The SAS METAANAL Macro

Ellen Hertzmark and Donna Spiegelman

March 31, 2017

Abstract

The %METAANAL macro is a SAS version 9 macro that produces the DerSimonian-Laird estimators for random effects or fixed effects models in pooled or metaanalysis. It can be used to pull results from two or three of the Channing cohorts and test for between-studies heterogeneity.

Keywords: SAS, macro, metaanalysis, DerSimonian-Laird, inhomogeneity, pooling, fixed effects model, random effects model

Contents

1   Description

This macro produces the Laird and DerSimonian estimators for fixed and random effects models in meta- or pooled analysis. It can be used to pull results from two or three of the Channing cohorts and test for between-studies heterogeneity. Depending on the type of pooling, the studies are weighted proportionately to the inverse of the sum of the study specific variance plus the common between-studies variance (random effects pooling) or the inverse of the study-specific variance (fixed effects pooling).

In using studies from the literature, the macro can take data in the form of 95% confidence intervals or coefficients with standard errors or variances (but not both). The macro can also handle continuous outcomes.
The macro does its calculations for one coefficient at a time.

2   Invocation and Details

To call %METAANAL, your program must know where to look for it. In the Channing Lab, the most efficient way is to include

```
options mautosource sasautos='/usr/local/channing/sasautos';
```

at the top of your program.

Below, values to the right of the '=' are the default values for the parameters, if any.

You type

```
%metaanal(
REQUIRED PARAMETERS
===================
  DATA =,   name of the dataset containing the coefficients
            and their variances (or however you are providing the data)
            It is desirable that the coefficients and
            their variances have at least 5 significant digits.
            If you are using values from the literature,
            do the best you can.
            REQUIRED

  STUDYLAB = ,  name of the variable for the labels of the studies.
               REQUIRED

  EXPLABEL = ,  descriptive title of exposure
               REQUIRED

  OUTCOMELABEL = ,  descriptive title of outcome
                   REQUIRED

OPTIONAL PARAMETERS
===================
```

```
COEFF=T,  T=results are given as coefficients and standard errors
          or variances;
          F=results are given as 95% confidence intervals.
          OPTIONAL


REQUIRED IF COEFF=T
===================
BETA = beta,  if COEFF=T, the name of the variable for the coefficients
              This parameter must contain a variable name.


SE_OR_VAR =V,  if COEFF=T, whether you are giving the macro the standard
               errors (S) or the variances (V) of the coefficients


VAR = var,  if COEFF=T, the name of the variable for the variances
            of the coefficients
            This parameter must contain a variable name.
            REQUIRED if COEFF is T and SE_OR_VAR is V.


SE  = se,  if COEFF=T, the name of the variable for the standard errors
           of the coefficients
           This parameter must contain a variable name.
           REQUIRED if COEFF is T and SE_OR_VAR is S.


REQUIRED IF COEFF=F
===================
VARNAMELCB = if COEFF=F, the name of the variable for the
             lower 95% confidence bound
             REQUIRED if COEFF is F


VARNAMEUCB = if COEFF=F, the name of the variable for the upper 95%
             confidence bound
             REQUIRED if COEFF is F

OPTIONAL PARAMETERS
===================
POOLTYPE = random,  method for pooling (random, fixed, or both)
                    The default has been set to random, as that
                    is what we recommend.
                    OPTIONAL


WT = 1,  The difference/change in the exposure for which to compute
         the odds ratios.
         If no value is given, the macro will set WT = 1.
         OPTIONAL


OUTDAT =,  name of a data set to which the results of %METAANAL
           will be output, if any.  If no name is given, no
           data set will be made.
           OPTIONAL
```

```
   NOPRINT = F,   F=print the output, T=do not print the output
                  You might want to use NOPRINT=F if you are making
                  an output data set (i.e. outdat has a value).
                  If OUTDAT does not have a value, the macro will
                  set NOPRINT=F and print a WARNING.
                  default=F
                  OPTIONAL


   LOGLINEAR = T,   T=results are from loglinear models
                    (logistic, phreg, log-binomial, poisson),
                    F=results are from linear models
                    If LOGLINEAR is T, the macro will display the pooled odds ratio or risk ratio
                    and its 95% confidence interval.
                    OPTIONAL


   PRINTCOEFF = F,   if LOGLINEAR is T, F=don't print the coefficients, only the RR/ORs.
                                        T=print the coefficients IN ADDITION TO the RR/ORs.
                     OPTIONAL


   PVALUEFORMAT = pvalue6.4   Any legitimate SAS format can be used.
                              Pvalue6.4 is sufficient for most purposes, but
                              can be overridden if you need more precision, e.g.
                              from GWAS.
                              OPTIONAL
);
```

