CIMT study
Ultrasound and metabolic data

SDC / LLCh
http://BendixCarstensen.com/SDC/CIMT
September 2013
1 (preliminary)

Compiled Tuesday 24th September, 2013, 00:34
from: C:/Bendix/Steno/LLCh/CIMT/EASD.tex

Bendix Carstensen Steno Diabetes Center, Gentofte, Denmark
& Department of Biostatistics, University of Copenhagen
bxc@steno.dk
http://BendixCarstensen.com

Contents

1 Construction of data **1**

2 Primary outcome (CIMT) **8**
2.1 Analysis as repeated measures 8
2.2 Analysis using baseline as covariate 12
2.2.1 Model 12
2.3 Relationship between approaches 14
2.3.1 Fitting the two models on complete data 15
2.3.2 Using the randomization assumption 18
2.3.3 Mean change in each group 19

Chapter 1

Construction of data

First we read the ultrasound data for *all* participants. Note that some of the numerical variables are with "." as decimal separator, even if we have a file from a Danish Locale:

```
> imt <- read.csv2("./data/imt.csv")
> str( imt )
```

```
'data.frame':        1544 obs. of  20 variables:
 $ subjid           : int  10002 70032 10002 10002 91089 60001 60001 80010 91081 91118 ...
 $ visit            : Factor w/ 2 levels "1a","7a": 1 2 1 2 1 1 1 2 1 1 ...
 $ cca_side         : Factor w/ 2 levels "L","R": 1 1 2 2 1 1 2 2 2 1 ...
 $ datescanned      : Factor w/ 805 levels "2008/05/06 - 10:45:19",..: 24 726 24 320 177 8 8 569 147
 $ mean_fimtavg     : Factor w/ 85 levels "0.44","0.46",..: 56 38 39 31 43 40 31 14 64 61 ...
 $ mean_fimtmin     : num  0.91 0.69 0.68 0.58 0.63 0.69 0.6 0.45 0.94 0.86 ...
 $ mean_fimtmax     : num  1.14 0.94 1 0.93 1.08 0.99 0.97 0.71 1.28 1.23 ...
 $ minvesseldia     : num  NA NA NA NA NA NA NA NA NA NA ...
 $ maxvesseldia     : Factor w/ 368 levels "","10","10.03",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ vesselareal      : num  NA NA NA NA NA NA NA NA NA NA ...
 $ lumenareal       : num  NA NA NA NA NA NA NA NA NA NA ...
 $ imtareal         : Factor w/ 1120 levels "","10.27300798",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ systolicpressure : int  115 163 115 140 119 145 145 151 145 124 ...
 $ diastolicpressure: int  69 84 69 72 73 89 89 97 68 69 ...
 $ ddpct            : num  NA NA NA NA NA NA NA NA NA NA ...
 $ csdpct           : num  NA NA NA NA NA NA NA NA NA NA ...
 $ dc               : num  NA NA NA NA NA NA NA NA NA NA ...
 $ csc1             : num  NA NA NA NA NA NA NA NA NA NA ...
 $ csc2             : num  NA NA NA NA NA NA NA NA NA NA ...
 $ iem              : num  NA NA NA NA NA NA NA NA NA NA ...
```

```
> names( imt )[c(5,9,12)]
```

```
[1] "mean_fimtavg" "maxvesseldia" "imtareal"
```

```
> for( i in c(5,9,12) ) imt[,i] <- as.numeric(as.character(imt[,i]))
> imt$datescanned <- as.Date( substr(imt$datescanned,1,10) )
> str(imt)
```

```
'data.frame':        1544 obs. of  20 variables:
 $ subjid           : int  10002 70032 10002 10002 91089 60001 60001 80010 91081 91118 ...
 $ visit            : Factor w/ 2 levels "1a","7a": 1 2 1 2 1 1 1 2 1 1 ...
 $ cca_side         : Factor w/ 2 levels "L","R": 1 1 2 2 1 1 2 2 2 1 ...
 $ datescanned      : Date, format: "2008-09-23" "2011-12-09" ...
 $ mean_fimtavg     : num  1.01 0.83 0.84 0.76 0.88 0.85 0.76 0.59 1.09 1.06 ...
```

```
 $ mean_fimtmin     : num  0.91 0.69 0.68 0.58 0.63 0.69 0.6 0.45 0.94 0.86 ...
 $ mean_fimtmax     : num  1.14 0.94 1 0.93 1.08 0.99 0.97 0.71 1.28 1.23 ...
 $ minvesseldia     : num  NA NA NA NA NA NA NA NA NA NA ...
 $ maxvesseldia     : num  NA NA NA NA NA NA NA NA NA NA ...
 $ vesselareal      : num  NA NA NA NA NA NA NA NA NA NA ...
 $ lumenareal       : num  NA NA NA NA NA NA NA NA NA NA ...
 $ imtareal         : num  NA NA NA NA NA NA NA NA NA NA ...
 $ systolicpressure : int  115 163 115 140 119 145 145 151 145 124 ...
 $ diastolicpressure: int  69 84 69 72 73 89 89 97 68 69 ...
 $ ddpct            : num  NA NA NA NA NA NA NA NA NA NA ...
 $ csdpct           : num  NA NA NA NA NA NA NA NA NA NA ...
 $ dc               : num  NA NA NA NA NA NA NA NA NA NA ...
 $ csc1             : num  NA NA NA NA NA NA NA NA NA NA ...
 $ csc2             : num  NA NA NA NA NA NA NA NA NA NA ...
 $ iem              : num  NA NA NA NA NA NA NA NA NA NA ...
```

```
> addmargins( with( imt, table(visit,cca_side) ) )
```

```
     cca_side
visit    L    R  Sum
  1a   411  427  838
  7a   343  363  706
  Sum  754  790 1544
```

```
> with( imt, table( table( subjid ) ) )
```

```
  1   2   3   4
  5  81  15 333
```

The we read the randomization codes for those actually randomized:

```
> rnd <- read.csv2("./data/PTNR_MetforminCode_InsulinTypeCode.csv" )
> str( rnd )
```

```
'data.frame':        412 obs. of  3 variables:
 $ PTNR           : int  70002 70001 70004 70005 70006 70007 70010 70008 70009 70011 ...
 $ MetforminCode  : int  1 1 0 0 0 0 0 1 0 0 ...
 $ InsulinTypeCode: int  3 3 3 2 2 3 2 2 2 1 ...
```

