

Meta-Analysis with Linear and Nonlinear Multilevel Models

Using PROC MIXED and PROC NL MIXED

Eugene Komaroff, MetaWorks Inc., Medford, MA
Russell D. Wolfinger, SAS Institute Inc., Cary, NC

ABSTRACT

Meta-analysis is both a theory and a toolbox of statistical techniques for combining summary statistics from similar studies. However, with the notable exception of the FDA New Drug Application, typical studies are not similar. Rather than replicate, they try to extend the generalizability of earlier findings. Consequently, between-studies heterogeneity continues to plague the theory and practice of meta-analysis. This paper uses a study level covariate in a multilevel model for meta-analysis to statistically shrink the between-studies heterogeneity and to estimate an overall treatment effect and confidence interval. The model appears in diverse scientific literature under a variety of names, e.g., mixed, hierarchical, empirical Bayes, and variance components. In SAS®, the model is available in PROC MIXED for continuous outcome measures and PROC NL MIXED for binomial outcome measures.

INTRODUCTION

R.A. Fisher was one of the first statisticians to combine experimental outcomes by pooling the p-values from “quite independent tests of significance.” He observed that “the aggregate, based on the product of the individually observed probabilities, gave an impression that the probabilities are on the whole lower than would often have been obtained by chance.” Nevertheless, Fisher was aware that his procedure did not take into account the “detailed composition of the data from which the p-values were derived, which may have been of very different kinds, to obtain a single test of significance” (Fisher, 1995).

Gene Glass, an educational psychologist, coined the term meta-analysis in 1974 referring to the statistical techniques that he used to combine outcomes (effect sizes) from individual studies on the effects of psychotherapy. Today, the term meta-analysis encompasses not only a variety of statistical techniques for pooling p-values and outcomes, but also rigorous standards for systematically reviewing and collecting the data. (Normand, 1999).

META-ANALYSIS IN MEDICAL RESEARCH

DerSimonian and Laird (1986) noted, almost 15 years ago, that meta-analysis is becoming popular in medical research where information on efficacy of a treatment is available from a number of clinical studies with

similar treatment protocols. The popularity of meta-analysis has grown considerably since then (Wang & Bushman, 1999). Undoubtedly, the growth has been fueled by the development, dissemination and accessibility of meta-analysis theory, techniques and software (e.g., Hedges & Olkin, 1985; Raudenbush & Bryk, 1985; DuMouchel, 1989; Laird & Mosteller, 1990, Fleiss, 1993; Cooper & Hedges, 1994; Normand 1995).

Nevertheless despite its popularity, meta-analysis is not free of controversy. A persistent criticism is heterogeneity of study outcomes. Meta-analysts have responded by exploring statistical methods to understand and explain the heterogeneity. Linear and non-linear multilevel models that permit relatively easy inclusion of study level covariates seem promising. Covariates have been used to statistically adjust for non-equivalent or non-randomized groups in observational research (Cook & Campbell, 1979). Similarly, they could be used to adjust for non-equivalent study outcomes in meta-analytic research and to gain substantive information for evaluating treatment efficacy (Thompson & Sharp, 1999).

MULTILEVEL MODELS FOR TREATMENT EFFECTS

Effects can be defined as the quantifiable outcomes from experimental or observational research studies. Here we focus on the odds ratio (OR); however, the methodology we present for pooling odds ratios can be used with other popular effects such as relative risk and difference between proportions.

Fixed and random effects models have been the most popular techniques for pooling effects in meta-analysis. The fixed effects model, usually in conjunction with an underlying assumption of homogeneity, uses the inverse variance of each within study effect as a weighting factor to derive an overall average effect size and confidence interval. The random effects model assumes that the variance of effects among studies can be modeled with a probability distribution and uses an estimated between-studies variance component and the within-study variances as the weighting factor.

