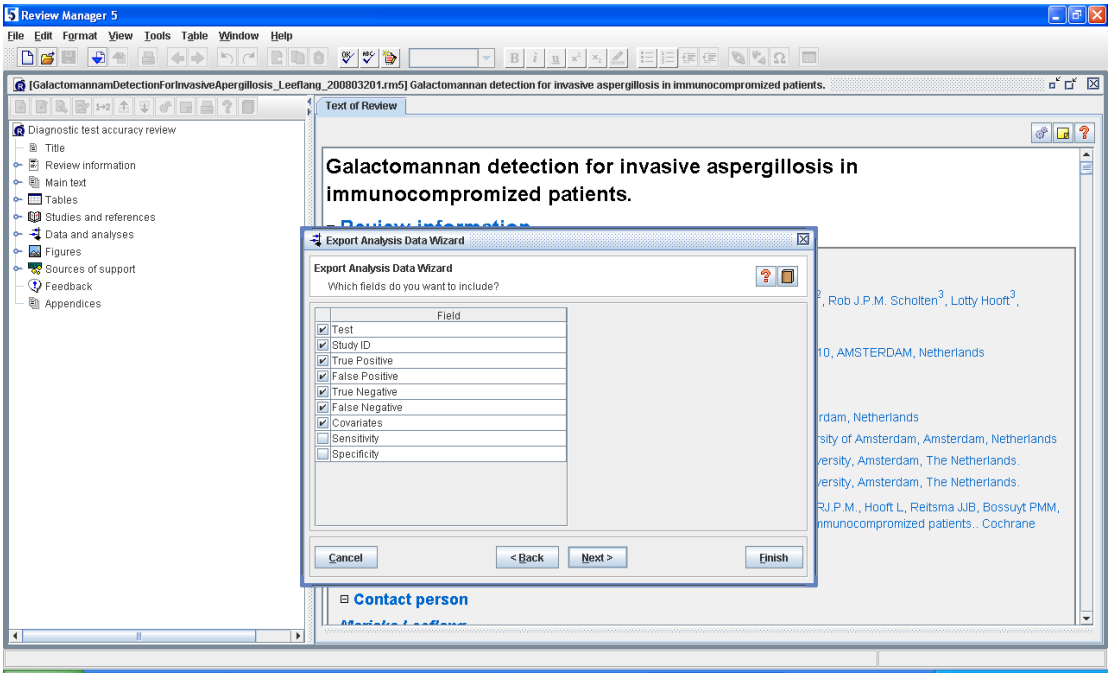
The export analysis data wizard is launched.



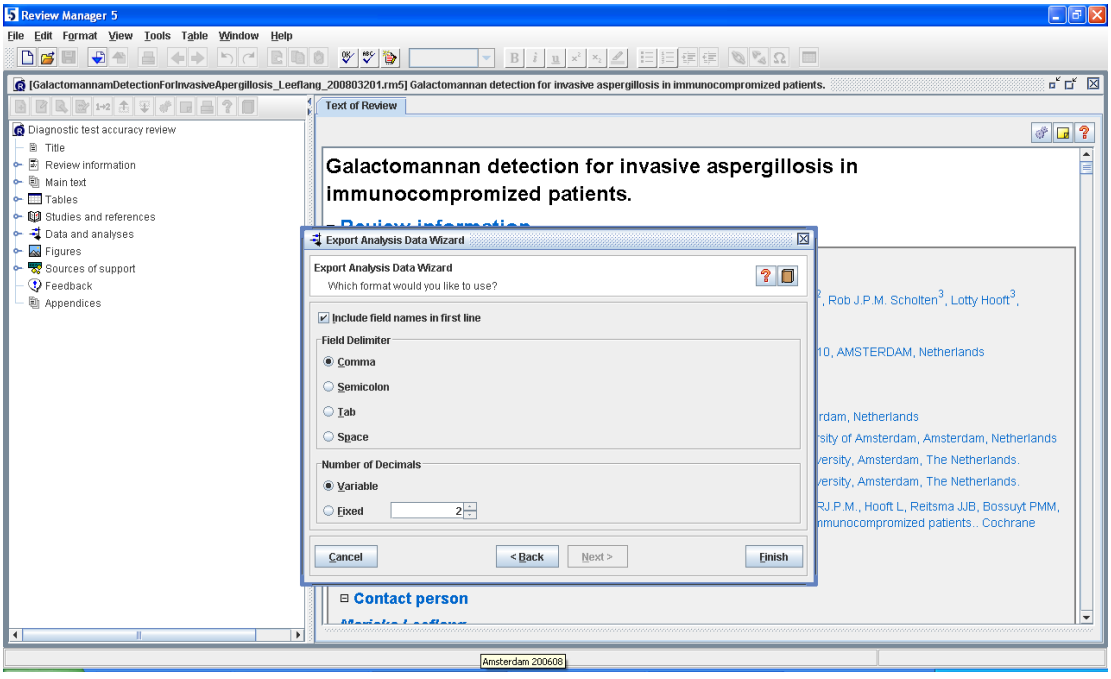
If you do not wish to select specific data tables by test or covariates then click next. If you wish to select then expand the tree and make your selection as shown below.



