Arm-based network meta-analysis

Hans-Peter Piepho

Biostatistics Unit
University of Hohenheim
Stuttgart, Germany

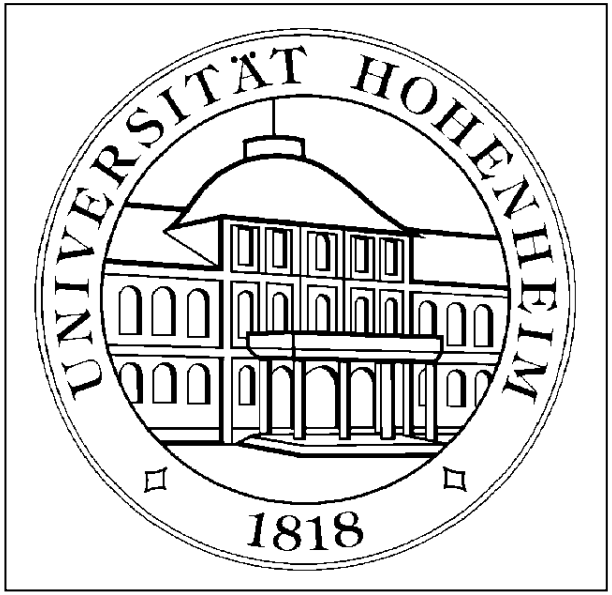


Table of contents

1. Introduction

Network meta-analysis

- More than two treatments tested in combined trials (studies)
- Need to combine direct and indirect evidence on treatment comparisons

- Direct comparison: Trials *A* vs *B*
- Indirect comparison: Trials *A* vs *C* and *B* vs *C*

- Other names:

  Mixed-treatment comparisons (MTC)

  Mixed-treatment meta-analysis (MTM)

1. Introduction

Table 1. Smoking Cessation Rates (r_{ik}/n_{ik}) (Hasselblad 1998)

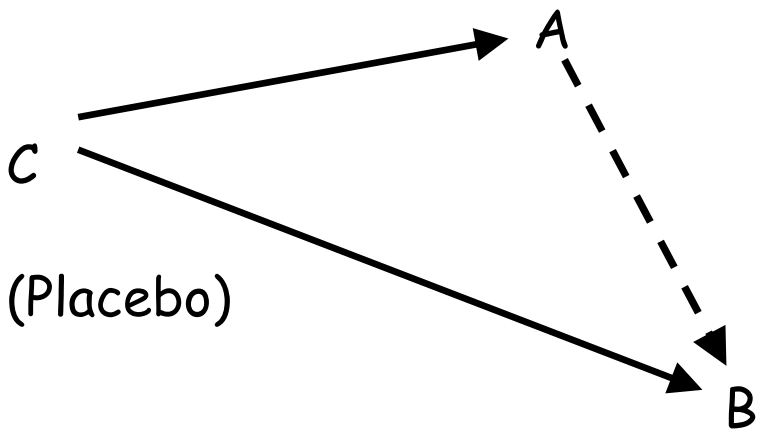
Baseline	Study number	No contact (A)	Self-help (B)	Individual counseling (C)	Group counseling (D)
$G_{(A)}$	1	9/140		23/140	10/138
$G_{(B)}$	2		11/78	12/85	29/170
$G_{(A)}$	3	79/702	77/694		
	4	18/671	21/535		
	5	8/116	19/146		
	6	75/731		363/714	
	7	2/106		9/205	
	8	58/549		237/1,561	
	9	0/33		9/48	
	10	3/100		31/98	
	11	1/31		26/95	
	12	6/39		17/77	
	13	95/1,107		134/1,031	
	14	15/187		35/504	
	15	78/584		73/675	
	16	69/1,177		54/888	
	17	64/642		107/761	
	18	5/62		8/90	
	19	20/234		34/237	
	20	0/20			9/20
$G_{(B)}$	21		20/49	16/43	
	22		7/66		32/127
$G_{(C)}$	23			12/76	20/74
	24			9/55	3/26

Direct comparison (A vs B)

Indirect comparison (via C)

Example 1:
Lu and Ades (2006)

1. Introduction



? Indirect comparison

Comparison	Mean difference (contrast)
A vs C	-0.34
B vs C	-0.19

$$MD_{AB} = MD_{AC} - MD_{BC} = -0.34 + 0.19 = -0.15$$

1. Introduction

Combining direct and indirect evidence

- Inverse variance method
- Each estimate of mean difference (MD) is 'weighted' by the inverse of its variance
- This leads to a 'mixed' result:

$$\text{'mixed } MD\text{'} = \frac{\frac{1}{\text{var}_{direct}} MD_{direct} + \frac{1}{\text{var}_{indirect}} MD_{indirect}}{\frac{1}{\text{var}_{direct}} + \frac{1}{\text{var}_{indirect}}}$$

(Georgia Salanti, Workshop Zurich 2011)

1. Introduction

Parallels with multi-environment trials (MET) in crop science

- Incomplete genotype × environment trials

  (treatments = genotypes, environments = trials, studies)

- Interested in genotype means across environments

- Heterogeneity between environments ⇒ genotype-environment interaction

- Modelling variance-covariance structure for heterogeneity
  ⇒ variance-covariance structures for genotype-environment interaction
  ⇒ variances and covariances not constant between genotypes
  ⇒ stability analysis, analysis of phenotypic stability

- Also similar to incomplete block designs

1. Introduction

Disclaimer

I am staying entirely in a frequentist framework!

2. Modelling individual patient data

Two modelling approaches

(1) Contrast-based models

- relative treatment effects compared to baseline (log relative risk, log odds ratio, mean difference)
- models for contrasts with baseline

(2) Arm-based models

- absolute treatment effects (log risk, log odds, treatment means)
- two-way ANOVA models for factors trial and treatment

2. Modelling individual patient data

2.1 Contrast-based approach

Linear predictors for two treatments *A* and *B*

A = baseline treatment

B = new medication

A: $\eta = \mu$

B: $\eta = \mu + d_{AB}$

μ = baseline effect for the trial

d_{AB} = effect of treatment *B* compared to baseline *A*

2. Modelling individual patient data

Linear predictors for three treatments A, B and C

(1) When A is baseline (A vs B and A vs C trials)

A: $\eta = \mu$

B: $\eta = \mu + d_{AB}$

C: $\eta = \mu + d_{AC}$

(2) When B is baseline (B vs C trials)