By merging with `all.y=TRUE`, we ensure that only those from the randomization set are included:

```
> dim( imt )
```

```
[1] 1544   20
```

```
> rimt <- merge( imt, rnd, by.x="subjid", by.y="PTNR", all.y=TRUE )
> dim( rimt )
```

```
[1] 1498   22
```

```
> str( rimt )
```

```
'data.frame':        1498 obs. of  22 variables:
 $ subjid           : int  10001 10001 10001 10001 10002 10002 10002 10003 10003 10003 ...
 $ visit            : Factor w/ 2 levels "1a","7a": 1 2 2 1 1 1 2 2 1 2 ...
 $ cca_side         : Factor w/ 2 levels "L","R": 2 1 2 1 1 2 2 1 2 2 ...
 $ datescanned      : Date, format: "2008-09-22" "2010-03-22" ...
 $ mean_fimtavg     : num  0.8 0.84 0.88 0.85 1.01 0.84 0.76 0.88 1.04 1.02 ...
 $ mean_fimtmin     : num  0.54 0.63 0.71 0.65 0.91 0.68 0.58 0.7 0.87 0.86 ...
 $ mean_fimtmax     : num  1 1.03 1.03 0.96 1.14 1 0.93 1.03 1.21 1.2 ...
 $ minvesseldia     : num  NA 10.8 NA 10.9 NA ...
 $ maxvesseldia     : num  NA 11.5 NA 11.5 NA ...
 $ vesselareal      : num  NA 104 NA 103 NA ...
 $ lumenareal       : num  NA 76 NA 74.8 NA ...
 $ imtareal         : num  NA 28.2 NA 28.3 NA ...
 $ systolicpressure : int  142 124 124 142 115 115 140 154 153 154 ...
 $ diastolicpressure: int  86 73 73 86 69 69 72 85 95 85 ...
 $ ddpct            : num  NA 7.05 NA 5.51 NA NA NA 5.39 3.89 5.57 ...
 $ csdpct           : num  NA 14.6 NA 11.3 NA ...
 $ dc               : num  NA 0.0149 NA 0.0107 NA NA NA 0.00633 0.00608 0.00693 ...
 $ csc1             : num  NA 0.26 NA 0.19 NA NA NA 0.0829 0.0883 0.0961 ...
 $ csc2             : num  NA 0.00286 NA 0.00202 NA NA NA 0.0016 0.00137 0.00166 ...
 $ iem              : num  NA 1914 NA 2680 NA ...
 $ MetforminCode    : int  1 1 1 1 0 0 0 0 0 0 ...
 $ InsulinTypeCode  : int  3 3 3 3 3 3 3 1 1 1 ...
```

```
> with( rimt, table(table(subjid)) )
```

```
  1   2   3   4
  5  60  15 332
```

Fishily enough there is one person with 4 observations in `imt` that is not randomized:

```
> tt <- with( imt, table( subjid ) )
> has4i <- names( tt[tt==4] )
> isrnd <- rnd$PTNR
> setdiff( has4i, isrnd )
```

```
[1] "91218"
```

```
> with( rimt, pairs( cbind( avg = as.numeric(as.character(mean_fimtavg)),
+                           min = mean_fimtmin,
+                           max = mean_fimtmax,
+                            mn = (mean_fimtmin+mean_fimtmax)/2 ),
+                    gap=0, pch=16, cex=0.5 ) )
```

We now take the average for left and right-sided measurements, but also merge in as separate variables the left and right side measurements:

```
> what <- 5:22
> names( rimt )[what]
```

```
 [1] "mean_fimtavg"      "mean_fimtmin"      "mean_fimtmax"
 [4] "minvesseldia"      "maxvesseldia"      "vesselareal"
 [7] "lumenareal"        "imtareal"          "systolicpressure"
[10] "diastolicpressure" "ddpct"             "csdpct"
[13] "dc"                "csc1"              "csc2"
[16] "iem"               "MetforminCode"     "InsulinTypeCode"
```