3 Examples

Below we give examples of %METAANAL using data from an analysis pooling NHS and HPFS data fron 1986 to 2000. The outcome of interest is confirmed lung cancer among subjects with known smoking status. There were 883 incident lung cancer cases among 330129 person-periods in NHS and 272 among 125371 person-periods in HPFS. For simplicity, the model contains only 3 variables: current smoking (currsmoke) early age (under 15) at start of smoking (agestr1), and obesity (bmige30). As usual, the Cox proportional hazards model is stratified by age in months and time period.
We consider the effect of obesity (bmi ge 30).
The analysis was done with the following code:

```
proc phreg data=alldat outest=men;
model ptime*lung19(0)=currsmoke agestr1 bmige30;
where gender eq 1 and bmim ne 1;
run;

proc phreg data=alldat outest=women;
model ptime*lung19(0)=currsmoke agestr1 bmige30;
where gender eq 2 and bmim ne 1;
run;
```

The dataset, OBESE, used in the calls to %METAANAL is

```
Obs      beta          se        p      study

1      -0.36254     0.11010     .001    nhs
2      -0.03305     0.22174     .88     hpfs
```

It was made by appending `MEN` to `WOMEN`.

3.1 Example 1. Obesity and lung cancer

The macro call was

```
title2 'example 1';
%metaanal(data=obese, beta=estimate, se_or_var=s, se=stderr, studylab=study,
loglinear=t, coeff=t,
printcoeff=t,
explabel=obesity (BMI ge 30), outcomelabel=lung cancer);
```

Note that we gave the macro the standard errors of the coefficients, rather than the variances, because this was is what we got in the output data set from `PROC PHREG`. The output was

----

```
The SAS System                                                                              1
example 1

Studies of obesity (BMI ge 30) and lung cancer
listing of input data

                                         weight in    weight in
                            std error      fixed        random
                Parameter    of study     effects      effects
Obs    study    Estimate    coefficient    model        model

1      NHS      -0.36254     0.11010       0.80222      0.67063
2      HPFS     -0.03305     0.22174       0.19778      0.32937
```

```
The SAS System                                                                              1
example 1

Studies of obesity (BMI ge 30) and lung cancer

Obs    Statistic    Value (95% CI)           P            Hypothesis being tested

3      OR/RR (R)    0.78 (0.57, 1.05)      0.1009    Is OR/RR  different from 1? (Random effects mo
4      Coeff. (R)   -0.25 (-0.56, 0.05)    0.1009    Is coeff. different from 0? (Random effects mo
5      Q            1.77                   0.1832    Is there heterogeneity among the studies?
```

```
6     tau2           0.0236                   .
7     r(i)           54.9 (0.0, 100.0)        .
8     CVB            0.605 (0.000, 2.663)     .
9     r(b)           49.3 (0.0, 100.0)        .
```

----

example 1

Studies of obesity (BMI ge 30) and lung cancer

```
Obs     Statistic     Description

 3      OR/RR (R)     OR/RR from random effects model (95% CI)
 4      Coeff. (R)    Coefficient from random effects model
 5      Q             Q-statistic for heterogeneity with 1 degree(s) of freedom
 6      tau2          Estimate of between-studies variance (tau squared)
 7      r(i)          Between-studies variance as a percent of total variance (r(i))
 8      CVB           SQRT(Between-studies variance)/coeff. from random-effects model (CVB)
 9      r(b)          New between-studies variance as a percent of total variance (r(b))
```

The title given by the macro used the parameters EXPLABEL and OUTCOMELABEL, so you will know
what the analysis was about when you look at it 6 months from now. Then the macro listed the
input data. This listing also shows the weights for the fixed effects and random effects models. The
big white space before the title is there to allow room for up to 3 user titles.
Then the macro gave the results of the random effects model (the default).

3.2   Example 2. Errors in the macro call

3.2.1   Example 2a. Did not specify SE_OR_VAR

In this example, we did not set se_or_var to S, so the macro used the default (v). However, there
is no variable VAR (the default name for the coefficient's variance) in the input dataset OBESE.


```
title2 'example 2';
%metaanal(data=obese, beta=estimate, se=stderr, studylab=study,
loglinear=t, coeff=t,
explabel=obesity (BMI ge 30), outcomelabel=lung cancer);
```

In the .log file, we got the following

----

example 2

```
There is not enough information to do the metaanalysis
All values of the variable VAR are missing.
All values of the variable STDERR are missing.
```

Note that the macro thought that all values of STDERR were missing, even though the input dataset
has non-missing values. This is because the macro tried to compute STDERR from the missing
default variable VAR, and so replaced the original values with missings.

