Introduction to Mixed Treatment Comparisons.

A E Ades, Nicky Welton, Guobing Lu.

What are Mixed Treatment Comparisons ?

Mixed Treatment Comparison (MTC) evidence structures are a generalisation of Meta-Analysis [1] evidence structures. Instead of simply analysing a set of Randomised Clinical Trials comparing treatment A vs treatment B, we might have A vs B trials, A vs C trials, A vs D, B vs C, etc.

Indirect comparisons are a special case: here a relative treatment effect - a Log Odds Ratio, risk difference, etc - which is not itself supported by "direct" comparisons of treatments B and C, is estimated "indirectly" from data on A vs B and A vc C comparisons. This is a special case in the sense that there no "mixture" of indirect and indirect evidence

These evidence structures have raised a series of questions:

1. Are indirect estimates biased ?
2. Should direct and indirect evidence be combined ?
3. What statistical methods should be used ?
4. How do we determine whether direct and indirect evidence is consistent ?
5. What is "inconsistency" ?

The purpose of this note is not to answer these questions, but to set out our models for MTC analysis. This is particularly relevant to those wishing to use the WinBUGS software available on this website.

Some useful literature

1. Empirical papers using *indirect comparisons* to make inferences about relative treatment effects. [2 3].

2. Empirical papers comparing direct and indirect estimates. [4 5]

3. Methodological papers in the medical statistics and medical decision making literature on how a single coherent and internally consistent set of estimates can be estimated from MTC evidence structures. [6 7 8 9 10 11 12 13 14]

4. Practical uses of some of these approaches in the medical literature [15], and in decision making applications.[16 17]

What kinds of structures can be analysed by the MTC methods discussed here?

1 These methods can be applied only to *connected* networks of RCTs. [11 14] For example in a dataset consisting of AB, AC, BC, AD, EF, EG, FG pair-wise comparisons, the A,B,C,D group of treatments is not connected with the E,F,G group.

2. *We very strongly recommend that single-arm studies, whether from RCTs or observational studies, are excluded.*

A worked example : Smoking Cessation

This data structure, [8] consists of 24 studies comparing 4 treatments for smoking cessation: A= no intervention, B= self-help, C = individual counselling; D= group counselling. There is evidence in all 6 possible pair-wise comparisons: AB, AC, AD, BC, BD, CD. Two of the trials are 3-arm trials, so there are 50 arms in total. Each cell in the table contains the numerator (successful smoking cessation) and the patient totals in each arm

Comparison	Study number	No contact A	Self-help B	Individual Counselling C	Group counselling D
AB (3)	1	79/702	77/694		
	2	18/671	21/535		
	3	8/116	19/146		
AC (15)	4	75/731		363/714	
	5	2/106		9/205	
	6	58/549		237/1561	
	7	0/33		9/48	
	8	3/100		31/98	
	9	1/31		26/95	
	10	6/39		17/77	
	11	64/642		107/761	
	12	5/62		8/90	
	13	20/234		34/237	
	14	95/1107		134/1031	
	15	15/187		35/504	
	16	78/584		73/675	
	17	69/1177		54/888	
ACD (1)	18	9/140		23/140	10/138
AD (1)	19	0/20			9/20
BC (1)	20		20/49	16/43	
BCD (1)	21		11/78	12/85	29/170
BD (1)	22		7/66		32/127
CD (2)	23			12/76	20/74
	24			9/55	3/26

Our modeling approach is based on the statistical model of Higgins and Whitehead [7]

[10], although the coding and notation is extended to apply to any MTC structure, including those where there is no single treatment to which all other treatments have been compared. Readers are referred the original papers and to [14] for theory.

To analyse data structures of this sort, our underlying model is:

$$\text{logit}(p_{jk}) = \begin{cases} \mu_{jb} & k = b; \quad b = A,B,C \\ \mu_{jb} + \delta_{jkb} & k > b; \quad b = A,B,C \end{cases}$$

where p_{jk} is the probability of smoking cessation in trial j under treatment k, μ_{jb} is the log odds of smoking cessation on baseline treatment b in trial j, and δ_{jkb} is the trial-specific Log Odds Ratio of treatment k relative to treatment b. ("$k > b$" indicates that k is *after* b in the alphabet). We assume a binomial likelihood:

$$r_{jk} \sim \text{Bin}(p_{jk}, n_{jk})$$

Study effects are treated as unrelated nuisance parameters with priors: $\mu_{jb} \sim N(0,10000)$. We take treatment A as baseline, and the treatment effects of B, C, and D relative to treatment A as our *basic* [18] parameters with vague priors:

$$d_{AB}, d_{AC}, d_{AD} \sim N(0,10000)$$

The remaining contrasts (*functional* parameters) can be expressed in terms of these basic parameters

$$d_{BC} = d_{AC} - d_{AB}; \quad d_{BD} = d_{AD} - d_{AB}; \quad d_{CD} = d_{AD} - d_{AC};$$