The multilevel model for meta-analysis can be viewed as an extension of the two “traditional” models. It permits the inclusion of study level covariates to explain or account for as much of the between-studies heterogeneity as possible. Any remaining heterogeneity is then modeled as a between-studies variance component.

Until the recent appearance of PROC NL MIXED, the multilevel random effects model has been difficult to execute with binomial outcomes. It required perhaps untenable corrections for overdispersion or the use of customized SAS macros with an implied or explicit caveat emptor (e.g., Berkey, Hoaglin, Mosteller & Colditz, 1995). PROC MIXED was also an option, but there was always the perturbing possibility that fundamental assumptions were violated or the reality of inadmissible predicted values when using a procedure designed for continuous scales on binomial data.

The purpose of this paper is to compare several methods for fitting multilevel random effects models in meta-analysis of binomial outcomes with SAS® procedures.

METHODS

Our example data are from Tables 4.4 and 4.6 in Wang and Bushman (1999). They consist of cell counts for 2 X 2 contingency tables from 17 studies that examined the effectiveness of a nicotine patch treatment on smoking cessation. The covariate is the proportion of females in each study, which we model in reference to the control group.

We begin by considering models for OR without a study level covariate. The traditional random-effects approach for this case is due to DerSimonian and Laird (1986). They proposed a non-iterative method of moments estimate for the between-studies variance component, which is then added to the weighting factor from the fixed effects model. The additional component produces the random effects model.

DerSimonian and Laird (D&L) showed that their non-iterative variance component estimate compared well to maximum likelihood (ML) and restricted maximum likelihood (REML) estimates, both of which require iterative algorithms. Similarly, we compare the D&L estimate to those produced by PROC MIXED and PROC NL MIXED.

The PROC MIXED analysis uses log OR as the dependent variable and is based on the assumption that it is approximately normally distributed. The PROC NL MIXED logit-normal analysis considers the binomial observations directly and fits a logit model with normally distributed random effects.

Next, we augment the previous two models by including the reported proportion of females enrolled in each study, centered at its mean, as a covariate. The hope here is for a better explanatory model with a smaller between-studies variance component.

Finally, we fit the more general bivariate multilevel

random effects model of van Houwelingen, Zwinderman, and Stijnen (1993), first without and then with the covariate. This model has different but correlated random effects for the placebo and treatment group and can be fit using PROC NL MIXED.

The data and SAS code are provided in the Appendix.

RESULTS

Parameter estimates and approximate 95% confidence limits from the various multilevel models are listed in the table below.

Source	Parameter	Estimate	Lower	Upper
<i>Random Effects Models No Study Level Covariate</i>				
D & L	OR	2.881	2.357	3.521
PROC MIXED	OR	2.886	2.318	3.589
PROC NL MIXED	OR	2.707	2.264	3.150
D & L	Variance Component	0.0564	N/A	N/A
PROC MIXED	Variance Component	0.0592	0.0180	1.1115
PROC NL MIXED	Variance Component	0.3044	0.0613	0.5475
<i>Random Effects Models With Study Level Covariate</i>				
PROC MIXED	OR	2.869	2.300	3.578
	OR Prop Females	0.323	0.029	3.611
	Variance Component	0.0594	0.0178	1.2285
PROC NL MIXED	OR	2.687	2.245	3.129
	OR Prop Females	0.442	-0.381	1.265
	Variance Component	0.3227	0.0618	0.5837
<i>Bivariate Random Effects Model No Study Level Covariate</i>				
PROC NL MIXED	OR	2.914	2.263	3.565
	Variance Component	0.0561	-0.0536	0.1659
<i>Bivariate Random Effects Model With Study Level Covariate</i>				
PROC NL MIXED	OR	2.871	2.251	3.490
	OR Prop Females	0.297	-0.371	0.966
	Variance Component	0.0404	-0.0537	0.1344

The OR estimates for treatment for all of the models are similar and support the same substantive conclusion about the efficacy of the treatment: Participants in the nicotine patch group were almost 3 times more likely to quit smoking. Although not statistically significant, the covariate parameter estimates are also similar across the multilevel models and indicate the increase in the odds of cessation for a unit increase in the proportion of females.