B: $\eta = \mu$

C: $\eta = \mu + d_{BC}$

2. Modelling individual patient data

Basic parameters and functional parameters

Basic parameters: d_{AB} , d_{AC}

Functional parameters: $d_{BC} = d_{AC} - d_{AB}$

(2) When B is baseline (B vs C trials)

B: $\eta = \mu$

C: $\eta = \mu + d_{AC} - d_{AB}$

2. Modelling individual patient data

The linear predictor for the k-th treatment in the i-th trial is given by

$$\eta_{ik} = \mu_i + U_{ik}\delta_{ib(i)k}$$

where

μ_i = baseline parameter in the i-th trial

= expected value of the baseline treatment $b(i)$ in the i-th trial

$\delta_{ib(i)k}$ = random effect of treatment k versus baseline $b(i)$ in the i-th trial

$$U_{ik} = \begin{cases} 1, & k \neq b(i) \\ 0, & k = b(i) \end{cases}$$

(Lu & Ades, 2006)

2. Modelling individual patient data

Random effects for baseline contrasts:

$$E\left(\delta_{ib(i)k}\right) = d_{b(i)k}$$

$d_{b(i)k}$ = treatment effects to be estimated across trials

Fixed effects-part of the model:

$$E\left(\eta_{ik}\right) = \mu_i + U_{ik} d_{b(i)k} \; .$$

2. Modelling individual patient data

Heterogeneity between trials

$\Rightarrow$ Variance-covariance structure for $\delta_{ib(i)k}$ in i-th trial, e.g.

$$\mathrm{var}\left\{\delta_{ib(i)k}\right\} = \left(I_{n(i)-1} + J_{n(i)-1}\right)\tau^2 / 2$$

where

I_n = n-dimensional identity matrix

J_n = $n \times n$ matrix of ones

τ^2 = a variance component for between-trial heterogeneity

$n(i)$ = number of treatments in the i-th trial

(Higgins & Whitehead, 1996; Lu & Ades, 2004)

2. Modelling individual patient data

Conditionally on the linear predictor, the observation y_{ijk} on the j-th individual in the i-th trial for the k-th treatment has expected value

$$E\left(y_{ijk} \mid \eta_{ib(i)k}\right) = g^{-1}\left(\eta_{ib(i)k}\right)$$

where $g(.)$ is a suitable link function

$\Rightarrow$ Generalized linear mixed model (GLMM)

$\Rightarrow$ use (e.g.) adaptive Gaussian quadrature (Pinheiro & Bates, 1995)

2. Modelling individual patient data

2.2 Arm-based approach

An alternative linear predictor

$$\eta_{ik} = \beta_i + \alpha_k + u_{ik}$$

where

β_i = fixed main effect of the i-th trial,

α_k = main effect of the k-th treatment, and

u_{ik} = random effect associated with η_{ik}

$$E(\eta_{ik}) = \beta_i + \alpha_k$$

2. Modelling individual patient data

Variance-covariance structure for heterogeneity

Let u_i = vector of random effects u_{ik} for the i-th trial

Then

$E(u_i) = 0$ and

$\text{var}(u_i) = \Sigma_i$

2. Modelling individual patient data

2.3 Relation between contrast-based and arm-based model

$$\eta_{ik} = \beta_i + \alpha_k + u_{ik}$$

$$\eta_{ik} = \underbrace{\beta_i + \alpha_{b(i)} + u_{ib(i)}}_{\mu_i} + \underbrace{\alpha_k - \alpha_{b(i)} + u_{ik} - u_{ib(i)}}_{\delta_{ib(i)k}} = \mu_i + U_{ik}\delta_{ib(i)k}$$

$$\mu_i = \beta_i + \alpha_{b(i)} + u_{ib(i)} \qquad \delta_{ib(i)k} = \alpha_k - \alpha_{b(i)} + \tilde{u}_{ik}$$

where

$\tilde{u}_{ik} = u_{ik} - u_{ib(i)}$ and $E\left(\delta_{ib(i)k}\right) = d_{b(i)k} = \alpha_k - \alpha_{b(i)}$

$b(i)$ = baseline treatment in *i*-th trial

2. Modelling individual patient data

Re-parameterized model has random effects:

$u_{ib(i)}$ and $\widetilde{u}_{ik} = u_{ik} - u_{ib(i)}$ $\left[k \neq b(i)\right]$

Transition from arm-based model to contrast-based model:

Conditioning on $u_{ib(i)}$!!

$\Rightarrow$ baseline treatment has no variance in i-th trial

2. Modelling individual patient data

Let

- u_i = vector of random effects u_{ik} for the i-th trial
- $\widetilde{u}_i$ = vector of random effects $\widetilde{u}_{ik}$ for the i-th trial
- $\mathrm{var}(u_i) = \Sigma_i$ and (without loss of generality) $b(i) = 1$

Then

$$\mathrm{var}(\widetilde{u}_i) = \widetilde{\Sigma}_i = D_i \Sigma_i D_i^T$$

where $D_i = \left(-1_{n(i)-1} \quad I_{n(i)-1}\right)$ is the matrix generating all contrasts relative to the baseline treatment in the i-th trial

2. Modelling individual patient data

Examples for variance-covariance structure of $\widetilde{u}_i$

Constant variance model:

$$\Sigma_i = I_{n(i)}\sigma_u^2 \Rightarrow \widetilde{\Sigma}_i = \left(I_{n(i)-1} + J_{n(i)-1}\right)\sigma_u^2$$

Diagonal model:

$$\Sigma_i = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, ..., \sigma_n^2\right) \Rightarrow \widetilde{\Sigma}_i = \mathrm{diag}\left(\sigma_2^2, \sigma_3^2, ..., \sigma_n^2\right) + J_{n-1}\sigma_1^2$$

Factor-analytic model (one factor):

$\Sigma_i = \lambda\lambda^T$, where $\lambda^T = \left(\lambda_1, \lambda_2, ...\right) \Rightarrow \widetilde{\Sigma}_i = \widetilde{\lambda}\widetilde{\lambda}^T$ with $\widetilde{\lambda}^T = \left(\lambda_2 - \lambda_1, \lambda_3 - \lambda_1, ...\right)$

Unstructured model:

Maximum $n_i\left(n_i - 1\right)/2$ free parameters for $\widetilde{\Sigma}_i$

2. Modelling individual patient data

Implement conditional model for $\widetilde{\Sigma}_i$ via unconditional model for Σ_i

$$\widetilde{u}_{ik} = u_{ik} - u_{ib(i)} \Rightarrow \widetilde{u}_{ik} = \sum_{k=1}^{n} x_{ik} u_{ik}$$