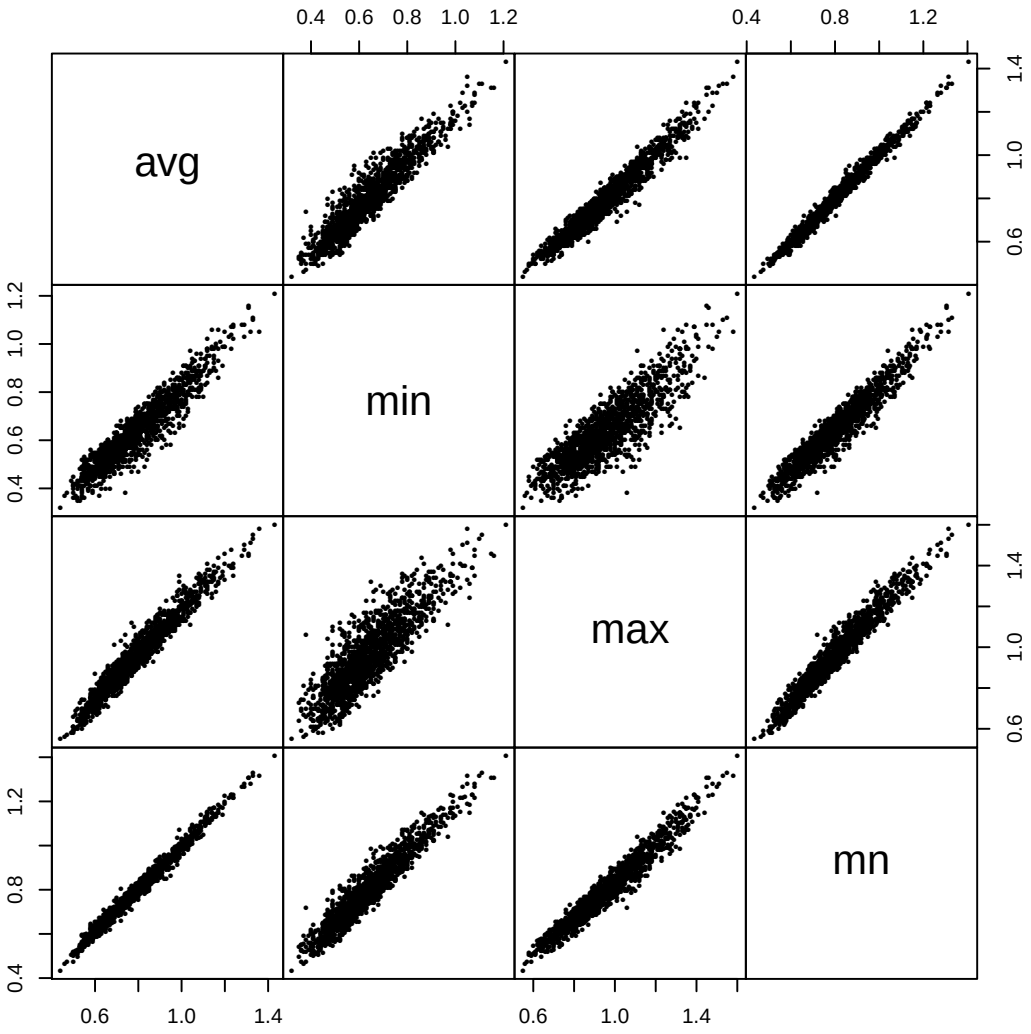

Figure 1.1: *Pairwise comparison of far wall intima media thickness as the average, min, max and the mean of min and max.*

```
> mimt <- aggregate( rimt[,what], list( subjid=rimt$subjid,
+                                       visit=rimt$visit),
+                    FUN=mean, na.rm=TRUE )
> wh <- c(1,2,5:12,18:20)
> names( imt )[wh]
```

```
 [1] "subjid"       "visit"        "mean_fimtavg" "mean_fimtmin" "mean_fimtmax"
 [6] "minvesseldia" "maxvesseldia" "vesselareal"  "lumenareal"   "imtareal"
[11] "csc1"         "csc2"         "iem"
```

```
> Limt <- subset( rimt, cca_side=="L" )[,wh]
> Rimt <- subset( rimt, cca_side=="R" )[,wh]
> names( Limt ) <- gsub( "mean_", "", names( Limt ) )
> names( Rimt ) <- gsub( "mean_", "", names( Rimt ) )
> names( Limt )[-(1:2)] <- paste( "left_" , names( Limt )[-(1:2)], sep="" )
> names( Rimt )[-(1:2)] <- paste( "right_", names( Rimt )[-(1:2)], sep="" )
> mimt <- merge( merge( mimt, Limt, all.x=TRUE ), Rimt, all.x=TRUE )
> head( mimt )
```

```
  subjid visit mean_fimtavg mean_fimtmin mean_fimtmax minvesseldia maxvesseldia
1  10001    1a        0.825        0.595        0.980       10.860       11.460
2  10001    7a        0.860        0.670        1.030       10.760       11.520
3  10002    1a        0.925        0.795        1.070          NaN          NaN
4  10002    7a        0.760        0.580        0.930          NaN          NaN
5  10003    1a        0.940        0.725        1.105        8.610        8.960
6  10003    7a        0.950        0.780        1.115        8.345        8.805
  vesselareal lumenareal imtareal systolicpressure diastolicpressure ddpct
1   103.14760   74.81514 28.33245              142                86 5.510
2   104.23050   76.04665 28.18386              124                73 7.050
3         NaN        NaN      NaN              115                69   NaN
4         NaN        NaN      NaN              140                72   NaN
5    63.21921   39.42228 23.79694              153                95 4.095
6    60.94144   37.45741 23.48403              154                85 5.480
  csdpct      dc    csc1    csc2      iem MetforminCode InsulinTypeCode
1 11.330 0.01070 0.19000 0.00202 2680.310             1               3
2 14.600 0.01490 0.26000 0.00286 1913.970             1               3
3    NaN     NaN     NaN     NaN      NaN             0               3
4    NaN     NaN     NaN     NaN      NaN             0               3
5  8.355 0.00606 0.08365 0.00144 3520.270             0               1
6 11.260 0.00663 0.08950 0.00163 3080.515             0               1
  left_fimtavg left_fimtmin left_fimtmax left_minvesseldia left_maxvesseldia
1         0.85         0.65         0.96             10.86              11.46
2         0.84         0.63         1.03             10.76              11.52
3         1.01         0.91         1.14                NA                 NA
4           NA           NA           NA                NA                 NA
5         0.84         0.58         1.00              8.15               8.50
6         0.88         0.70         1.03              8.11               8.55
  left_vesselareal left_lumenareal left_imtareal left_csc1 left_csc2 left_iem
1        103.14760        74.81514      28.33245    0.1900   0.00202  2680.31
2        104.23050        76.04665      28.18386    0.2600   0.00286  1913.97
3               NA              NA            NA        NA        NA       NA
4               NA              NA            NA        NA        NA       NA
5         56.74502        36.53075      20.21426    0.0790   0.00151  3262.22
6         57.41457        36.21008      21.20449    0.0829   0.00160  3099.45
  right_fimtavg right_fimtmin right_fimtmax right_minvesseldia
1          0.80          0.54          1.00                NA
2          0.88          0.71          1.03                NA
3          0.84          0.68          1.00                NA
4          0.76          0.58          0.93                NA
5          1.04          0.87          1.21              9.07
6          1.02          0.86          1.20              8.58
  right_maxvesseldia right_vesselareal right_lumenareal right_imtareal
```

```
1                NA                NA               NA             NA
2                NA                NA               NA             NA
3                NA                NA               NA             NA
4                NA                NA               NA             NA
5              9.42          69.69341         42.31380       27.37961
6              9.06          64.46831         38.70474       25.76357
  right_csc1 right_csc2 right_iem
1         NA         NA        NA
2         NA         NA        NA
3         NA         NA        NA
4         NA         NA        NA
5     0.0883    0.00137   3778.32
6     0.0961    0.00166   3061.58
```

```
> subset( rimt, subjid==10001 )
```

```
  subjid visit cca_side datescanned mean_fimtavg mean_fimtmin mean_fimtmax
1  10001    1a        R  2008-09-22         0.80         0.54         1.00
2  10001    7a        L  2010-03-22         0.84         0.63         1.03
3  10001    7a        R  2010-03-22         0.88         0.71         1.03
4  10001    1a        L  2008-09-22         0.85         0.65         0.96
  minvesseldia maxvesseldia vesselareal lumenareal imtareal systolicpressure
1           NA           NA          NA         NA       NA              142
2        10.76        11.52    104.2305   76.04665 28.18386              124
3           NA           NA          NA         NA       NA              124
4        10.86        11.46    103.1476   74.81514 28.33245              142
  diastolicpressure ddpct csdpct     dc csc1    csc2     iem MetforminCode
1                86    NA     NA     NA   NA      NA      NA             1
2                73  7.05  14.60 0.0149 0.26 0.00286 1913.97             1
3                73    NA     NA     NA   NA      NA      NA             1
4                86  5.51  11.33 0.0107 0.19 0.00202 2680.31             1
  InsulinTypeCode
1               3
2               3
3               3
4               3
```

```
> subset( mimt, subjid==10001 )
```

```
  subjid visit mean_fimtavg mean_fimtmin mean_fimtmax minvesseldia maxvesseldia
1  10001    1a        0.825        0.595         0.98        10.86        11.46
2  10001    7a        0.860        0.670         1.03        10.76        11.52
  vesselareal lumenareal imtareal systolicpressure diastolicpressure ddpct
1    103.1476   74.81514 28.33245              142                86  5.51
2    104.2305   76.04665 28.18386              124                73  7.05
  csdpct     dc csc1    csc2     iem MetforminCode InsulinTypeCode left_fimtavg
1  11.33 0.0107 0.19 0.00202 2680.31             1               3         0.85
2  14.60 0.0149 0.26 0.00286 1913.97             1               3         0.84
  left_fimtmin left_fimtmax left_minvesseldia left_maxvesseldia
1         0.65         0.96             10.86             11.46
2         0.63         1.03             10.76             11.52
  left_vesselareal left_lumenareal left_imtareal left_csc1 left_csc2 left_iem
1         103.1476        74.81514      28.33245      0.19   0.00202  2680.31
2         104.2305        76.04665      28.18386      0.26   0.00286  1913.97
  right_fimtavg right_fimtmin right_fimtmax right_minvesseldia
1          0.80          0.54          1.00                 NA
2          0.88          0.71          1.03                 NA
  right_maxvesseldia right_vesselareal right_lumenareal right_imtareal
1                 NA                NA               NA             NA
2                 NA                NA               NA             NA
  right_csc1 right_csc2 right_iem
1         NA         NA        NA
2         NA         NA        NA
```