The between-studies variance component from the PROC MIXED multilevel model was almost identical to the D&L variance component. The PROC NL MIXED logit-normal variance component is different, but because the random effect enters in a different form in this model, the associated variance component is not directly comparable. The covariate produced essentially no change in the PROC MIXED variance component, and a 6% increase in the PROC NL MIXED logit-normal variance component

(0.3044 to 0.3227).

The PROC NLMIXED bivariate model has distinct random effects for the placebo and treatment group, and they are assumed to be normally distributed with mean 0 and an unstructured 2 x 2 covariance matrix. This matrix consists of three unknown parameters: the variance of the placebo group, the variance of the treatment group, and the covariance between the two. Instead of listing all three of these parameters, the table presents the sum of the variances minus twice the covariance. This quantity is comparable to the D&L and PROC MIXED variance components, as it is the variance of the difference of the two random effects in the bivariate model. The parameter estimates were very similar to the other models and the covariate produced a 28% reduction in the variance component (0.0561 to 0.0404).

CONCLUSION

All models that were considered in this paper produced the same substantive conclusion about treatment efficacy. Including covariates in multilevel random effects models, which is technically much easier than adjusting the D&L variance component, appears to be useful technique for minimizing and/or explaining between-studies heterogeneity. The bivariate random effects model, which permits separate random effects for treatment and control groups, seems to be even more tractable for understanding both between-studies heterogeneity and treatment effects.

REFERENCES

Berkey CS, Hoaglin DC, Mosteller F, Colditz GA (1995). A random-effects regression model for meta-analysis. *Statistics in Medicine*, 14, 395-411.

Cooper HM and Hedges LV (1994). *The Handbook of Research Synthesis*. New York:Russell Sage Foundation.

Cook TD and Campbell DT (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin Company.

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

DuMouchel W. (1989). Bayesian meta-analysis in Berry, D.A. (ed.) *Statistical Methods in the Pharmaceutical Sciences*: New York: Marcel Dekker:

Fisher RA. (1995). *Statistical Methods, Experimental Design and Scientific Inference*. New York: Oxford University Press

Fleiss, JL (1993). The statistical basis of meta-analysis. *Statistical Methods in Medical Research*, 2, 121-145.

Hedges LV and Olkin I (1985). *Statistical Methods for Meta-Analysis*. New York: Academic Press.

Laird N and Mosteller F (1990). Some statistical methods for combining experimental results. *International Journal of Technology Assessment in Health Care*, 6, 5-30.

Normand SL (1999). Meta-analysis: Formulating, evaluating, combining, and reporting. *Statistics in Medicine*, 18, 321-359.

Normand SL (1995). Meta-analysis software: A comparative review. *American Statistician*, 49, 298-309.

Raudenbush SW and Bryk AS (1985). Empirical Bayes meta-analysis. *Journal of Educational Statistics*, 10, 75-98.

Thompson SG and Sharp SJ (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine*, 18, 2693-2708.

van Houwelingen HC, Zwinderman AH, and Stijnen Th. (1993). A bivariate approach to meta-analysis. *Statistics in Medicine*, 12, 2273-2284.

Wang MC and Bushman BJ (1999). *Integrating Results Through Meta-Analytic Review Using SAS Software*. Cary, NC: SAS Institute Inc.