The trial-specific LORs are now drawn from one of the six Random Effects distributions :

$$\delta_{jXY} \sim N(d_{XY}, \sigma_{XY}^2)$$

If $\sigma^2 = 0$ we obtain a fixed effects model. For Random Effects, we make the assumption of *homogeneous variance*: $\sigma_{XY}^2 = \sigma^2$. A vague prior is provided for the common variance term, for example $\sigma \sim \text{Uniform}(0,2)$. *This prior should not be used unthinkingly when evidence is sparse* [19 20], and it may be worth considering informative priors based on literature - see [7]. For *heterogeneous variance* models see [13 14].

Multi-arm trials on treatments A,B,C induce a covariance between δ_{jAB} and δ_{jAC}. Under homogeneous variance the covariance is $\sigma^2/2$ [7 10 13].

WinBUGS 1.4 Programmes for MTC analysis.

There are currently four programmes available on this website
1. FIXED EFFECTS MODEL
2. SIMPLE RANDOM EFFECTS MODEL
3. RANDOM EFFFECTS MODEL : up to 3-arm trials

4. RANDOM EFFECTS MODEL : multi-arm trials

We provide the full code as text, and the odc files can also be accessed from the Mixed Treatment Comparisons web-page and used by anyone with WinBUGS 1.4.3 [21]. This software can be downloaded free of charge from http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml

DATA STRUCTURE

Constants to be entered
N= Number of arms
NS = Number of Studies
NT = Number of Treatments

Labelling the treatments
The programs below are designed to be completely general, but there are constraints on which set of paired comparisons can be represented by *basic* parameters [14]. The following rules should be adhered to:
1. The treatments labelled A,B,C, … etc. become treatments number 1,2,3, … etc in the WinBUGS code
2. Choose any treatment as treatment A. (It may be useful if treatment A represents 'standard care').
3. Set out the trials systematically, as in Table 1, and maintain this order in the WinBUGS data listing.
(a) start with all the trials including treatment A,
(b) list the AB first, then the AC, AD, etc
(c) then the trials including B *but not A.*
(d) then those including C, but not A or B, etc …

The data list for WinBUGS
Vectors of length N are set out in columns; each row represents a single arm. Arms from each trial should be consecutive and in alphabetical order.

s[] indicating the study
t[] the treatment
r[] the numerator
n[] the denominator
b[] the comparator treatment (baseline) for that trial, $b[i] \leq t[i]$

list(N=50, NS=24, NT=4)

s[]	t[]	r[]	n[]	b[]
1	1	79	702	1
1	2	77	694	1
2	1	18	671	1
2	2	21	535	1
3	1	8	116	1
3	2	19	149	1
4	1	75	731	1
4	3	363	714	1
5	1	2	106	1
5	3	9	205	1
6	1	58	549	1
6	3	237	1561	1
7	1	0	33	1
7	3	9	48	1

```
8  1    3  100  1
8  3   31   98  1
9  1    1   31  1
9  3   26   95  1
10 1    6   39  1
10 3   17   77  1
11 1   64  642  1
11 3  107  761  1
12 1    5   62  1
12 3    8   90  1
13 1   20  234  1
13 3   34  237  1
14 1   95 1107  1
14 3  143 1031  1
15 1   15  187  1
15 3   36  504  1
16 1   78  584  1
16 3   73  675  1
17 1   69 1177  1
17 3   54  888  1
18 1    9  140  1
18 3   23  140  1
18 4   10  138  1
19 1    0   20  1
19 4    9   20  1
20 2   20   49  2
20 3   16   43  2
21 2   11   78  2
21 3   12   85  2
21 4   29  170  2
22 2    7   66  2
22 4   32  127  2
23 3   12   76  3
23 4   20   74  3
24 3    9   55  3
24 4    3   26  3
END
```

FIXED EFFECTS MODEL

```
for(i in 1:N)  {     logit(p[s[i],t[i]])<-mu[s[i]]+ d[t[i]] - d[b[i]]                    # model
                     r[i]~dbin(p[s[i],t[i]],n[i])  }                          #  binomial likelihood

for(j in 1:NS)  {   mu[j]~dnorm(0,.0001)}                  # vague priors for 24 trial baselines

d[1]<-0
for (k in 2:NT)  {d[k] ~ dnorm(0,.0001) }          # vague priors for 3 basic LOR parameters
```