Finally we attach the stratification variables, and grom the variable names to something slightly more handy:

```
> st <- read.csv2( "./data/strata.csv" )[,-2]
> names( st )[2:4] <- c("o.65","pre.ins","SDC")
> mimt <- merge( mimt, st, by.x="subjid", by.y="ptnr" )
> names( mimt ) <- gsub( "mean_", "", names(mimt) )
```

The resulting data set has one observation per visit among the randomized patients:

```
> mimt <- transform( mimt, subjid = factor(subjid),
+                          met = factor(MetforminCode,labels=c("Plc","Met")),
+                          ins = factor(InsulinTypeCode,labels=c("NR+Lv","Lv","NM30")) )
> with( mimt, table(MetforminCode, met ))


             met
MetforminCode Plc Met
            0 388   0
            1   0 395


> with( mimt, table(InsulinTypeCode,ins))


               ins
InsulinTypeCode NR+Lv  Lv NM30
              1   263   0    0
              2     0 257    0
              3     0   0  263


> with( mimt, ftable(addmargins(table(Met=met, Ins=ins, visit ))) )


          visit  1a  7a Sum
Met Ins
Plc NR+Lv        73  66 139
    Lv           66  55 121
    NM30         67  61 128
    Sum         206 182 388
Met NR+Lv        65  59 124
    Lv           71  65 136
    NM30         70  65 135
    Sum         206 189 395
Sum NR+Lv       138 125 263
    Lv          137 120 257
    NM30        137 126 263
    Sum         412 371 783
```

Finally for the benefit of the client, we write it both an `Rda` format and as a SAS data-file with corresponding script:

```
> save( mimt, file="./data/mimt.Rda" )
> write.foreign( mimt,
+             datafile = "./sas/mimt.dat",
+             codefile = "./sas/mimt.sas",
+              package = "SAS" )
```

Chapter 2

Primary outcome (CIMT)

2.1 Analysis as repeated measures

We reload the data and the two necessary packages from R, and convert the subject indicator to a factor as is needed for use in `lmer`:

```
> library( lme4 )
> library( Epi )
> load( file="./data/mimt.Rda" )
```

We analyze data with a random effects model, for the CIMT-mean y_{it} on individual i at time $t = 1, 7$ (`mean_fimtavg`), randomized to treatment $m = M, P$ using subject as random, and with a separate metformin by time (`visit`) interaction:

$$\begin{aligned} y_{it} &= \mu + \beta_t + \gamma_{mt} \\ &\quad + \alpha_1 \texttt{o.65} + \alpha_2 \texttt{per.ins} + \alpha_1 \texttt{SDC} \\ &\quad + a_i + e_{it}, \quad t = 1, 7 \\ a_i &\sim \mathcal{N}(0, \tau^2), \quad e_{it} \sim \mathcal{N}(0, \sigma^2), \quad \text{all independent} \end{aligned} \tag{2.1}$$

The model states that persons have an average baseline level of CIMT depending on randomization and stratification group; the mean at baseline and follow-up are:

$$\begin{aligned} \text{baseline:} &\quad \mu + \beta_1 + \gamma_{m1} + \alpha_1 \texttt{o.65} + \alpha_2 \texttt{per.ins} + \alpha_3 \texttt{SDC} \\ \text{follow-up:} &\quad \mu + \beta_7 + \gamma_{m7} + \alpha_1 \texttt{o.65} + \alpha_2 \texttt{per.ins} + \alpha_3 \texttt{SDC} \\ \text{change:} &\quad \beta_7 - \beta_1 + (\gamma_{m7} - \gamma_{m1}) \end{aligned}$$

So in the changes in CIMT are:

$$\begin{aligned} \text{metformin:} &\quad \beta_7 - \beta_1 + (\gamma_{\text{met}7} - \gamma_{\text{met}1}) \\ \text{placebo:} &\quad \beta_7 - \beta_1 + (\gamma_{\text{plc}7} - \gamma_{\text{plc}1}) \end{aligned}$$

The model as stated here is grossly overparametrized, it can be identified if all parameters relation to baseline or placebo were set to 0, the changes in CIMT over the follow-up are

$$\begin{aligned} \text{metformin:} &\quad \beta_7 + \gamma_{\text{met}7} \\ \text{placebo:} &\quad \beta_7 \\ \text{difference:} &\quad \gamma_{\text{met}7} \end{aligned}$$

These three parameters are those of interest which should be extracted.