CONTACT INFORMATION

Question or comments, please contact:

Eugene Komaroff
MetaWorks Inc.
10 Presidents Landing
Medford, MA 02155
781-395-0700
Fax: 781-395-7336
ekomaroff@metawork.com

APPENDIX

```
/*---data from Wang and Bushman (1999,
Tables 4.4 and 4.6)---*/
data wb_h;
input n11 n12 n21 n22 prop_fem;
or = (n11*n22)/(n21*n12);
lor = log(or);
v_lor = 1/n11 + 1/n12 + 1/n21 + 1/n22;
w_lor = 1/v_lor;
d_pfem = 0.5564118 - prop_fem;
row = _n_;
col = _n_;
value = v_lor;
study = _n_;
datalines;
```

```

36  64  22  77  0.402
22  34  11  45  0.268
29  13  22  21  0.504
19  36  7   45  0.532
20  31  7   45  0.532
45  110 22  142 0.584
48  91  30  107 0.645
21  36  11  44  0.679
56  64  24  96  0.538
121 721 73  771 0.551
19  21  6   34  0.525
70  330 15  185 0.612
46  67  17  90  0.591
43  102 7   137 0.700
61  60  29  95  0.603
37  91  11  118 0.627
23  55  7   73  0.566
run;

/*---random effects with no
covariates---*/
proc mixed data=wb_h cl;
class study;
model lor = / ddfm=sat s cl;
random study / gdata=wb_h;
run;

/*---alternative specification---*/
proc mixed data=wb_h cl;
class study;
model lor = / ddfm=sat s cl;
random study;
weight w_lor;
parms 1 1 / hold=2;
run;

/*---random effects with covariate---*/
proc mixed data=wb_h cl;
class study;
model lor = d_pfem / ddfm=sat s cl;
random study / gdata=wb_h;
run;

/*---construct binomial observations---*/
data wb_v;
set wb_h;
event = n11;
tot = n11+n12;
treat = 1;
a_fem = d_pfem;
b_fem = d_pfem;
output;
event = n21;
tot = n21+n22;
treat = 0;
a_fem = 0;
b_fem = d_pfem;
output;
keep study treat event tot a_fem b_fem;
run;

```

```

/*---logit-normal model with no
covariate---*/
proc nlmixed data=wb_v cov corr;
parms beta0=1 betal=-2 vc=1;
bounds vc >= 0;
eta = beta0 + betal*treat + u;
expeta = exp(eta);
p = expeta/(1+expeta);
model event ~ binomial(tot,p);
random u ~ normal(0,vc) subject=study;
estimate 'OR' exp(betal);
run;

/*---logit-normal model with
covariate---*/
proc nlmixed data=wb_v cov corr;
parms beta0=1 betal=-2 beta2=1 vc=1;
bounds vc >= 0;
eta = beta0 + betal*treat +
beta2*a_fem + u;
expeta = exp(eta);
p = expeta/(1+expeta);
model event ~ binomial(tot,p);
random u ~ normal(0,vc) subject=study;
estimate 'OR' exp(betal);
estimate 'OR Prop Fem' exp(beta2);
run;

/*---bivariate model with no
covariate---*/
proc nlmixed data = wb_v;
parms beta0=-1 betal=-2 var0=1 cov01=1
var1=2;
bounds var0 >= 0, var1 >= 0;
if (treat=0) then eta = beta0 + u0;
else eta = betal + u1;
p = exp(eta)/(1+exp(eta));
model event ~ binomial(tot,p);
random u0 u1 ~ normal([0,0],
[var0,cov01,var1]) subject=study;
estimate 'OR' exp(betal-beta0);
estimate 'Var Comp' var0 + var1 - 2*cov01;
run;

/*---bivariate model with
covariate---*/
proc nlmixed data = wb_v;
parms beta0=-1 betal=-2 beta20=1 beta21=1
var0=1 cov01=1 var1=2;
bounds var0 >= 0, var1 >= 0;
if (treat=0) then eta = beta0 +
beta20*b_fem + u0;
else eta = betal + beta21*b_fem + u1;
p = exp(eta)/(1+exp(eta));
model event ~ binomial(tot,p);
random u0 u1 ~ normal([0,0],
[var0,cov01,var1]) subject=study;
estimate 'OR' exp(betal-beta0);
estimate 'OR prop fem' exp(beta21-beta20);
estimate 'Var Comp' var0 + var1 - 2*cov01;
run;

```