SIMPLE RANDOM EFFECTS MODEL

The simple RE model treats M-arm trials (M>2) without taking account of the correlations between the (M-1) trial-specific LORs that they estimate

```
for(i in 1:N)  {      logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))                  #  model
                      r[i]~dbin(p[i],n[i])                                                #  binomial likelihood
                      delta[i] ~ dnorm(md[i],tau)                    # random effects: trial-specific LORs
                      md[i] <- d[t[i]] - d[b[i]]   }                       # means of trials-specific LORS

for(j in 1:NS)  { mu[j]~dnorm(0,.0001) }                          # vague priors for 24 trial baselines

d[1]<-0
for (k in 2:NT)  {d[k] ~ dnorm(0,.0001) }                         # vague priors for basic parameters

sd~dunif(0,2)                                  #  vague prior for random effects standard deviation
```

```
tau<-1/pow(sd,2)
```

SOME ADDITIONAL CODE

Some additional code can be added to each program to:

1. To Form a Baseline

Add the LORs to the baseline to find the absolute efficacy T[k] of each treatment k, given some assumed baseline probability of success on treatment A. The treatment A baseline used here is taken from a separate WinBUGS analysis of the 19 treatment A arms appearing in the dataset. However, this should NOT be taken as an endorsement of such a procedure in general. A baseline can be defined in many other ways: cohort studies, a single trial or set of trials considered to reflect contemporary outcomes under treatment A, expert opinion, etc. There is, in addition, the possibility of putting a model on the baseline as well as the relative treatment effect, though this runs the risk of biasing the treatment effect estimates, if the baseline model is not correct.

```
# Absolute log odds(success) on Treatment A,  based on a seperate model on the
#  19 trials Treatment A arms.
mA ~ dnorm(-2.585,2.763)
# Absolute pr(success) Treatments B,C,D based on T[1] and the
#   MEAN  Relative treatment effects
for (k in 1:NT)   { logit(T[k])<- mA  +d[k] }
```

2. To Rank the treatments in efficacy, and calculate the probability that each is best: best[]. With smoking cessation, higher values of T[] are 'better'. If the data were fatalities, for example, this code would require adjustment, otherwise it would give the probability that each treatment was the worst.

```
# Ranking and prob{treatment k is best}
for (k in 1:NT) { rk[k]<- NT+1 - rank(T[],k)
                  best[k]<-equals(rk[k],1)}
```

3. Calculate all the pair-wise odds ratios between treatments in the MTC analysis: This code generate all the possible LORs and ORs

```
# pairwise ORs
for (c in 1:(NT-1))
        {  for (k in (c+1):NT)
             {  lor[c,k] <- d[k] - d[c]
                log(or[c,k]) <- lor[c,k]
             }
         }
```

FULL RANDOM EFFECTS MODELS

The full RE model takes into account correlation structure induced by multi-arm trials. [7 10] Although the correlation structure may not make a great difference if the number and/or size of multi-arm trials is small, it is advisable to take correlation into account.

1. Extension to datasets with 3-arm trials

Arms belonging to the same trial must appear consecutively in the data list, and a further data vector m[] takes the values 1,2, for the arms of a two arm trials, 1,2,3 for a 3 arm, and so on:

```
s[]  t[]  r[]   n[]   b[]  m[]
1    1    79    702    1   1
1    2    77    694    1   2
2    1    18    671    1   1
2    2    21    535    1   2
3    1     8    116    1   1
.
.
20   3    16     43    2   2
21   2    11     78    2   1
21   3    12     85    2   2
21   4    29    170    2   3
```

The code relies on a realisation of the bivariate normal distribution as a univariate marginal distribution and a univariate conditional distribution:

$$\text{If} \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2/2 \\ \sigma^2/2 & \sigma^2 \end{pmatrix} \right]$$