The model is a random effects model that is very close to using the baseline (y_{i1}) as covariate in an analysis of the follow-up (y_{i7}) as outcome:

```
> m0 <- lmer( fimtavg ~ -1 + visit:met + o.65 + pre.ins + SDC +(1|subjid),
+             data = mimt )
> summary( m0 )


Linear mixed model fit by REML
Formula: fimtavg ~ -1 + visit:met + o.65 + pre.ins + SDC + (1 | subjid)
   Data: mimt
   AIC   BIC logLik deviance REMLdev
 -1365 -1323  691.5    -1436   -1383
Random effects:
 Groups   Name        Variance  Std.Dev.
 subjid   (Intercept) 0.0144863 0.120359
 Residual             0.0026388 0.051369
Number of obs: 783, groups: subjid, 412

Fixed effects:
               Estimate Std. Error t value
o.65           0.089748   0.013846    6.48
pre.ins       -0.006157   0.013800   -0.45
SDC           -0.014061   0.012718   -1.11
visit1a:metPlc 0.785330   0.014019   56.02
visit7a:metPlc 0.771776   0.014117   54.67
visit1a:metMet 0.773305   0.014082   54.91
visit7a:metMet 0.772499   0.014178   54.48

Correlation of Fixed Effects:
            o.65   pre.ns SDC    vst1:P vst7:P vst1:M
pre.ins     -0.075
SDC         -0.031 -0.204
vist1:mtPlc -0.208 -0.566 -0.302
vist7:mtPlc -0.206 -0.560 -0.302  0.927
visit1:mtMt -0.221 -0.569 -0.295  0.579  0.574
visit7:mtMt -0.220 -0.566 -0.294  0.576  0.571  0.931


> round( ee <- ci.lin( m0, subset="visit" ), 4 )


              Estimate StdErr       z P   2.5%  97.5%
visit1a:metPlc   0.7853 0.0140 56.0180 0 0.7579 0.8128
visit7a:metPlc   0.7718 0.0141 54.6710 0 0.7441 0.7994
visit1a:metMet   0.7733 0.0141 54.9139 0 0.7457 0.8009
visit7a:metMet   0.7725 0.0142 54.4843 0 0.7447 0.8003


> CM <- rbind(diag(4),c(0,0,-1,1),
+                     c(-1,1,0,0),
+                     c(1,-1,-1,1))
> row.names(CM) <- c("Plc 1a","Plc 7a",
+                    "Met 1a","Met 7a",
+                    "Met 7a-1a",
+                    "Plc 7a-1a",
+                    "dMet - dPlc" )
> colnames(CM) <- row.names(ee)
> CM
```

```
            visit1a:metPlc visit7a:metPlc visit1a:metMet visit7a:metMet
Plc 1a                   1              0              0              0
Plc 7a                   0              1              0              0
Met 1a                   0              0              1              0
Met 7a                   0              0              0              1
Met 7a-1a                0              0             -1              1
Plc 7a-1a               -1              1              0              0
dMet - dPlc              1             -1             -1              1
```

```
> round( e0 <- ci.lin( m0, subset="visit", ctr.mat=CM ), 4 )
```

```
            Estimate StdErr       z      P    2.5%   97.5%
Plc 1a        0.7853 0.0140 56.0180 0.0000  0.7579  0.8128
Plc 7a        0.7718 0.0141 54.6710 0.0000  0.7441  0.7994
Met 1a        0.7733 0.0141 54.9139 0.0000  0.7457  0.8009
Met 7a        0.7725 0.0142 54.4843 0.0000  0.7447  0.8003
Met 7a-1a    -0.0008 0.0053 -0.1529 0.8785 -0.0111  0.0095
Plc 7a-1a    -0.0136 0.0054 -2.5283 0.0115 -0.0241 -0.0030
dMet - dPlc   0.0127 0.0075  1.6963 0.0898 -0.0020  0.0275
```

```
> elab <- c("Metformin 18m - baseline",
+           "Placebo 18m - baseline",
+           "Metformin vs. Placebo change")
> rownames( e0 )[5:7] <- elab
> round( e0, 4 )
```

```
                            Estimate StdErr       z      P    2.5%   97.5%
Plc 1a                        0.7853 0.0140 56.0180 0.0000  0.7579  0.8128
Plc 7a                        0.7718 0.0141 54.6710 0.0000  0.7441  0.7994
Met 1a                        0.7733 0.0141 54.9139 0.0000  0.7457  0.8009
Met 7a                        0.7725 0.0142 54.4843 0.0000  0.7447  0.8003
Metformin 18m - baseline     -0.0008 0.0053 -0.1529 0.8785 -0.0111  0.0095
Placebo 18m - baseline       -0.0136 0.0054 -2.5283 0.0115 -0.0241 -0.0030
Metformin vs. Placebo change  0.0127 0.0075  1.6963 0.0898 -0.0020  0.0275
```

For the sake of plotting we start out by defining the colors to be used for Metformin, Placebo, differences and the axes etc on the transparent plots:

```
> eclr <- c("green","yellow","orange","white")
```

```
> win.metafile( "forest1.emf", width=10, height=5, pointsize=24 )
> par( mar=c(3,1,1,1), mgp=c(3,1,0)/1.6,
+      bg=#"black",
+         "transparent",
+      col.axis=eclr[4], col.lab=eclr[4] )
> plotEst( e0[5:7,c(1,5,6)], lwd=7,cex=1.5,
+          xlab="", col.txt="transparent", col=eclr,
+          restore.par=FALSE )
> for( i in 1:3 ) axis( side=2, at=4-i, labels=elab[i], col="transparent",
+                       col.axis= eclr[i], las=1 )
> abline(v=0,col=eclr[4])
> axis( side=1, col=eclr[4] )
> mtext( "Carotid intima-media thickness (mm)", side=1, line=3/1.6, col=eclr[4] )
> dev.off()
```

```
null device
          1
```

```
> par( mar=c(3,1,1,1), mgp=c(3,1,0)/1.6 )
> plotEst( e0[5:7,c(1,5,6)], lwd=3, cex=1.5, vref=0,
+          xlab="Carotid intima-media thickness (mm)", txt=elab )
```

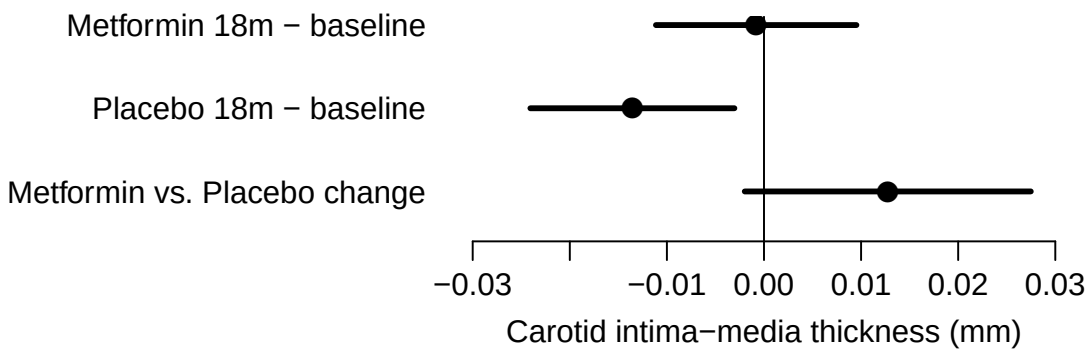

Figure 2.1: *Estimated contrasts from the random effects model.*

2.2 Analysis using baseline as covariate

Another way of looking at the analysis is to use the measurement at the second occasion using the baseline measurement as covariate. In section **??** is a brief account of the relationship between the two approaches.

