Mixed effects linear modelling

This type of modelling was recommended by the director of data science at GSK. A mixed effect linear model allows for random as well as fixed effects on the variable of interest whereas a standard linear model only allows for fixed effects.

In this data set, each arm of the trial is a subset of patients. Not every patient has been in each arm of the trial which means the way the patients have been split up is a random effect. Patients also didn't have their glucose taken on exactly the same day, also a random effect.

The model allows for these random effects and the variance that may be happening because of them (for example, there may be an unusual amount of people with naturally high glucose in one arm of the trial) to be taken into account.

The summaries of the three models are shown below:

```
Linear mixed-effects model fit by REML
 Data: Trial1A
       AIC      BIC    logLik
  16868.24 16931.89 -8424.119

Random effects:
 Formula: ~1 | usubjid
        (Intercept) Residual
StdDev:    1.679846  1.23485

Fixed effects: lbstresn ~ lbdy + actarm + lbdy:actarm
                              Value  Std.Error   DF  t-value p-value
(Intercept)                5.860045 0.11867186 2772 49.38024  0.0000
lbdy                       0.000797 0.00064896 2772  1.22782  0.2196
actarmUMEC 62.5            0.157112 0.15345081 1523  1.02386  0.3061
actarmUMEC/VI 62.5/25     -0.000034 0.15368099 1523 -0.00022  0.9998
actarmVI 25                0.244164 0.15299698 1523  1.59588  0.1107
lbdy:actarmUMEC 62.5       0.000109 0.00083203 2772  0.13149  0.8954
lbdy:actarmUMEC/VI 62.5/25 -0.000019 0.00082600 2772 -0.02277  0.9818
lbdy:actarmVI 25           -0.000209 0.00082991 2772 -0.25189  0.8011
 Correlation:
                           (Intr) lbdy   aUMEC62 aUMEC/6 acVI25 l:UMEC62 l:UMEC/6
lbdy                       -0.369
actarmUMEC 62.5            -0.773  0.285
actarmUMEC/VI 62.5/25      -0.772  0.285  0.597
actarmVI 25                -0.776  0.286  0.600   0.599
lbdy:actarmUMEC 62.5        0.288 -0.780 -0.369  -0.222  -0.223
lbdy:actarmUMEC/VI 62.5/25  0.290 -0.786 -0.224  -0.371  -0.225  0.613
lbdy:actarmVI 25            0.289 -0.782 -0.223  -0.223  -0.370  0.610    0.614

Standardized Within-Group Residuals:
        Min          Q1         Med          Q3         Max
-6.26040330 -0.35648115 -0.07097604  0.22262025 14.67328191

Number of Observations: 4303
Number of Groups: 1527
```

```
Linear mixed-effects model fit by REML
 Data: Trial2A
       AIC      BIC    logLik
  26380.86 26450.18 -13180.43

Random effects:
 Formula: ~1 | usubjid
        (Intercept) Residual
StdDev:    1.359359 1.098172

Fixed effects: lbstresn ~ lbdy + actarm + lbdy:actarm
                            Value  Std.Error   DF  t-value p-value
(Intercept)              5.594588 0.07534142 5948 74.25647  0.0000
lbdy                     0.000622 0.00019083 5948  3.25802  0.0011
actarmFF/VI 200/25       0.060449 0.10596900 1620  0.57044  0.5685
actarmFF/VI 50/25        0.156012 0.10581065 1620  1.47444  0.1406
actarmVI 25              0.096673 0.10597413 1620  0.91223  0.3618
lbdy:actarmFF/VI 200/25 -0.000627 0.00026713 5948 -2.34896  0.0189
lbdy:actarmFF/VI 50/25  -0.000626 0.00026730 5948 -2.34044  0.0193
lbdy:actarmVI 25        -0.000412 0.00027036 5948 -1.52550  0.1272
 Correlation:
                        (Intr) lbdy   aFF/V2 aFF/V5 acVI25 l:FF/2 l:FF/5
lbdy                    -0.254
actarmFF/VI 200/25      -0.711  0.180
actarmFF/VI 50/25       -0.712  0.181  0.506
actarmVI 25             -0.711  0.180  0.505  0.506
lbdy:actarmFF/VI 200/25  0.181 -0.714 -0.255 -0.129 -0.129
lbdy:actarmFF/VI 50/25   0.181 -0.714 -0.129 -0.251 -0.129  0.510
lbdy:actarmVI 25         0.179 -0.706 -0.127 -0.127 -0.252  0.504  0.504

Standardized Within-Group Residuals:
       Min         Q1        Med         Q3        Max
-8.0654330 -0.3787987 -0.0735543  0.2569946 11.5295104

Number of Observations: 7576
Number of Groups: 1624
```

```
Linear mixed-effects model fit by REML
 Data: Trial3A
       AIC      BIC   logLik
  26602.79 26672.13 -13291.4

Random effects:
 Formula: ~1 | usubjid
        (Intercept) Residual
StdDev:    1.459194 1.096653

Fixed effects: lbstresn ~ lbdy + actarm + lbdy:actarm
                            Value  Std.Error   DF  t-value p-value
(Intercept)              5.767015 0.07991961 5973 72.16020  0.0000
lbdy                     0.000274 0.00018703 5973  1.46597  0.1427
actarmFF/VI 200/25      -0.049152 0.11277208 1610 -0.43585  0.6630
actarmFF/VI 50/25       -0.083182 0.11253593 1610 -0.73916  0.4599
actarmVI 25             -0.159183 0.11244271 1610 -1.41568  0.1571
lbdy:actarmFF/VI 200/25  0.000170 0.00026478 5973  0.64021  0.5221
lbdy:actarmFF/VI 50/25   0.000233 0.00026365 5973  0.88413  0.3767
lbdy:actarmVI 25         0.000253 0.00026560 5973  0.95245  0.3409
 Correlation:
                        (Intr) lbdy   aFF/V2 aFF/V5 acVI25 l:FF/2 l:FF/5
lbdy                    -0.244
actarmFF/VI 200/25      -0.709  0.173
actarmFF/VI 50/25       -0.710  0.173  0.503
actarmVI 25             -0.711  0.173  0.504  0.505
lbdy:actarmFF/VI 200/25  0.172 -0.706 -0.242 -0.122 -0.122
lbdy:actarmFF/VI 50/25   0.173 -0.709 -0.123 -0.244 -0.123  0.501
lbdy:actarmVI 25         0.172 -0.704 -0.122 -0.122 -0.241  0.497  0.500

Standardized Within-Group Residuals:
        Min          Q1         Med          Q3         Max
-7.01435721 -0.37454169 -0.06663456  0.25715937 11.90392058

Number of Observations: 7591
Number of Groups: 1614
```