$$\text{then} \qquad x_1 \sim N(\mu_1, \sigma^2), \quad \text{and} \quad x_2 \mid x_1 \sim N(\mu_2 + \frac{1}{2}(x_1 - \mu_1), \frac{3}{4}\sigma^2)$$

```
sw[1] <- 0
for(i in 1:N)  {
      logit(p[i])<-mu[s[i]]+ delta[i]  * (1-equals(t[i],b[i]))                    #  model
      r[i]~dbin(p[i],n[i])                                                       #  binomial likelihood
       delta[i] ~ dnorm(md[i],taud[i])                                           # trial-specific LOR distributions
       taud[i] <- tau * (1 + equals(m[i],3) /3)                                  # precisions of LOR distributions
      md[i] <- d[t[i]] - d[b[i]]  +  equals(m[i],3) * sw[i]                      # means of LOR distributions
      }
for (i in 2:N)  {   sw[i] <- (delta[i-1] -  d[t[i-1]] + d[b[i-1]])/2}           # adjustment for 3-arm trials

for(j in 1:NS){ mu[j]~dnorm(0,.0001) }                                          # vague priors for 24 trial baselines

d[1]<-0
for (k in 2:NT)  {d[k] ~ dnorm(0,.0001) }                                       # vague priors for basic parameters

sd~dunif(0,2)                                                                   #  vague prior for random effects standard deviation
tau<-1/pow(sd,2)
```

2. General code for multi-arm trials

A completely general programme is given below. Like the above this is based on decomposition of multivariate normal as a series of conditional univariate distributions. We rely on a more highly structured data listing, including vectors of coefficients, and a data layout with each record represent a trial and the columns r[,k] and n[,k] giving numerators and denominators treatment k on each trial, a set of indicators t[,] show which treatments were compared, and the number of arms na[]. If

$$\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2/2 & \dots & \sigma^2/2 \\ \sigma^2/2 & \sigma^2 & \cdots & \sigma^2/2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2/2 & \sigma^2/2 & \cdots & \sigma^2 \end{pmatrix} \right)$$

then the conditional univariate distributions are:

$$x_i \mid \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \end{pmatrix} \sim N(\mu_i + \frac{1}{i}\sum_{j=1}^{i-1}(x_j - \mu_j), \frac{(i+1)}{2i}\sigma^2)$$

```
# NT=no. treatments, NS=no. studies;
# NB : set up M vectors each r[,]. n[,] and t[,],  where M is the Maximum number of treatments
#         per trial in the dataset. In this dataset M is 3.

list(NT=4,NS=24)

r[,1]        n[,1]        r[,2]        n[,2]        r[,3]        n[,3]        t[,1]  t[,2] t[,3]   na[]
 9            140          23           140          10           138          1     3     4      3
11            78           12           85           29    170             2     3     4      3
75            731          363          714          NA           1            1     3     NA     2
 2            106          9            205          NA           1            1     3     NA     2
58            549          237          1561         NA           1     1      3     NA     2
 0            33           9            48           NA           1            1     3     NA     2
 3            100          31           98           NA           1            1     3     NA     2
 1            31           26           95           NA           1            1     3     NA     2
 6            39           17           77           NA           1            1     3     NA     2
79            702          77           694          NA           1     1      2     NA     2
18            671          21           535          NA           1            1     2     NA     2
64            642   107                 761          NA           1            1     3     NA     2
 5            62           8            90           NA           1            1     3     NA     2
20            234          34           237          NA           1            1     3     NA     2
 0            20           9            20           NA           1            1     4     NA     2
 8            116          19           149          NA           1            1     2     NA     2
95           1107  143    1031         NA           1            1     3     NA     2
15            187          36           504          NA           1            1     3     NA     2
78            584          73           675          NA           1            1     3     NA     2
69    1177 54              888          NA           1            1     3     NA     2
20            49           16           43           NA           1            2     3     NA     2
 7            66           32    127    NA           1            2     4     NA     2
12            76           20           74           NA           1            3     4     NA     2
 9            55           3            26           NA           1            3     4     NA     2
END

#Random effects model for multi-arm trials (any number of arms)

model{
for(i in 1:NS){
        w[i,1] <-0
                delta[i,t[i,1]]<-0
                mu[i] ~ dnorm(0,.0001)                                  # vague priors for 24 trial baselines
                for (k in 1:na[i])  {
                     r[i,k] ~ dbin(p[i,t[i,k]],n[i,k])                                  # binomial likelihood
                                logit(p[i,t[i,k]])<-mu[i] + delta[i,t[i,k]] }                                  #
model
                          for (k in 2:na[i]) {
              delta[i,t[i,k]] ~ dnorm(md[i,t[i,k]],taud[i,t[i,k]])          # trial-specific LOR distributions
              md[i,t[i,k]] <-  d[t[i,k]] - d[t[i,1]]  + sw[i,k]             # mean of LOR distributions
               taud[i,t[i,k]] <- tau *2*(k-1)/k                             #precision of LOR distributions
               w[i,k] <- (delta[i,t[i,k]]  - d[t[i,k]] + d[t[i,1]])         #adjustment, multi-arm RCTs
               sw[i,k] <-sum(w[i,1:k-1])/(k-1) }              # cumulative adjustment for multi-arm trials
  }

d[1]<-0
for (k in 2:NT){d[k] ~ dnorm(0,.0001) }                 #  vague priors for basic parameters

sd~dunif(0,2)                                   #  vague prior for random effects standard deviation
tau<-1/pow(sd,2)
```

An interesting feature of this code is that it can readily be made to generate estimates of *all* the p_{jk}, ie the probability of success on each treatment in each trial. This can be contrived by arranging the dataset so that there are 4 four treatments in each trial; where a treatment is in reality missing an NA is put into the numerator column, and 1 into the denominator column.