2.2.1 Model

The main difference is that the absolute change in the response from baseline to follow-up in each of the groups does not appear as a parameter in this approach. The model is (for 7—follow-up and 1—baseline):

$$\begin{aligned} y_{i7} &= \mu + \theta y_{i1} + \delta_t \\ &\quad + \alpha_1 \texttt{o.65} + \alpha_2 \texttt{per.ins} + \alpha_1 \texttt{SDC} \\ &\quad + e_i, \\ e_i &\sim \mathcal{N}(0, \sigma^2) \end{aligned} \tag{2.2}$$

In order to use this approach we restructure the data set so that we have only one row per person with follow-up and baseline as separate variables:

```
> names( mimt )
```

```
 [1] "subjid"             "visit"              "fimtavg"
 [4] "fimtmin"            "fimtmax"            "minvesseldia"
 [7] "maxvesseldia"       "vesselareal"        "lumenareal"
[10] "imtareal"           "systolicpressure"   "diastolicpressure"
[13] "ddpct"              "csdpct"             "dc"
[16] "csc1"               "csc2"               "iem"
[19] "MetforminCode"      "InsulinTypeCode"    "left_fimtavg"
[22] "left_fimtmin"       "left_fimtmax"       "left_minvesseldia"
[25] "left_maxvesseldia"  "left_vesselareal"   "left_lumenareal"
[28] "left_imtareal"      "left_csc1"          "left_csc2"
[31] "left_iem"           "right_fimtavg"      "right_fimtmin"
[34] "right_fimtmax"      "right_minvesseldia" "right_maxvesseldia"
[37] "right_vesselareal"  "right_lumenareal"   "right_imtareal"
[40] "right_csc1"         "right_csc2"         "right_iem"
[43] "o.65"               "pre.ins"            "SDC"
[46] "met"                "ins"
```

```
> wimt <- reshape( mimt[,c("fimtavg","visit","subjid",
+                          "met","o.65","pre.ins","SDC")],
+                 direction = "wide",
+                   v.names = "fimtavg",
+                   timevar = "visit",
+                     idvar = "subjid" )
> subset( mimt[,c("fimtavg","visit","subjid",
+                          "MetforminCode","o.65","pre.ins","SDC")],
+         subjid=="10002" )
```

```
  fimtavg visit subjid MetforminCode o.65 pre.ins SDC
3   0.925    1a  10002             0    1       0   0
4   0.760    7a  10002             0    1       0   0
```

```
> subset( wimt, subjid=="10002" )
```

```
  subjid met o.65 pre.ins SDC fimtavg.1a fimtavg.7a
3  10002 Plc    1       0   0      0.925       0.76
```

We now have a dataset with the relevant variables, where we can estimate the difference in changes between the two groups:

```
> mf <- lm( fimtavg.7a ~ fimtavg.1a + met + o.65 + pre.ins + SDC,
+           data = wimt )
> summary( mf )


Call:
lm(formula = fimtavg.7a ~ fimtavg.1a + met + o.65 + pre.ins +
    SDC, data = wimt)

Residuals:
      Min        1Q    Median        3Q       Max
-0.246178 -0.039118  0.000996  0.042370  0.218314

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.111079   0.023192   4.790 2.43e-06
fimtavg.1a   0.839373   0.027631  30.378  < 2e-16
metMet       0.010418   0.007252   1.437   0.1517
o.65         0.020962   0.008412   2.492   0.0132
pre.ins      0.005352   0.008056   0.664   0.5069
SDC         -0.011232   0.007432  -1.511   0.1316

Residual standard error: 0.06966 on 365 degrees of freedom
  (41 observations deleted due to missingness)
Multiple R-squared:  0.7439,        Adjusted R-squared:  0.7404
F-statistic:   212 on 5 and 365 DF,  p-value: < 2.2e-16


> mm <- lm( fimtavg.7a ~ fimtavg.1a + met,
+           data = wimt )
> summary( mm )


Call:
lm(formula = fimtavg.7a ~ fimtavg.1a + met, data = wimt)

Residuals:
      Min        1Q    Median        3Q       Max
-0.259533 -0.038533  0.000097  0.042784  0.218626

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.098472   0.021941   4.488 9.63e-06
fimtavg.1a  0.859657   0.026706  32.190  < 2e-16
metMet      0.011424   0.007296   1.566    0.118

Residual standard error: 0.07018 on 368 degrees of freedom
  (41 observations deleted due to missingness)
Multiple R-squared:  0.7379,        Adjusted R-squared:  0.7365
F-statistic: 518.1 on 2 and 368 DF,  p-value: < 2.2e-16


> round( cf <- ci.lin( mf, subset=c("fimt","Met") ), 4 )


           Estimate StdErr       z      P    2.5%  97.5%
fimtavg.1a   0.8394 0.0276 30.3782 0.0000  0.7852 0.8935
metMet       0.0104 0.0073  1.4366 0.1508 -0.0038 0.0246
```

```
> round( cm <- ci.lin( mm, subset=c("fimt","Met") ), 4 )
```

```
           Estimate StdErr       z      P    2.5%  97.5%
fimtavg.1a   0.8597 0.0267 32.1899 0.0000  0.8073 0.9120
metMet       0.0114 0.0073  1.5658 0.1174 -0.0029 0.0257
```

These estimates can now be compared with those from the random effects model:

```
> round( e0, 4 )
```

```
                              Estimate StdErr       z      P    2.5%   97.5%
Plc 1a                          0.7853 0.0140 56.0180 0.0000  0.7579  0.8128
Plc 7a                          0.7718 0.0141 54.6710 0.0000  0.7441  0.7994
Met 1a                          0.7733 0.0141 54.9139 0.0000  0.7457  0.8009
Met 7a                          0.7725 0.0142 54.4843 0.0000  0.7447  0.8003
Metformin 18m - baseline       -0.0008 0.0053 -0.1529 0.8785 -0.0111  0.0095
Placebo 18m - baseline         -0.0136 0.0054 -2.5283 0.0115 -0.0241 -0.0030
Metformin vs. Placebo change    0.0127 0.0075  1.6963 0.0898 -0.0020  0.0275
```

The relevant quantities to compare are the additive effects of treatment:

```
> round( rbind( "Simple"=cm[2,],
+               "Contrl"=cf[2,],
+               "RanEff"=e0[7,] ), 4 )
```

```
        Estimate StdErr      z      P    2.5%  97.5%
Simple    0.0114 0.0073 1.5658 0.1174 -0.0029 0.0257
Contrl    0.0104 0.0073 1.4366 0.1508 -0.0038 0.0246
RanEff    0.0127 0.0075 1.6963 0.0898 -0.0020 0.0275
```

Clearly, the conclusion for the models are not substantially different; CIMT difference is some 0.01 mm larger between the groups at follow-up in favor of the Placebo group.