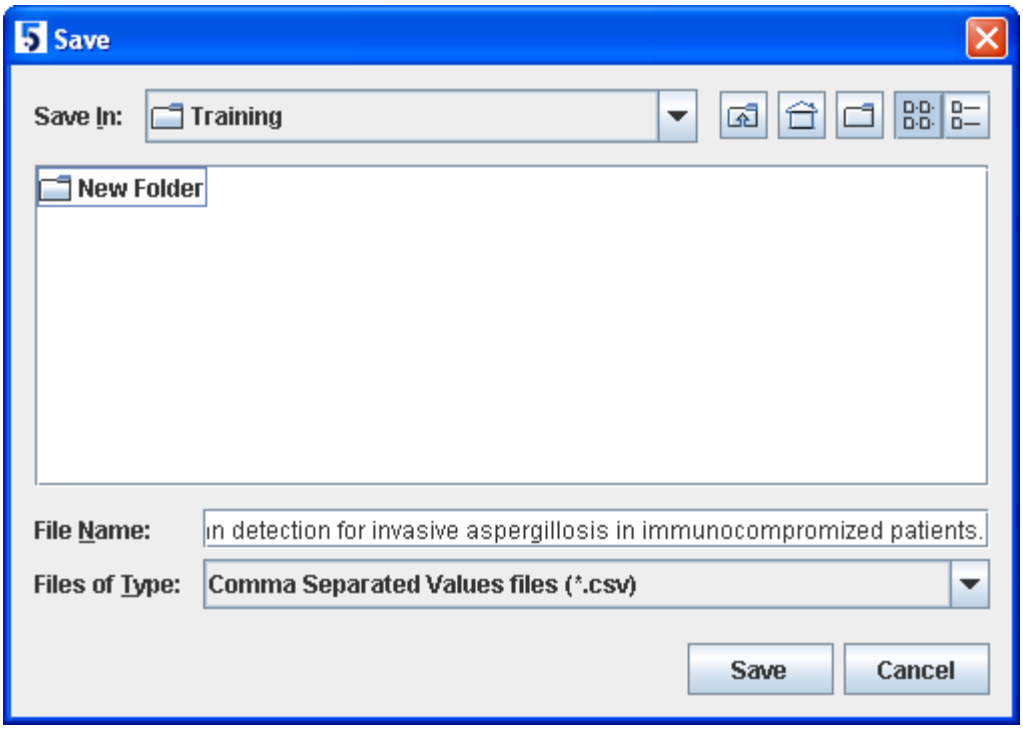
On the next page of the wizard select the fields you wish to export. Click on next.



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. Name your file as you wish (ensure that the .csv extension is included) and click save.



A sample of the extracted data is shown in figure 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 1 Sample of data from the Excel .csv file

2.2 Run *MetaDAS*

2.2.1 Sample SAS statements to run the macro

```
%include 'C:\My SAS programs\Metadas v1.3 beta.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.