The full random effects models code reflect the fundamental assumption that every trial is a sample from a multivariate-normal distribution of the (basic) relative treatment effects, with an overlaid missing at random process. For example, a trial of treatments B, C, and D will provide more information about treatment A (in that trial) than a trial of just treatments B and C, and this too is captured in the coding.

References

1. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for meta-analysis in medical research*. London: Wiley, 2000.

2. Bucher HC, Griffith L, Guyatt GH, Opravil M. Meta-analysis of prophylactic treatments against Pneumocystis Carinii pneumonia and toxoplasma encephalitis in HIV-infected patients. *J AIDS and Human Retrovir* 1997;**15**:104-114.

3. Yazdanpanah Y, Sissoko D, Egger M, Mouton Y, Zwahlen M, Chene G. Clinical efficacy of antiretroviral combination therapy based on protease inhibitors or non-nucleoside analogue reverse transcriptase onhibitors: indirect comparison of controlled trials. *British Medical Journal* 2004;**328**:249-253.

4. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The Results of Direct and Indirect Treatment Comparisons in Meta-Analysis of Randomized Controlled Trials. *Journal of Clinical Epidemiology* 1997;**50**:683-691.

5. Song F, Altman D, Glenny M-A, Deeks J. Validity of indirect comparison for estimating efficacy of competing interventions: evidence from published meta-analyses. *British Medical Journal* 2003;**326**:472-476.

6. Gleser LJ, Olkin I. Stochastically dependent effect sizes. In: Cooper H, Hedges LV, eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994;339-355.

7. Higgins JPT, Whitehead J. Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine* 1996;**15**:2733-2749.

8. Hasselblad V. Meta-analysis of multi-treatment studies. *Medical Decision Making* 1998;**18**:37-43.

9. Domenici F, Parmigiani G, Wolpert RL, Hasselblad V. Meta-analysis of migraine headache treatments: combining information from heterogenous designs. *Journal of the American Statistical Association* 1999;**94**:16-28.

10. Whitehead A. *Meta-analysis of controlled clinical trials*. Chichester, UK: 2002.

11. Lumley T. Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine* 2002;**21**:2313-2324.

12. Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter evidence synthesis and consistency of evidence. *Statistics in Medicine* 2003;**22**:2995-3016.

13. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 2004;**23**:3105-3124.

14. Lu G, Ades AE. Assessing evidence consistency in mixed treatment comparisons. *Journal of the American Statistical Association*, 2006; 101: 447-459.

15. Psaty BM, Lumley T, Furberg CD, et al. Health outcomes associated with various antihypertensive therapies used as first-line agents: a network meta-analysis. *JAMA*, 2003;**289**:2534-2544.

16. Bridle C, Bagnall A, Duffy S, et al. A rapid and systematic review of the clinical and cost-effectiveness of newer drugs for treatment of mania associated with bipolar affective disorder. *Health Technology Assessment* 2004;**8**:

17. Wilby, J, Kainth, A, McDaid, D, McIntosh, H, Golder, S, O'Meara, S, Forbes, C, Hawkins, N., Epstein, D, Mason, A, Sculpher M, and Drummond M. A rapid and systematic review of the clinical effectiveness, tolerability and cost effectiveness of newer drugs for epilepsy in adults. *Health Technology Assessment*, 2005,9:(15).

18. Eddy DM, Hasselblad V, Shachter R. *Meta-analysis by the Confidence Profile Method*. London: Academic Press, 1992.

19. Spiegelhalter DJ, Abrams KR, Myles J. *Bayesian approaches to clinical trials and Health-Care Evaluation*. New York: Wiley, 2004.

20. Lambert P, Sutton A, Burton P, Abrams KJ. How vague is vague? Assessment of the use of vague prior distributions for variance componenets. *Statistics in Medicine* 2005;(in press)

21. Spiegelhalter DJ, Thomas A, Best N, Lunn D. *WinBUGS User Manual: Version 1.4*. Cambridge, UK: MRC Biostatistics Unit, 2001.

Document last up-dated 20th December 2007