3.2.2   Example 2b. Omitted one of the required parameters

The macro call was

```
title2 'example 3';
%metaanal(data=obese, beta=estimate, studylab=study,
loglinear=t, coeff=t,
printcoeff=t,
explabel=obesity (BMI ge 30));
```

The call did not specify OUTCOMELABEL. This resulted in an error message in the .LOG and the
.LST.

3.3   Example 3. Using the confidence limits instead of the coefficient and standard error

For purposes of this example, we calculated the confidence limits from the coefficients and the
standard error, then gave these to the macro instead of the coeff, using *COEFF=F* and giving the
values for the upper and lower confidence limits. The macro call is

```
title2 'example with confidence intervals';
title3 'example 4';
%metaanal(data=obese, coeff=f, loglinear=t, varnamelcb=lcl, varnameucb=ucl,
studylab=study,
pooltype=both,
printcoeff=t,
explabel=obesity (BMI ge 30), outcomelabel=lung cancer);
```

and the output is

```
The SAS System                                                                                          1
example with confidence intervals
example 4
Studies of obesity (BMI ge 30) and lung cancer
listing of input data
```

7

Obs	study	study coefficient	std error of study coefficient	weight in fixed effects model	weight in random effects model
1	NHS	-0.36254	0.11010	0.80222	0.67063
2	HPFS	-0.03305	0.22174	0.19778	0.32937

----

example with confidence intervals
example 4
Studies of obesity (BMI ge 30) and lung cancer

Obs	Statistic	Value (95% CI)	P	Hypothesis being tested
1	OR/RR (F)	0.74 (0.61, 0.90)	0.0026	Is OR/RR different from 1? (Fixed effects mod
2	Coeff. (F)	-0.30 (-0.49, -0.10)	0.0026	Is coeff. different from 0? (Fixed effects mod
3	OR/RR (R)	0.78 (0.57, 1.05)	0.1009	Is OR/RR different from 1? (Random effects mo
4	Coeff. (R)	-0.25 (-0.56, 0.05)	0.1009	Is coeff. different from 0? (Random effects mo
5	Q	1.77	0.1832	Is there heterogeneity among the studies?
6	tau2	0.0236	.	
7	r(i)	54.9 (0.0, 100.0)	.	
8	CVB	0.605 (0.000, 2.663)	.	
9	r(b)	49.3 (0.0, 100.0)	.	

----

example with confidence intervals
example 4
Studies of obesity (BMI ge 30) and lung cancer

Obs	Statistic	Description
1	OR/RR (F)	OR/RR from fixed effects model (95% CI)
2	Coeff. (F)	Coefficient from fixed effects model
3	OR/RR (R)	OR/RR from random effects model (95% CI)
4	Coeff. (R)	Coefficient from random effects model
5	Q	Q-statistic for heterogeneity with 1 degree(s) of freedom
6	tau2	Estimate of between-studies variance (tau squared)
7	r(i)	Between-studies variance as a percent of total variance (r(i))
8	CVB	SQRT(Between-studies variance)/coeff. from random-effects model (CVB)
9	r(b)	New between-studies variance as a percent of total variance (r(b))

4 Frequently Asked Questions

4.1 Q: The macro formats low p-values as '< .0001', but I need to see the real value

A: Use the *OUTDAT* parameter to make an output dataset of the results, then reformat the p-value. I suggest making a permanent dataset, so you can play around. Alternatively, you can set pvalueformat=best12. .

4.2 Q: What should I do about a variable that is represented as a set of indicators?

A: You need to make small datasets for each of the indicators. For example, if you have quintile indicators, you have 4 coefficients for each study. You need to make 4 datasets, each of which has the coefficients for one of the quintile indicators.

5 References

DerSimonian, R. and Laird, N.: Meta-analysis in clinical trials. Controlled Clinical Trials 7:177-188, 1986.

Smith-Warner, S.A., Spiegleman, D., Ritz, J., *et al.*: Methods for pooling results of epidemiologic studies: the pooling project of prospective studies of diet and cancer. Am J Epidemiol 163:1053-1064, 2006.

Takkouche, B., Cadarso-Suarez, C., Spiegelman, D.: Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis. Am J Epidemiol 150(2): 206-215, 1999.

Takkouche, B., Khudyakov, P., Costa-Bouzas, J., Spiegelman, D.: Confidence intervals for heterogeneity measures in meta-analysis. American journal of epidemiology. 2013 Aug 6.

Crippa, A., Khudyakov, P., Wang, M., Orsini, N., Spiegelman, D.: A new measure of betweenstudies heterogeneity in meta-analysis. Statistics in medicine. 2016 Sep 20;35(21):3661-75.

6 Credits

Written by Ellen Hertzmark and Donna Spiegelman for the Channing Laboratory. Questions can be directed to Polyna Khudyakov,
`stpok@channing.harvard.edu`, (617) 432-1822.