To understand the estimates pertaining to the baseline variable, we can compare them to the ratio of the main effects from the linear mixed model.

2.3 Relationship between approaches

This section is a digression alien to the main analysis in the report; it aims at explaining the relationship between the random effects model and the conditional model.

The usual approach to analysis repeated measures with a baselaine and one follow-up measurement is to use the baseline as covariate, as done in the latter analysis above.

This is a corollary of the basic statistical principle that inference should be made in the conditional distribution given the sufficient statistics for the ancillary parameters, which in this case is the overall individaul-specific value for each person (a_i).

The model (2.1) induces a 2-dimensional normal distribution of the measuremnts y_1 and y_7:

$$\begin{pmatrix} y_1 \\ y_7 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu_1 \\ \mu_7 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_7 \\ \rho\sigma_1\sigma_7 & \sigma_7^2 \end{pmatrix} \right]$$

From standard statistical theory we know that under this model, the conditional distribution of y_7 given y_1 is:

$$y_7|y_1 \sim \mathcal{N}\big(\mu_7 + \frac{\rho\sigma_7}{\sigma_1}(y_1 - \mu_1), \sigma_7^2(1-\rho^2)\big)$$

Now in the model (2.1) we have the following values for the parameters μ_1, μ_7, σ_1^2, σ_7^2 and ρ as above. For convenience we set $\eta = \alpha_1 \texttt{o.65} + \alpha_2 \texttt{per.ins} + \alpha_1 \texttt{SDC}$, and so we have:

$$\begin{aligned}
\mu_1 &= \mu + \delta_m + \eta \\
\mu_7 &= \mu + \delta_m + \beta_7 + \gamma_{m7} + \eta \\
\sigma_1^2 &= \tau^2 + \sigma^2 \\
\sigma_7^2 &= \tau^2 + \sigma^2 \\
\rho &= \frac{\tau^2}{\sigma^2 + \tau^2}
\end{aligned}$$

Note that we have a term in the model, δ_m, allowing the two randomization groups to have different means at baseline; accounting for baseline imbalance despite the randomization.

Using this in the formulae for the conditional distribution, gives the conditional distribution of y_7 given y_1 in terms of the model parameters from (2.1) (well we maintain ρ):

$$\begin{aligned}
y_7|y_1 &\sim \mathcal{N}\big(\mu + \delta_m + \beta_7 + \gamma_{m7} + \eta + \rho\big(y_1 - (\mu + \delta_m + \eta)\big), (\sigma^2 + \tau^2)(1 - \rho^2)\big) \\
&= \mathcal{N}\Big(\big((1-\rho)\mu + \beta_7\big) + \big((1-\rho)\delta_m + \gamma_{m7}\big) + (1-\rho)\eta + \rho y_1, (\sigma^2 + \tau^2)(1 - \rho^2)\Big)
\end{aligned}$$

where the term $(1-\rho)\mu + \beta_7$ should show us as the intercept in the conditional analysis, and the term $(1-\rho)\delta_m + \gamma_{m7}$ as the coefficient to the treatment indicator.

2.3.1 Fitting the two models on complete data

Now fitting the joint model to the complete (y_1, y_7) data and fitting the conditional model to y_7 using y_1 as covariate does not necessarily give the same results, but here is a comparison. Note that as opposed to tha analysis above, we only use persons with complete information:

```
> compl <- mimt$subjid[mimt$visit=="7a"]
> cmimt <- subset( mimt, subjid %in% compl )
> with( cmimt, table( visit ) )
```

```
visit
 1a  7a
371 371
```

```
> mc <- lmer( fimtavg ~ visit*met + o.65 + pre.ins + SDC +(1|subjid),
+             data = cmimt )
> summary( mc )
```

```
Linear mixed model fit by REML
Formula: fimtavg ~ visit * met + o.65 + pre.ins + SDC + (1 | subjid)
   Data: cmimt
   AIC   BIC logLik deviance REMLdev
 -1313 -1271  665.5    -1383   -1331
Random effects:
 Groups   Name        Variance Std.Dev.
 subjid   (Intercept) 0.014578 0.120739
 Residual             0.002640 0.051381
Number of obs: 742, groups: subjid, 371
```

```
Fixed effects:
                Estimate Std. Error t value
(Intercept)     0.787233   0.014677   53.64
visit7a        -0.013544   0.005386   -2.51
metMet         -0.014940   0.013638   -1.10
o.65            0.088787   0.014619    6.07
pre.ins        -0.006101   0.014574   -0.42
SDC            -0.017517   0.013436   -1.30
visit7a:metMet  0.013200   0.007546    1.75

Correlation of Fixed Effects:
            (Intr) visit7 metMet o.65   pre.ns SDC
visit7a     -0.183
metMet      -0.452  0.197
o.65        -0.192  0.000 -0.027
pre.ins     -0.548  0.000 -0.028 -0.086
SDC         -0.306  0.000  0.010 -0.041 -0.217
visit7:mtMt  0.131 -0.714 -0.277  0.000  0.000  0.000
```

We now extract the quantities to campare with the results from the conditional analysis:

```
> sig2  <- attr(VarCorr(mc),"sc")^2
> tau2  <- as.numeric(VarCorr(mc)$subjid)
> rho   <- tau2/(sig2+tau2)
> int   <- fixef( mc )["(Intercept)"]
> beta  <- fixef( mc )["visit7a"]
> delta <- fixef( mc )["metMet"]
> gamma <- fixef( mc )["visit7a:metMet"]
> alpha <- fixef( mc )[4:6]
> cbind( c(rho=rho, int, delta, beta, gamma, alpha) )
```

```
                       [,1]
rho             0.846672253
(Intercept)     0.787232800
metMet         -0.014939592
visit7a        -0.013543956
visit7a:metMet  0.013200041
o.65            0.088787347
pre.ins        -0.006100712
SDC            -0.017516980
```