2.2.2 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 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 2 Log content with input parameter *logfile*

b. Data import

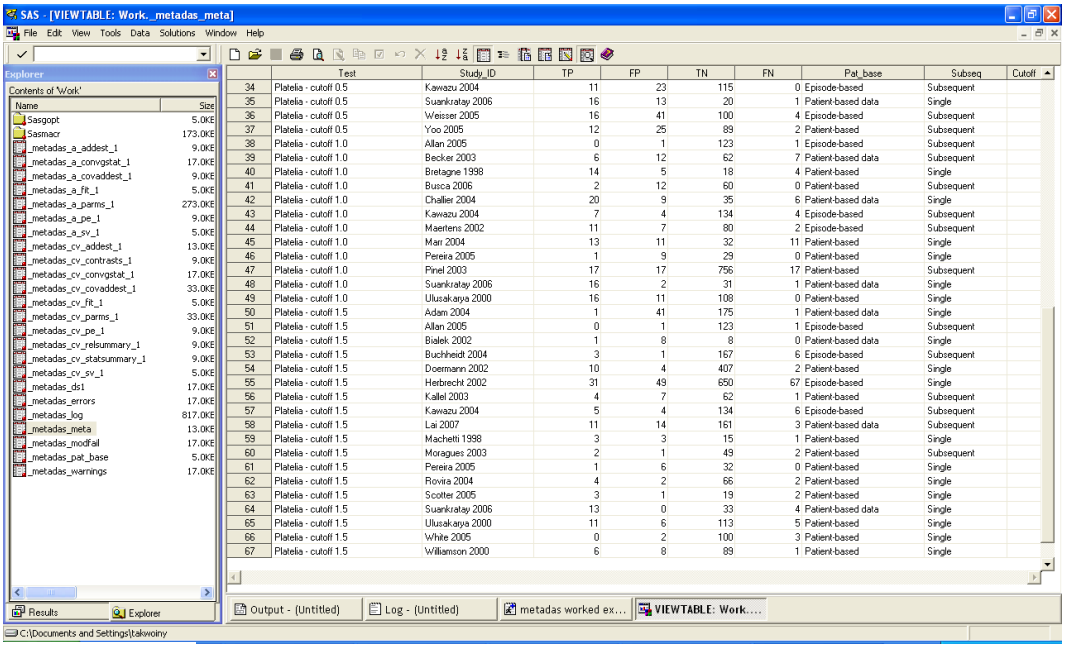
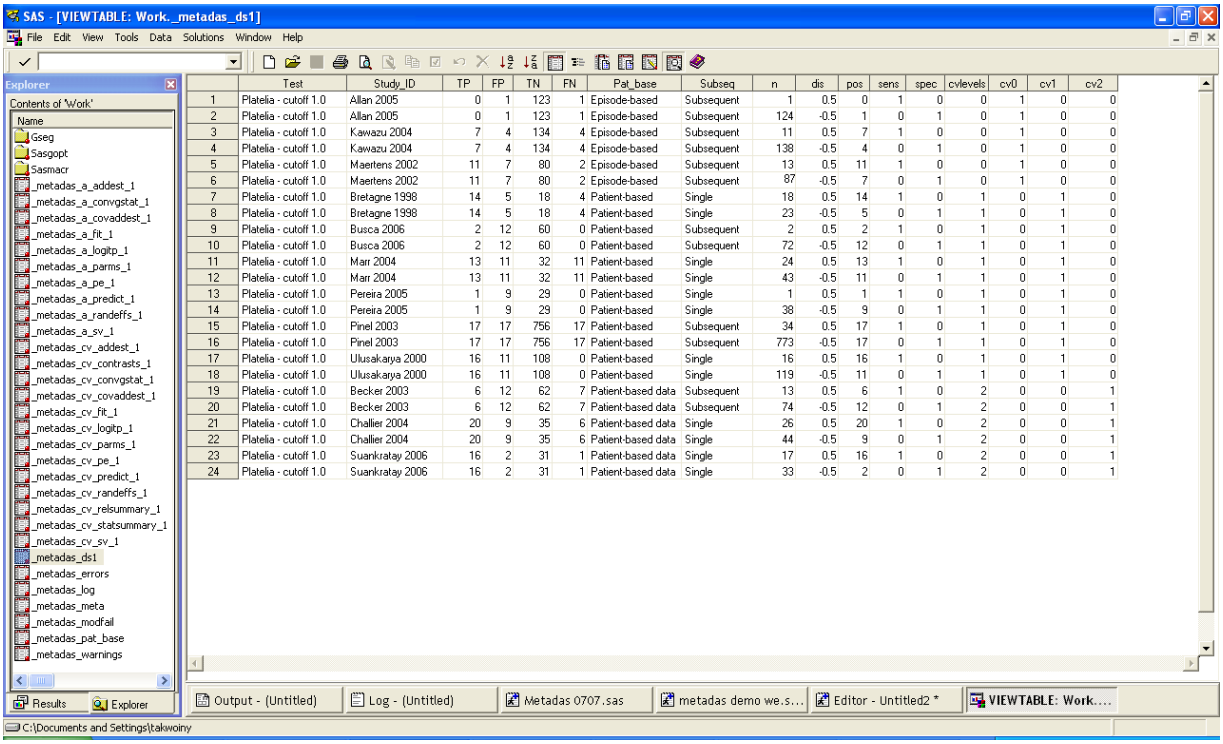


Figure 3 Sample of contents of table _metadas_meta

The data set _metadas_meta contains all the data from the input data file without any modification as shown in figure 3. Figure 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 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.