Example 1: Smoking cessation data

		Dummy variables			
Baseline treatment	Treatment	x_{i1}	x_{i2}	x_{i3}	x_{i4}
A	A	0	0	0	0
	B	–1	1	0	0
	C	–1	0	1	0
	D	–1	0	0	1

2. Modelling individual patient data

Baseline treatment	Treatment	Dummy variables x_{i1}	x_{i2}	x_{i3}	x_{i4}
B	*A*	1	–1	0	0
	B	0	0	0	0
	C	0	–1	1	0
	D	0	–1	0	1
C	*A*	1	0	–1	0
	B	0	1	–1	0
	C	0	0	0	0
	D	0	0	–1	1

2. Modelling individual patient data

2.4 Equivalence of conditional and unconditional model

Conditional model:

$\mathrm{var}(\eta_i \mid u_{i1}) = 0 \oplus \widetilde{\Sigma}_i$, where $\eta_i^T = (\eta_{i1}, \eta_{i2}, \ldots)$ and $b(i) = 1$

Unconditional model:

$\mathrm{var}(\eta_i) = \Sigma_i$

Both models are equivalent in the sense that for any contrast $c^T \eta_i$

$$\mathrm{var}(c^T \eta_i \mid u_{i1}) = c^T (0 \oplus \widetilde{\Sigma}_i) c = c^T \Sigma_i c = \mathrm{var}(c^T \eta_i)$$

2. Modelling individual patient data

Equivalence (continued)

$$\mathrm{var}\left(c^T \eta_i \mid u_{i1}\right) = c^T \left(0 \oplus \widetilde{\Sigma}_i\right) c = c^T \Sigma_i c = \mathrm{var}\left(c^T \eta_i\right)$$

To see this, let $c^T = \left(c_1, c_2^T\right)$, where c_1 is the first element of c and c_2 is the remainder. Then $c^T \left(0 \oplus \widetilde{\Sigma}_i\right) c = c_2^T \widetilde{\Sigma}_i c_2 = c_2^T D_i \Sigma_i D_i^T c_2 = \left(c_1, c_2^T\right) \Sigma_i \left(c_1, c_2^T\right)^T$.

2. Modelling individual patient data

Equivalence (continued)

- Models fully equivalent with identity link and normal distribution
- Models not equivalent with other link functions and distributions

Example 1:

- Smoking cessation data
- Changed baseline treatment in some trials
- Used adaptive Gaussian quadrature (GLIMMIX procedure of SAS)
- $\Sigma_i = I_{n(i)}\sigma_u^2 \Rightarrow \widetilde{\Sigma}_i = \left(I_{n(i)-1} + J_{n(i)-1}\right)\sigma_u^2$

2. Modelling individual patient data

Table 1: Smoking cessation data (Example 1)

	Estimate	Standard error
Baseline contrasts using original baseline treatments (*A*)		
d_{AB}	0.4192	0.2959
d_{AC}	0.7407	0.1738
d_{AD}	0.9484	0.3292
Baseline contrasts taking *B* as baseline treatment in trials 3-5		
d_{AB}	0.4415	0.2982
d_{AC}	0.7449	0.1751
d_{AD}	0.9580	0.3315

2. Modelling individual patient data

Table 1: Smoking cessation data (Example 1 continued)

	Estimate	Standard error
Baseline contrasts (2) taking *C* as baseline treatment in trials 6-15		
d_{AB}	0.4407	0.3154
d_{AC}	0.7773	0.1868
d_{AD}	0.9821	0.3493
Two-way model estimates		
$\alpha_B - \alpha_A$	0.3865	0.2387
$\alpha_C - \alpha_A$	0.7166	0.1374
$\alpha_D - \alpha_A$	0.9199	0.2720

2. Modelling individual patient data

Table 2: Smoking cessation data (Example 1 continued); constant variance model for u_{ij}

	Estimate	Standard error
Adjusted means $		
$\alpha_A + \overline{\beta}_\bullet$	-2.4235 a	0.1107
$\alpha_B + \overline{\beta}_\bullet$	-2.0366 ab	0.2106
$\alpha_C + \overline{\beta}_\bullet$	-1.7068 b	0.0971
$\alpha_D + \overline{\beta}_\bullet$	-1.5047 b	0.2273

$ Adjusted means (computed on the logit scale) followed by a common letter are not significantly different at $\alpha = 5\%$ according to a Wald-test.

2. Modelling individual patient data

Table 3: Analysis of smoking cessation data based on two-way model.

Parameter	Estimate	Standard error	AIC
Constant variance:			
σ_u^2	0.09068	0.02810	391.20
Diagonal (treatment-specific variance):			
$\sigma_{u(1)}^2$	0.5599	0.2626	365.91
$\sigma_{u(2)}^2$	0	-	
$\sigma_{u(3)}^2$	0	-	
$\sigma_{u(4)}^2$	0.1292	0.2411	

2. Modelling individual patient data

Table 4: Analysis of smoking cessation data based on two-way model.

Parameter	Estimate	Standard error	AIC
Constant variance:			
σ_u^2	0.09068	0.02810	391.20
Factor-analytic:			
λ_1	0.4969	0.1736	364.02
λ_2	0	-	
λ_3	-0.2423	0.1157	
λ_4	0.05856	0.1985	

2. Modelling individual patient data

Fitting the two-way ANOVA model with SAS

```
proc glimmix data=a maxopt=100
             method=quad(qpoints=6);
class trial trt;
model m/n = trial trt / ddfm=none solution chisq;
random trt*trial;
lsmeans trt / pdiff lines;
run;
```

2. Modelling individual patient data

Fitting the FA model with SAS

```
proc glimmix data=a maxopt=100
             method=quad(qpoints=6);
class trial trt;
model m/n = trial trt
                      / ddfm=none solution chisq;
random trt / sub=trial type=fa1(1);
lsmeans trt / pdiff lines;
run;
```

2. Modelling individual patient data

2.5 Trial effects fixed or random?

Trial effects fixed

- Inference based on within-trial information
- Inference protected by randomization
- Obeys principle of concurrent control (Stephen Senn, 2010)
- Can only assess *relative* treatment effects

Trial effects random

- Recovery of inter-trial information
- Need to assume that trials in NMA are random sample
- Can also assess *absolute* treatment effects

2. Modelling individual patient data

Recent discussion on arm-based (AB) versus contrast-based (CB) models

- The discussion focusses much on estimation of relative treatment effects (CB) versus absolute treatment effects (AB)
- I think this becomes a non-issue when a trial main effect is included in the AB model
- The main issue is whether or not to recover the inter-trial information, i.e. whether the trial main effect is taken as fixed or random

Dias S, Ades AE 2016 Absolute or relative effects? Arm-based synthesis of trial data (Commentary). Research Synthesis Methods 7, 23-28.