The corresponding analysis on the complete cases from the `wimt` dataset (after making sure that it is the same persons):

```
> summary( wimt )
```

```
     subjid     met           o.65          pre.ins           SDC
 10001  :  1   Plc:206   Min.   :0.000   Min.   :0.0000   Min.   :0.0000
 10002  :  1   Met:206   1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.0000
 10003  :  1             Median :0.000   Median :1.0000   Median :0.0000
 10004  :  1             Mean   :0.284   Mean   :0.6917   Mean   :0.4927
 10005  :  1             3rd Qu.:1.000   3rd Qu.:1.0000   3rd Qu.:1.0000
 10006  :  1             Max.   :1.000   Max.   :1.0000   Max.   :1.0000
 (Other):406
   fimtavg.1a       fimtavg.7a
 Min.   :0.4850   Min.   :0.5200
 1st Qu.:0.6950   1st Qu.:0.6850
 Median :0.7850   Median :0.7700
 Mean   :0.7936   Mean   :0.7849
 3rd Qu.:0.8712   3rd Qu.:0.8700
 Max.   :1.2800   Max.   :1.1850
                  NA's   :41
```

```
> cwimt <- wimt[complete.cases(wimt),]
> dim( cwimt )

[1] 371   7

> length( intersect( compl, cwimt$subjid ) )

[1] 371

> cc <- lm( fimtavg.7a ~ fimtavg.1a + met + o.65 + pre.ins + SDC,
+           data=cwimt )
> summary( cc )

Call:
lm(formula = fimtavg.7a ~ fimtavg.1a + met + o.65 + pre.ins +
    SDC, data = cwimt)

Residuals:
      Min        1Q    Median        3Q       Max
-0.246178 -0.039118  0.000996  0.042370  0.218314

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.111079   0.023192   4.790 2.43e-06
fimtavg.1a   0.839373   0.027631  30.378  < 2e-16
metMet       0.010418   0.007252   1.437   0.1517
o.65         0.020962   0.008412   2.492   0.0132
pre.ins      0.005352   0.008056   0.664   0.5069
SDC         -0.011232   0.007432  -1.511   0.1316

Residual standard error: 0.06966 on 365 degrees of freedom
Multiple R-squared:  0.7439,        Adjusted R-squared:  0.7404
F-statistic:   212 on 5 and 365 DF,  p-value: < 2.2e-16
```

We see that the coefficient of the baseline measuremeent is pretty close to the derived ρ, and the other parameters are pretty close too:

```
> c( coef(cc)["fimtavg.1a"], rho )

fimtavg.1a
 0.8393728  0.8466723

> c( summary(cc)$sigma^2, (sig2+tau2)*(1-rho^2) )

[1] 0.004852344 0.004875144

> c( coef(cc)["metMet"], (1-rho)*delta+gamma )

    metMet     metMet
0.01041848 0.01090939

> c( coef(cc)["(Intercept)"], (1-rho)*int + beta )

(Intercept) (Intercept)
  0.1110794   0.1071607

> cbind( coef( cc )[4:6], (1-rho)*alpha )

                [,1]          [,2]
o.65     0.020961773  0.0136135639
pre.ins  0.005351978 -0.0009354084
SDC     -0.011232145 -0.0026858391
```

We see that the only possible exception is the coefficients of the stratification parameters.

2.3.2 Using the randomization assumption

If we rely on the assumption that there is no difference between the groups at baseline, that is that $\delta_m = 0$ we fit the model:

```
> mr <- lmer( fimtavg ~ visit + I((visit=="7a")*(met=="Met")) + o.65 + pre.ins + SDC +(1|subjid),
+             data = cmimt )
> summary( mr )


Linear mixed model fit by REML
Formula: fimtavg ~ visit + I((visit == "7a") * (met == "Met")) + o.65 +      pre.ins + SDC + (1 | su
   Data: cmimt
   AIC   BIC logLik deviance REMLdev
 -1320 -1284  668.2    -1382   -1336
Random effects:
 Groups   Name        Variance  Std.Dev.
 subjid   (Intercept) 0.0145858 0.120771
 Residual             0.0026401 0.051382
Number of obs: 742, groups: subjid, 371

Fixed effects:
                                    Estimate Std. Error t value
(Intercept)                         0.779973   0.013099   59.55
visit7a                            -0.012379   0.005280   -2.34
I((visit == "7a") * (met == "Met")) 0.010914   0.007252    1.50
o.65                                0.088362   0.014618    6.04
pre.ins                            -0.006546   0.014572   -0.45
SDC                                -0.017363   0.013439   -1.29

Correlation of Fixed Effects:
            (Intr) visit7 I="*(=" o.65   pre.ns
visit7a     -0.108
I((=="7*(="  0.007 -0.700
o.65        -0.229  0.005 -0.008
pre.ins     -0.628  0.006 -0.008  -0.086
SDC         -0.338 -0.002  0.003  -0.040 -0.217
```

and the resulting derived parameters are then:

```
> sig2  <- attr(VarCorr(mr),"sc")^2
> tau2  <- as.numeric(VarCorr(mr)$subjid)
> rho   <- tau2/(sig2+tau2)
> int   <- fixef( mr )["(Intercept)"]
> beta  <- fixef( mr )["visit7a"]
> delta <- 0
> gamma <- fixef( mr )[3]
> alpha <- fixef( mr )[4:6]
> cbind( c(rho=rho, int, delta, beta, gamma, alpha) )


                                          [,1]
rho                                  0.846737608
(Intercept)                          0.779972561
                                     0.000000000
visit7a                             -0.012379420
I((visit == "7a") * (met == "Met")) 0.010914100
o.65                                 0.088361776
pre.ins                             -0.006546396
SDC                                 -0.017363258
```

and the resultin comparison with the linear model is then:

```
> c( coef(cc)["fimtavg.1a"], rho )
```

```
fimtavg.1a
 0.8393728  0.8467376


> c( summary(cc)$sigma^2, (sig2+tau2)*(1-rho^2) )


[1] 0.004852344 0.004875518


> c( coef(cc)["metMet"], (1-rho)*delta+gamma )


                            metMet I((visit == "7a") * (met == "Met"))
                        0.01041848                          0.01091410


> c( coef(cc)["(Intercept)"], (1-rho)*int + beta )


(Intercept) (Intercept)
  0.1110794   0.1071610


> cbind( coef( cc )[4:6], (1-rho)*alpha )


                [,1]         [,2]
o.65     0.020961773  0.013542537
pre.ins  0.005351978 -0.001003316
SDC     -0.011232145 -0.002661134
```

Pretty much the same sort of agreement as seen above.

However, it is only in the latter case, assuming no mean difference between the baseline measurements in the two randomization groups that we have the conditional model estimating the same quantity (γ_{7m}) as in the random effects model.

2.3.3 Mean change in each group

If we want to report the individual changes in each group, that is β_7 in the placebo group and $\beta_7 + \gamma_{m7}$ in the metformin group, they are not available for the conditional model. This is because the intercept in the conditional model is $(1 - \rho)\mu + \beta_7$, with three parameters of which only ρ is known as the regression coefficient to y_1. The two others we have no handle on. Hence if we want to report the individual changes in each group, we *must* use a model that explicitly addresses these parameters. The conditional model deems these group changes irrelevant.