It's clear from the intercepts and also the p values that there is no difference in glucose through the trials.

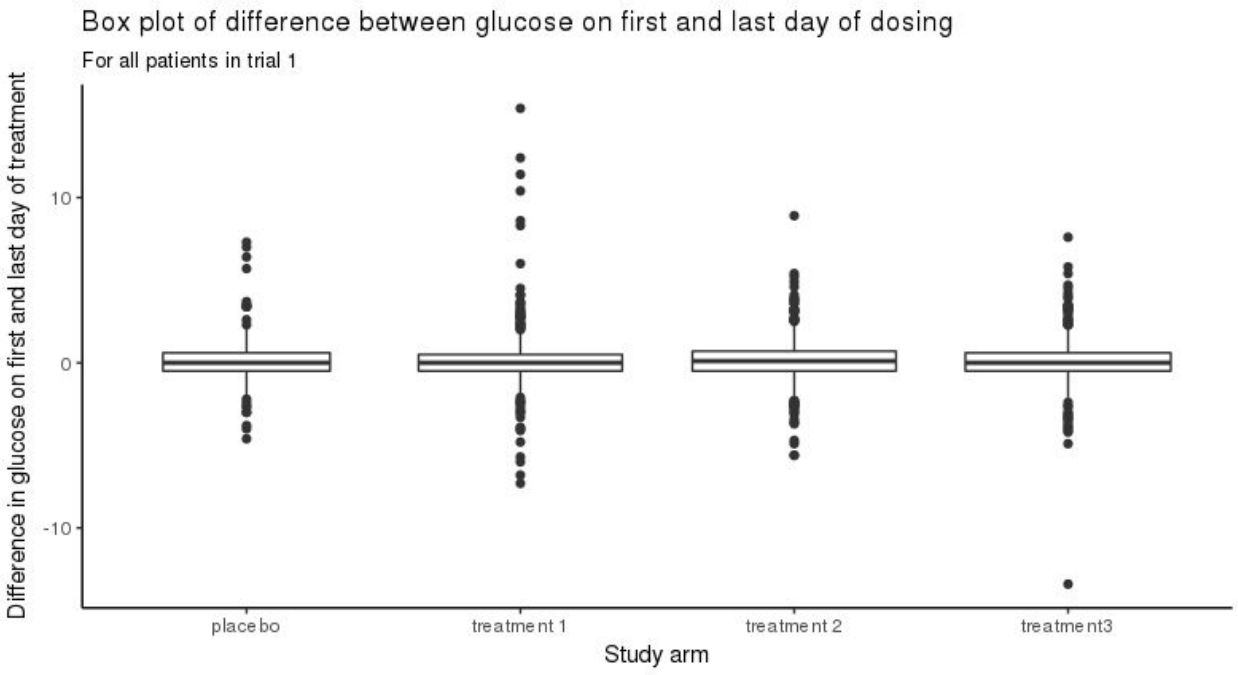
Detailed statistical examination of one trial - further use of statistics

As well as using the model above, GSK asked for one trial to be examined statistically, looking at the absolute differences in glucose values for patients on the first and last days of dosing. This was why the dm data was added to the dataset. The first trial was chosen.

The first dose was identified by pulling the glucose result for the lbdy that was closest to zero. The mean glucose was assigned to any patients where there were two tests done on the same lbdy.

The last dose was identified by subtracting the first dose date (rfxendtc) from the last dose date (rfxstdtc) to calculate which lbdy the last dose was given. Then the glucose result for that lbdy was pulled out.

The differences between these first and last glucose measures were taken for all patients and examined:



This box plot shows that the majority of the differences were near zero.

A further numerical analysis of the differences in glucose on first and last day of dosing for all arms of the trial was performed as well:

actarm	No of patients	Mean of differences	SD of differences	Median of differences
PLACEBO	279	0.08956833	1.347520	0.0000000
UMEC 62.5	416	0.17142856	1.947642	0.0000000
UMEC/VI 62.5/25	413	0.15496368	1.448704	0.0999999
VI 25	421	0.06380953	1.530213	0.0000000

This statistical analysis also shows that the differences were mostly near to zero, showing that nothing interesting is happening to glucose in any arm of the trial.