Hong, H., Chu, H., Zhang, J., Carlin, B.P. 2016 Rejoinder to the discussion of "a Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons," by S. Dias and A.E. Ades. Research Synthesis Methods 7, 29-33.

3. Treatment summaries and contrasts thereof

Table 5: Summary measures analysis (empirical logits) for smoking cessation data (REML) (Example 1). We assumed $\Sigma_i = I_{n(i)}\sigma_u^2$ for heterogeneity under the two-way model. This is equivalent to fitting $\widetilde{\Sigma}_i = \left(I_{n(i)-1} + J_{n(i)-1}\right)\sigma_u^2$ for the baseline-contrast model.

Estimable function		Estimate	Standard error
Baseline contrasts §	Two-way model §		
d_{AB}	$\alpha_B - \alpha_A$	0.3978	0.3305
d_{AC}	$\alpha_C - \alpha_A$	0.7013	0.1972
d_{AD}	$\alpha_D - \alpha_A$	0.8642	0.3749

§ Results are identical for both analyses $\Rightarrow$ Piepho et al. (2012)

3. Treatment summaries and contrasts thereof

Table 5 (continued)

Contrast		Estimate	Standard error
Baseline contrasts	Two-way model		
		Adjusted means $	
-	$\alpha_A + \overline{\beta}_\bullet$	-2.3792 a	0.1553
-	$\alpha_B + \overline{\beta}_\bullet$	-1.9815 ab	0.2886
-	$\alpha_C + \overline{\beta}_\bullet$	-1.6779 b	0.1352
-	$\alpha_D + \overline{\beta}_\bullet$	-1.5150 b	0.3100

$ Adjusted means followed by a common letter are not significantly different at $\alpha = 5\%$ according to a t-test using the Kenward-Roger (1997) method for approximating the denominator degrees of freedom and variance adjustments

3. Treatment summaries and contrasts thereof

Take home message up to here

Compared:

- Contrast-based model (conditional) $\eta_{ik} = \mu_i + U_{ik}\delta_{ib(i)k}$
- Arm-based model (unconditional) $\eta_{ik} = \beta_i + \alpha_k + u_{ik}$

Full equivalence:

- Summary data
- Individual patient data with identity link and normal errors

Very similar results:

- All other cases
- But: Contrast-based model is not invariant to choice of baseline!

4. Testing inconsistency

Example

- Trial network with three treatments (*A*, *B*, *C*)
- Three types of trial: *A* vs *B*, *A* vs *C* and *B* vs *C*
- Consider evidence on *B* vs *C*
- Need to combine direct and indirect evidence on treatment comparisons

  Direct comparison: Trials *B* vs *C*

  Indirect comparison: Trials *A* vs *B* and *A* vs *C*

- Inconsistency (incoherence):

  $\Rightarrow$ direct and indirect comparisons for *B* vs *C* do not agree

4. Testing inconsistency

Extending the notion of inconsistency

- Comparison of direct and indirect evidence on a contrast
- Presence of a new treatment in a trial may well modify the direct difference between *A* and *B* (Lu et al., 2011)

  $\Rightarrow$ need to also compare direct comparisons from different types of trial

Idea

$\Rightarrow$ Test interaction in trial type $\times$ treatment classification

4. Testing inconsistency

Model to test for inconsistency

$$\eta_{ijk} = \delta_j + \beta_{ij} + \alpha_k + (\alpha\delta)_{jk} + u_{ijk}$$

δ_j = fixed main effect for the j-th trial type (design)

$(\alpha\delta)_{jk}$ = fixed effect for the interaction jk-th trial type $\times$ treatment

- Heterogeneity u_{ijk} can be separated from inconsistency $(\alpha\delta)_{jk}$ provided there are several trials per trial type (design)
- Heterogeneity is a property of variation among trials within the same trial type (design), while inconsistency affects variation between trial types

(Piepho, Madden and Williams, 2012, *Biometrics*)

5. Estimation of variance components

5.1 The bias problem illustrated

$$\eta_{ij} = \beta_i + \alpha_j + u_{ij}$$

Assume here: balanced data, normality

Use:

MS_{ST} = ANOVA mean square for the trial-by-treatment interaction

MS_E = error mean square

s = number of trials (studies)

t = number of treatments

n = number of replications

5. Estimation of variance components

ML estimation of $\sigma_u^2 = \mathrm{var}(u_{ij})$ is biased:

$$\hat{\sigma}_u^2 = \frac{(1-s^{-1})(1-t^{-1})MS_{ST} - MS_E}{n}$$

By comparison, the ANOVA estimator is unbiased:

$$\hat{\sigma}_u^2 = \frac{MS_{ST} - MS_E}{n}$$

$\Rightarrow$ identical to REML when $\hat{\sigma}_u^2 \geq 0$

Bias problem with ML is potentially worse with unbalanced data and GLMMs

5. Estimation of variance components

5.2 Alternatives to full ML

$\Rightarrow$ penalized quasi-likelihood (PQL)/pseudo-likelihood (PL)

(Brewlow & Clayton 1993; O'Connell & Wolfinger 1993)

$\Rightarrow$ ML with random trial effects

$\Rightarrow$ REML-like modifications of ML

(Piepho et al. 2018 *Pharmaceutical Statistics*)

5. Estimation of variance components

Pseudo-Likelihood (PL)

- Linearization of GLMM by Taylor series expansion
- Analysis of linearized model by residual ML (REML)

$\Rightarrow$ Expansion around $X\beta$ $\Rightarrow$ RMPL (M = marginal)

$\Rightarrow$ Expansion around $X\beta + Zu$ $\Rightarrow$ RSPL (S = subject-specific)

RMPL and RSPL are implemented in the GLIMMIX procedure of SAS

(O'Connell & Wolfinger 1993)

5. Estimation of variance components

ML with random trial effects β_i

$$\hat{\sigma}_u^2 = \frac{\left(1 - s^{-1}\right)MS_{ST} - MS_E}{n}$$ (balanced data, normal)