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

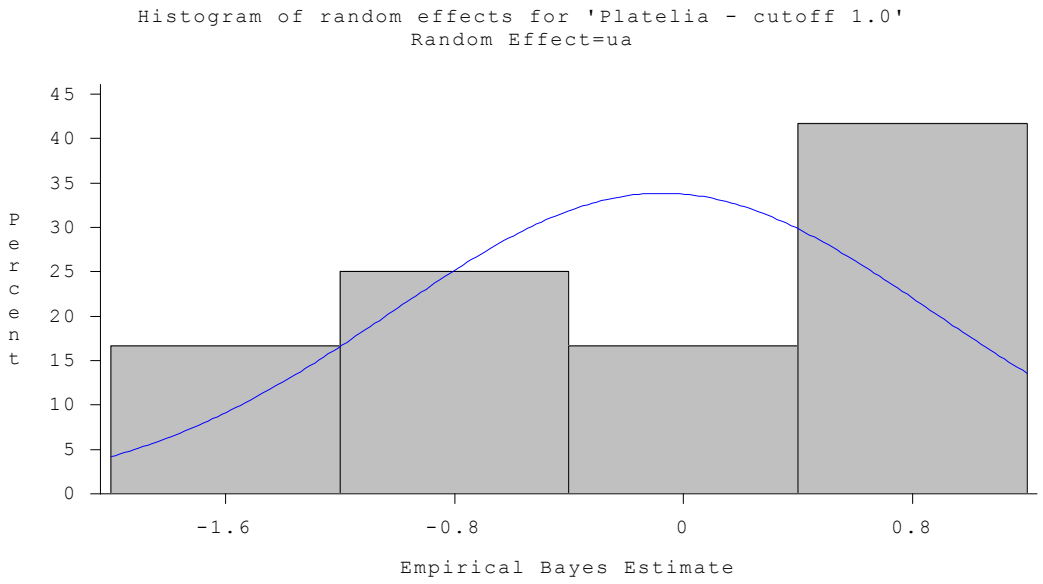
Predicted values of sensitivity and specificity based on parameter and empirical Bayes estimates

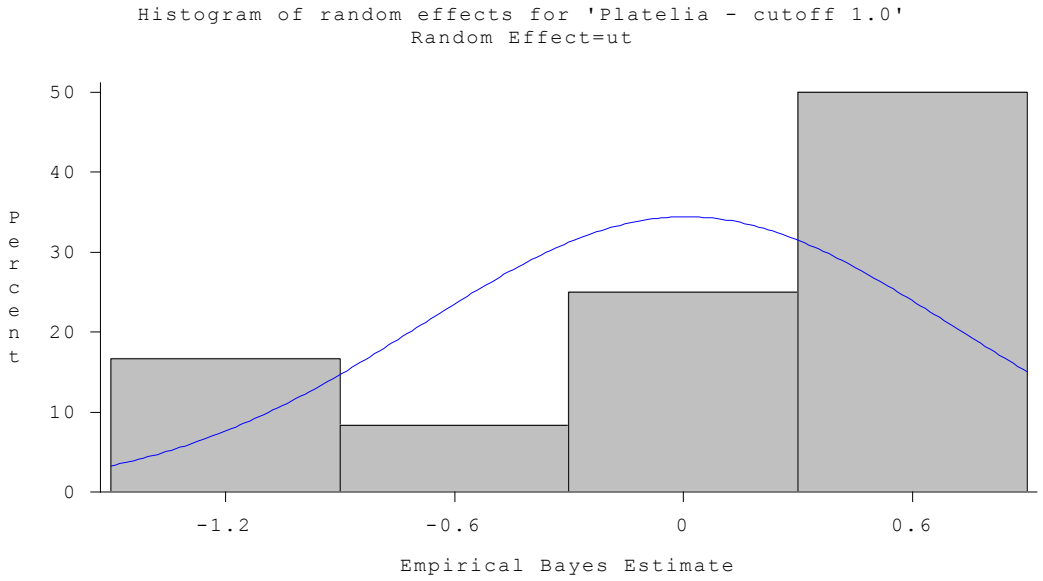
Study_ID	Pat_base	Observed sensitivity	Predicted sensitivity	Lower confidence limit for predicted sensitivity
Allan 2005	Episode-based	0.00000	0.55200	0.12668
Becker 2003	Patient-based data	0.46154	0.55750	0.31057
Bretagne 1998	Patient-based	0.77778	0.78085	0.57079
Busca 2006	Patient-based	1.00000	0.82596	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

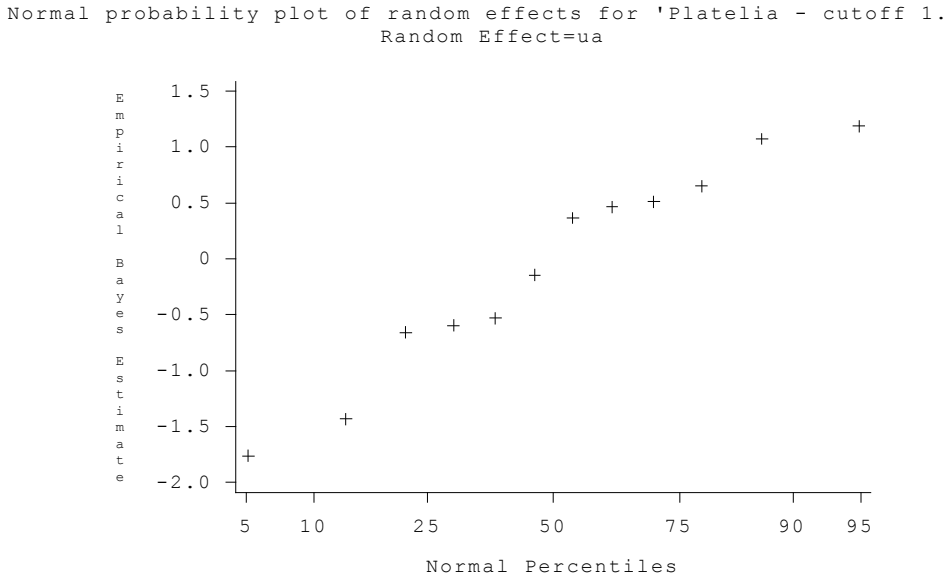