Problem with unbalanced data and with GLMMs:

$\Rightarrow$ inter-trial information recovered

Two *ad hoc* remedies to prevent recovery of inter-trial information:

(1) fix $\sigma_\beta^2 = \mathrm{var}(\beta_i)$ at large value

(2) estimate σ_u^2 assuming random trial effects, then take trial effects fixed

5. Estimation of variance components

A useful representation of REML as a modification of ML

$$Y = X\beta + Zu + e$$

where

β and u = fixed and random effects vectors; $\mathrm{var}(u) = G$

X and Z = design matrices, and

e = vector of residual errors; $\mathrm{var}(e) = R$

5. Estimation of variance components

Re-write the model as

$$Y = X\beta + \widetilde{Z}\widetilde{u}$$

where $\widetilde{Z} = \begin{pmatrix} Z & I \end{pmatrix}$ and $\widetilde{u}^T = \begin{pmatrix} u^T & e^T \end{pmatrix}$.

5. Estimation of variance components

Equivalent form:

$$Y = X\beta^* + \left(I - X\left(X^T X\right)^- X^T\right)\tilde{Z}\tilde{u}$$

where

$$\beta^* = \beta + \left(X^T X\right)^- X^T \tilde{Z}\tilde{u}$$

and M^- denotes a g-inverse of M .

5. Estimation of variance components

Now take β^* as fixed:

$$E[Y] = X\beta^*$$

$$V[Y] = K\Sigma K^T$$

where $K = \left(I - X\left(X^T X\right)^{-} X^T\right)$ and $\Sigma = ZGZ^T + R$.

5. Estimation of variance components

Estimation of this model by ML is equivalent to the classic representation of REML as marginal ML for KY, where we maximize the likelihood for $KY \sim N\left(0, K\Sigma K^T\right)$, which is synonymous with the likelihood for $Y \sim N\left(X\left(X^T X\right)^- X^T Y, K\Sigma K^T\right)$, which in turn is synonymous with ML for $Y \sim N\left(X\beta^*, K\Sigma K^T\right)$ profiled over β^* choosing $\hat{\beta}^* = \left(X^T X\right)^- X^T Y$.

The important point to observe about this representation of REML is that by reducing the dimensionality of the random effects space (i.e., substitution of $\tilde{Z}$ by $K\tilde{Z}$), we effectively inflate the variance component estimates.

5. Estimation of variance components

REML-like modification of ML for GLMM

$$E[Y \mid u] = g^{-1}[X\beta + Zu]$$

where

$g(.)$ = link function and

$Y \mid u$ ~ some distribution in the exponential family

$\Rightarrow$ This time, cannot include the residual error e in the random effects

5. Estimation of variance components

But we can take the modification as far as this:

$$E[Y \mid u] = g^{-1}\left[X\beta^{**} + \left(I - X\left(X^T X\right)^{-} X^T\right)Zu\right]$$

where $\beta^{**} = \beta + \left(X^T X\right)^{-} X^T Zu$.

In this modification, we have replaced Z by KZ in a fashion analogous to REML in the normal case, thereby reducing the dimensionality of the random effects

5. Estimation of variance components

Example 2: Diabetes data (Elliot & Meyer, 2007, Lancet)

- Incidence of diabetes with various antihypertensive drugs
- Binomial response (cases/total counts)
- 6 treatments:
  ACE Inhibitor, ARB, CCB, Diuretic, Placebo, Beta-blocker
- 21 studies
- Data very incomplete

5. Estimation of variance components

Table 6: Treatments tested in 21 studies of diabetes dataset.

Trial ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Treatment																					
ACE inhibitor	x	x	x		x		x				x						x			x	
ARB						x								x				x			x
CCB	x	x		x								x	x			x				x	x
Diuretic		x	x					x		x		x			x				x		
Placebo						x	x	x	x		x						x	x	x		
Beta-blocker	x			x	x					x			x	x	x	x				x	

5. Estimation of variance components

5.3 Simulation study

- Fit binomial GLMM with logit link using pseudo-likelihood (PL)
- Simulate from fitted model (10,000 runs)
- Assess:

  $\Rightarrow$ bias and mean squared error (MSE) of parameter estimates

  $\Rightarrow$ coverage probabilities of 95% confidence intervals for contrasts

5. Estimation of variance components

Table 7: Simulation results for variance estimate

Method	Bias [× 10^{-2}]	MSE [× 10^{-4}]
Trial main effects β_i fixed		
ML	-4.582	22.188
ML (KZ for Z)§	-0.1720	6.580
RSPL	0.0124	6.629
RMPL	-0.2635	5.754
Trial main effect β_i random		
ML ($\sigma^2_\beta = 10^4$)	-1.6100	6.853
ML (β_i fixed for estimating α_j	-1.5630	6.843

§ REML-like modification of ML

5. Estimation of variance components

Table 8: Simulation results for contrast $\alpha_1 - \alpha_2$ (A – B)

Method	Bias [× 10^{-2}]	MSE [× 10^{-4}]	Coverage (%)
Trial main effects β_i fixed			
ML	0.06656	287.9	0.7945
ML (KZ for Z)§	-0.00758	286.4	0.9437
RSPL	0.04517	285.0	0.9467
RMPL	0.05967	283.8	0.9436
Trial main effect β_i random			
ML ($\sigma^2_\beta = 10^4$)	0.00449	286.4	0.9102
ML (β_i fixed for estimating α_j	0.00716	286.1	0.9187

§ REML-like modification of ML

6. Summary

Compared:

- Contrast-based model (conditional) $\eta_{ik} = \mu_i + U_{ik}\delta_{ib(i)k}$
- Arm-based model (unconditional) $\eta_{ik} = \beta_i + \alpha_k + u_{ik}$

Full equivalence:

- Summary data
- Individual patient data with identity link and normal errors

Very similar results:

- All other cases
- But: Contrast-based model is not invariant to choice of baseline!

6. Summary

- Arm-based (two-way ANOVA) model invariant to choice of baseline
- Arm-based (two-way) model much easier to fit using standard software
- Easy to fit two-way variance-covariance models for heterogeneity
- Inconsistency = treatment x trial design interaction
- PL/PQL & REML-like modification of ML are preferred methods for variance estimation

Lesson for multi-environment variety trials:

- Consider testing inconsistency in trials networks

References:

Madden, L.V., Piepho, H.P., Paul, P.A. (2016): Models and methods for network meta-analysis. *Phytopathology* **106**, 792-806.

Piepho, H.P. (2014): Network-meta analysis made easy: Detection of inconsistency using factorial analysis-of-variance models. *BMC Medical Research Methodology* **14**, 61.

Piepho, H.P., Madden, L.V., Roger, J., Payne, R., Williams, E.R. (2018): Estimating the variance for heterogeneity in arm-based network meta-analysis. *Pharmaceutical Statistics* (forthcoming).

Piepho, H.P., Madden, L.V., Williams, E.R. (2015): Multiplicative interaction in network meta-analysis. *Statistics in Medicine* **34**, 582-594.

Piepho, H.P., Möhring, J., Schulz-Streeck, T., Ogutu, J.O. (2012): A stage-wise approach for analysis of multi-environment trials. *Biometrical Journal* **54**, 844-860.

Piepho, H.P., Williams, E.R., Madden, L.V. (2012): The use of two-way mixed models in multi-treatment meta-analysis. *Biometrics* **68**, 1269-1277.

Thanks!

4. Testing inconsistency

Example

- Trial network with three treatments (*A*, *B*, *C*)
- Three types of trial: *A* vs *B*, *A* vs *C* and *B* vs *C*
- Consider evidence on *B* vs *C*
- Need to combine direct and indirect evidence on treatment comparisons

  Direct comparison: Trials *B* vs *C*

  Indirect comparison: Trials *A* vs *B* and *A* vs *C*

- Inconsistency (incoherence):

  $\Rightarrow$ direct and indirect comparisons for *B* vs *C* do not agree

4. Testing inconsistency

Reasons for inconsistency

- A new drug may be tested on a population of patients, for which a standard drug did not show a satisfactory effect. The effect relative to a placebo in such a selected population may differ from the effect in a population that is not selected in this way.

- Inconsistency may also occur in open-label or imperfectly blinded trials (Lumley, 2002)

Other term

- Incoherence (Lumley, 2002)

4. Testing inconsistency

Inconsistency relation

- Assume that *B* is baseline treatment in trials *B* vs *C*
- Use functional parameter to model effect of *C* :

$$d_{BC} = d_{AC} - d_{AB}$$

- Modification in case of inconsistency :

$$d_{BC} = d_{AC} - d_{AB} + w_{ABC} \qquad \text{(inconsistency relation)}$$

$\Rightarrow$ use this for treatment C in trials where B is baseline

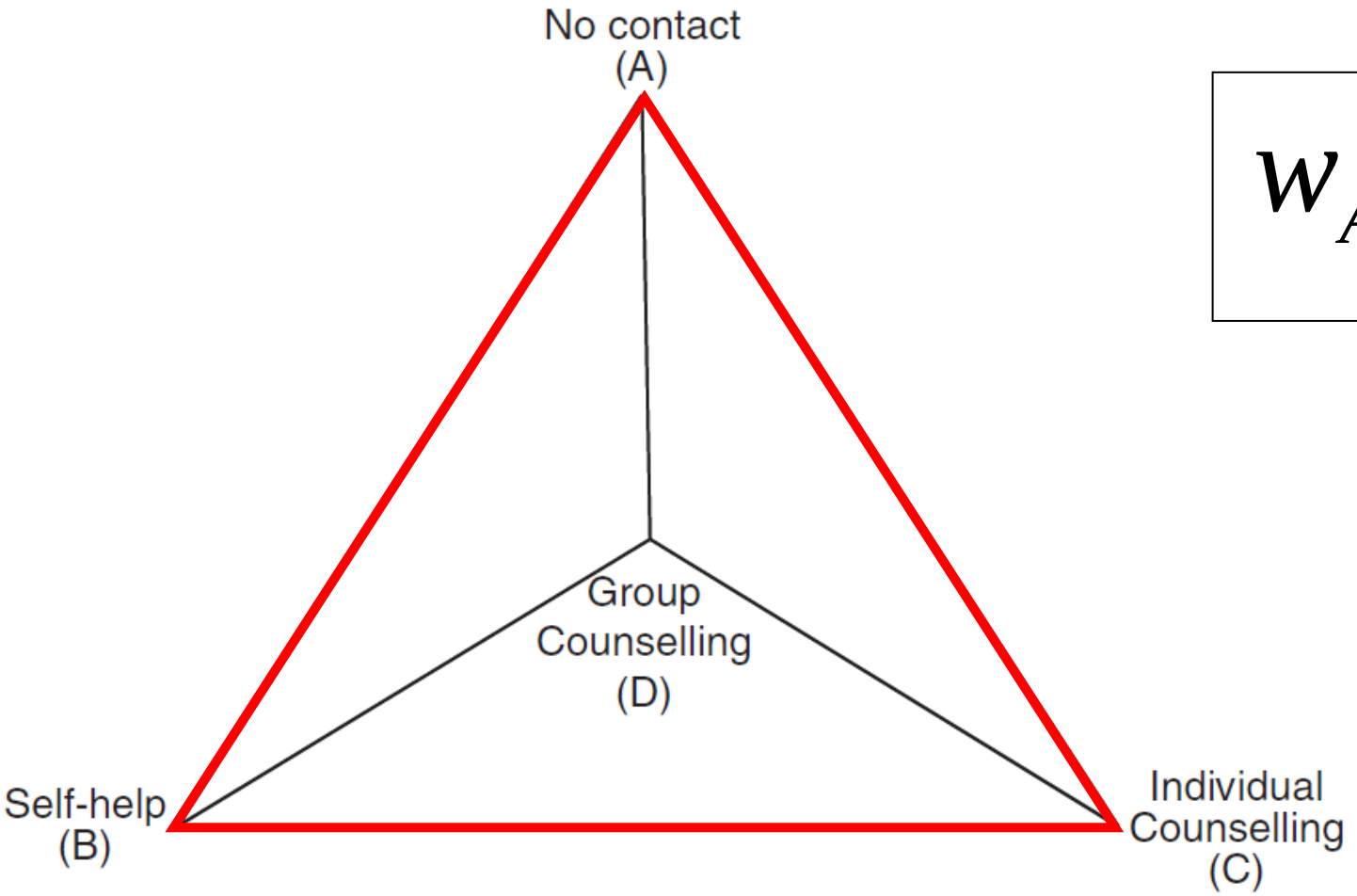
- If w_{ABC} is significant, inconsistency is established

4. Testing inconsistency

Loops

Network forms a closed loop between *A*, *B* and *C* in an undirected graph with vertices corresponding to treatments and edges representing direct comparisons between treatments (Lu and Ades, 2006)

4. Testing inconsistency



$$W_{ABC}$$

(Dias et al., 2010)

Undirected graph: Vertices = treatments

Edges = direct comparisons

4. Testing inconsistency

Using inconsistency factors is not easy!