Upper confidence limit for predicted sensitivity	Observed specificity	Predicted specificity	Lower confidence limit for predicted specificity	Upper confidence limit for predicted specificity
0.91278	0.99194	0.97996	0.93807	0.99371
0.77894	0.83784	0.85021	0.75337	0.91340
0.90519	0.78261	0.81845	0.63506	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

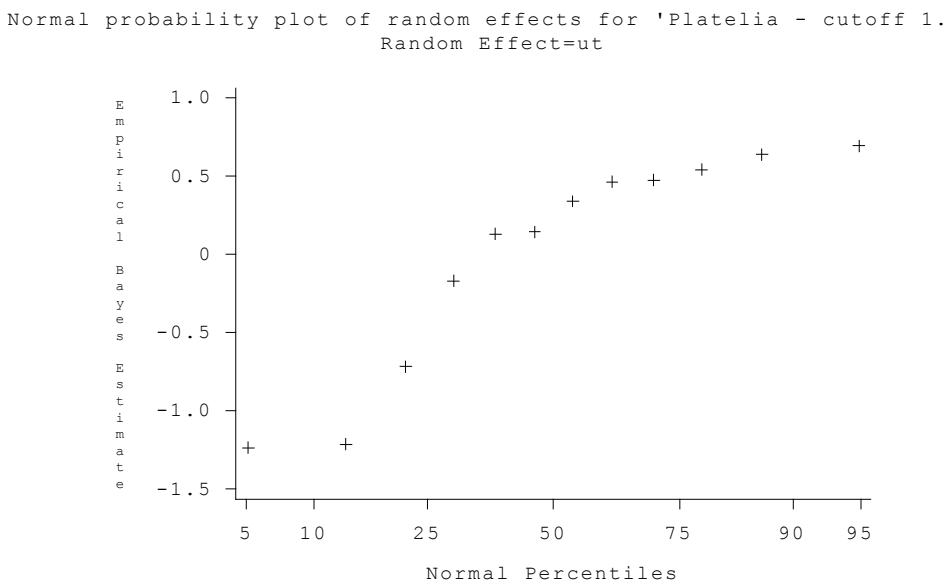
Model checking - distribution of random effects

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

References

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

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