- Modeling and interpretation of inconsistency become more difficult in the presence of multi-arm trials, and fitting the model may require careful programming
- The types of inconsistency that can be tested using inconsistency factors are not invariant to the choice of basic parameters
- "… we have not managed to find a general formula of a mechanical routine to count [the number of independent consistency relations]" (Lu & Ades, 2006)
- "In practice, an inconsistency model must be programmed very carefully, and the [number of independent inconsistencies] may have to be counted by hand." (Lu & Ades, 2006)

4. Testing inconsistency

Extending the notion of inconsistency

- Comparison of direct and indirect evidence on a contrast
- Presence of a new treatment in a trial may well modify the direct difference between *A* and *B* (Lu et al., 2011)

  $\Rightarrow$ need to also compare direct comparisons from different types of trial

Idea

$\Rightarrow$ Test interaction in trial type $\times$ treatment classification

4. Testing inconsistency

Model to test for inconsistency

$$\eta_{ijk} = \delta_j + \beta_{ij} + \alpha_k + (\alpha\delta)_{jk} + u_{ijk}$$

δ_j = fixed main effect for the j-th trial type

$(\alpha\delta)_{jk}$ = fixed effect for the interaction jk-th trial type $\times$ treatment

- Heterogeneity u_{ijk} can be separated from inconsistency $(\alpha\delta)_{jk}$ provided there are several trials per trial type (design)
- Heterogeneity is a property of variation among trials within the same trial type, while inconsistency affects variation between trial types

(Piepho, Madden and Williams, 2012, *Biometrics*)

4. Testing inconsistency

	Treatment		
Trial type	*A*	*B*	*C*
1	X	X	
2	X		X
3		X	X

Fig. 2: Trial type × treatment classification for network {*A* vs *B*, *A* vs *C*, *B* vs *C*}.

- $n = 3$ treatments
- $m = 3$ trial types
- $c = 6$ cells filled

$\Rightarrow \; c - n - m + 1 = 1$ d.f. for interaction trial type × treatment

4. Testing inconsistency

	Treatment		
Trial type	*A*	*B*	*C*
1	X	X	X
2	X	X	

Fig. 3: Trial type × treatment classification for network {*A* vs *B* vs *C*, *A* v. *B*}.

- $n = 3$ treatments
- $m = 2$ trial types
- $c = 5$ cells filled

$\Rightarrow \; c - n - m + 1 = 1$ d.f. for interaction trial type × treatment

4. Testing inconsistency

	Treatment		
Trial type	*A*	*B*	*C*
1	X	X	
2	X		X
3	X	X	X

Fig. 4: Trial type × treatment classification for network {*A* vs *B*, *A* vs *C*, *A* vs *B* vs *C*}.

- $n = 3$ treatments
- $m = 3$ trial types
- $c = 7$ cells filled

$\Rightarrow$ $c - n - m + 1 = 2$ d.f. for interaction trial type × treatment

4. Testing inconsistency

Example 3:

- Diabetes study of Senn et al. (2013)
- 26 trials
- 15 different designs (one three-arm trial)
- 10 treatments, mostly involving glucose-lowering agent added to baseline sulfonylurea treatment
- Continuous outcome: blood glucose change

4. Testing inconsistency

Factor symbol	Factor description
G	Group of trials, trial type, design
S	Study, trial
T	Treatment

Two-way ANOVA

$S \times T = S + T + S.T$

Model for inconsistency

$(G/S) \times T = G + G.S + T + G.T + G.S.T$

inconsistency heterogeneity

4. Testing inconsistency

Locating inconsistency by detachment of individual designs

(Krahn et al. 2013)

Factor symbol	Factor description
D1	D1 = 1 for design 1, D1 = 0 otherwise
G	Group of trials, trial type, design
S	Study, trial
T	Treatment

$$(D1/G/S) \times T = D1 + D1.G + D1.G.S + T + D1.T + D1.G.T + D1.G.S.T$$

D1.T: detach design 1

D1.G.T: inconsistency

D1.G.S.T: heterogeneity

4. Testing inconsistency

Design	Design no. (*k*)	No. of trials	D.f. for D*k*.T	Effect G.S.T fixed			
				Detachment D*k*.T		Inconsistency D*k*.G.T	
				Wald statistic	*p*-value	Wald statistic	*p*-value
acar:plac	1	1	1	0.09	0.7699	**22.45**	**0.0010**
acar:SUal	2	1	1	0.01	0.9091	**22.52**	**0.0010**
metf:plac	4	3	1	0.46	0.4976	**22.07**	**0.0012**
metf:acar:plac	5	1	2	0.15	0.9297	**22.39**	**0.0004**
metf:SUal	6	1	1	**15.02**	**0.0001**	7.52	0.2758
piog:plac	8	1	1	**5.28**	**0.0215**	**17.25**	**0.0084**
piog:metf	9	1	1	**5.40**	**0.0201**	**17.13**	**0.0088**
piog:rosi	10	1	1	0.05	0.8280	**22.49**	**0.0010**
rosi:plac	11	6	1	**6.24**	**0.0125**	**16.30**	**0.0122**
rosi:metf	12	2	1	0.01	0.9199	**22.52**	**0.0010**
rosi:SUal	13	1	1	**15.76**	**<0.0001**	6.77	0.3424

4. Testing inconsistency

Design	Design no. (*k*)	No. of trials	D.f. for D*k*.T	Effect G.S.T random			
				Detachment D*k*.T		Inconsistency D*k*.G.T	
				Wald statistic	*p*-value	Wald statistic	*p*-value
acar:plac	1	1	1	0.02	0.8889	2.25	0.8782
acar:SUal	2	1	1	0.01	0.9430	2.26	0.8765
metf:plac	4	3	1	0.04	0.8379	2.22	0.8814
metf:acar:plac	5	1	2	0.07	0.9634	2.18	0.8129
metf:SUal	6	1	1	1.63	0.2343	0.92	0.9835
piog:plac	8	1	1	0.43	0.5299	1.96	0.9062
piog:metf	9	1	1	0.43	0.5318	1.94	0.9081
piog:rosi	10	1	1	0.01	0.9065	2.27	0.8751
rosi:plac	11	6	1	0.74	0.4112	1.87	0.9168
rosi:metf	12	2	1	0.01	0.9276	2.25	0.8795
rosi:SUal	13	1	1	1.79	0.2146	0.66	0.9930

4. Testing inconsistency

Case-deletion plots and residual diagnostics

(1) Fit model (G/S) × T and compute G.T means

(2) Fit model G + T to G.T means

⇒ Drop a G.T mean and compute T means based on model G + T

⇒ Compute studentized residuals for G.T means from model G + T

4. Testing inconsistency

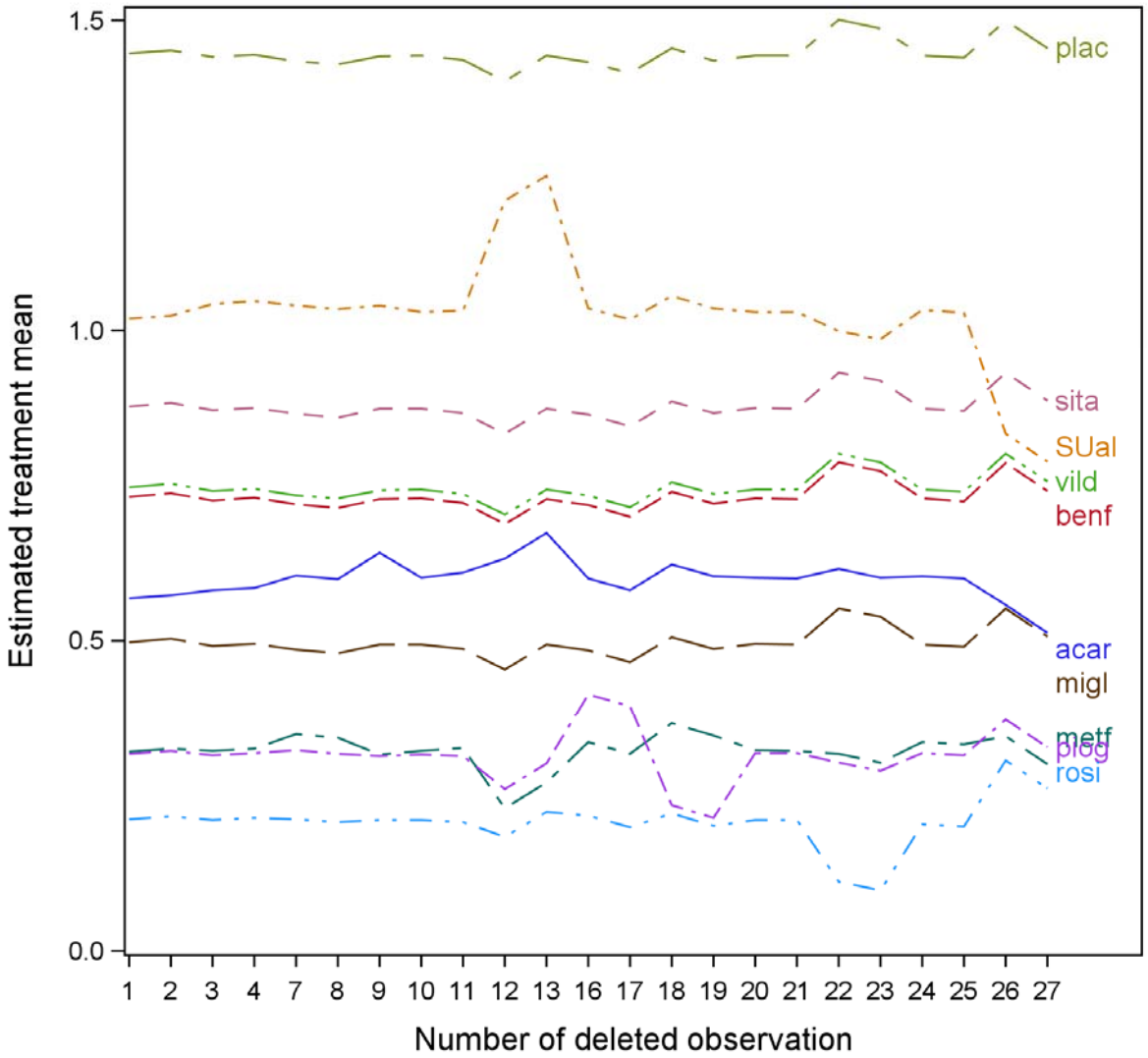


Fig. 4: Case-deletion plot of treatment means. Case-deletion means based on a fit of the model G + T using design × treatment mean estimates obtained from fitting model (2) taking heterogeneity G.S.T as random. To obtain diagnostics for treatment means (factor T), we prevented an intercept from being fitted and imposed a sum-to-zero restriction on the design effects G.

4. Testing inconsistency

Design	Observation	Treatment	G.S.T random	
			PRESS residual	Studentized res.
1	1	Acar	0.0785	0.1453
	2	plac	-0.0785	-0.1453
2	3	acar	0.0619	0.1056
	4	SUal	-0.0619	-0.1056
3	5	benf	.	.
	6	plac	.	.
4	7	metf	-0.0781	-0.2282
	8	plac	0.0781	0.2282
5	9	acar	-0.1507	-0.2601
	10	metf	0.0036	0.0075
	11	plac	0.1193	0.2273
6	**12**	**metf**	**0.6095**	**1.1614**
	13	**SUal**	**-0.6095**	**-1.1614**
7	14	migl	.	.
	15	plac	.	.

4. Testing inconsistency

Design	Observation	Treatment	G.S.T random	
			PRESS residual	Studentized res.
8	16	piog	-0.2802	-0.5585
	17	plac	0.2802	0.5585
9	18	metf	-0.2927	-0.5779
	19	piog	0.2927	0.5779
10	20	piog	-0.0073	-0.0141
	21	rosi	0.0073	0.0141
11	22	plac	-0.2100	-0.6391
	23	rosi	0.2100	0.6391
12	24	metf	-0.0616	-0.1610
	25	rosi	0.0616	0.1610
13	**26**	**rosi**	**-0.6733**	**-1.2693**
	27	**SUal**	**0.6733**	**1.2693**
14	28	plac	.	.
	29	sita	.	.
15	30	plac	.	.
	31	vild	.	.

4. Testing inconsistency