Interindividual differences in response to treatment: fact, fiction and erroneous analyses

Eamonn O'Brien

20 August, 2016

1 Introduction

An example is presented demonstrating why separately analysing treatment and control arms from a randomised controlled trial (RCT) and testing for an association of a covariate (biomarker) with observed 'responders' and 'non responders' in each arm is erroneous, wasteful and a seriously misleading analysis. This analysis scenario is never advisable and demonstrates a misunderstanding of the basics of randomised clinical trials.

Analysts working in the personalised medicine field need to be knowledgeable of variance components analysis, understand the counterfactual premise that underpins RCTs, that is, what would have happened to the same patients in the treatment arm had they been in the control arm and be knowledgeable of appropriate statistical analyses for RCT data. This basic counterfactual tenet indicates that responders and non responders cannot be identified by examining trial arms separately.

It is advisable to first estimate the true individual response using parallel information from both arms, and if present, judge if it is clinically relevant. If so, covariates that potentially modify or mediate response can be included in the statistical model of the two trial arms.

2 Simulate a RCT

Simulate a randomised controlled trial with a baseline and follow up measurement and a constant treatment effect. A constant treatment effect means **everybody** in the treatment arm responds by a **constant** amount. **There are NO interindividual differences!** In the control arm the true baseline and true follow up are the same.

```
n <- 5000
noise <- 100           # add noise (within person var & meas. error) to the baseline & foll. up
beta.treatment <- -250 # all trt'd subjects exp same trt effect, so no resp - non responders!!
# beta.treatment <- runif(n,-20,-5)  # subjects vary in response to treatment
clin.rel.diff <- 200

pop_mu <- 0            # population mean
pop_sd <- 200          # between person SD
ur.eligible <- -1000   # eligibility criteria for trial

y.0true <- rnorm(n, pop_mu, pop_sd)                  # true baseline
y.0observed <- y.0true + rnorm(n, 0, 1*noise)        # observed baseline
eligible <- ifelse(y.0observed > ur.eligible, 1, 0)  # Recruit all n as ur.eligible tiny
treat <- 1*(runif(n)<.5)                             # random treatment allocation
y.1true <- y.0true + (treat*beta.treatment)          # true follow up, treated only respond
y.1observed <- y.1true + rnorm(n, 0, 1*noise)        # observed follow up, noise added
delta.observed <- y.1observed - y.0observed

d <- data.frame(y.0true, y.0observed, eligible, treat , beta.treatment,
                y.1true, y.1observed, delta.observed)

# prob that a member of pop observed baseline is eligible
# pnorm(ur.eligible, mean= pop_mu, sd=sqrt(pop_sd^2 + noise^2))
# 1- pnorm( (pop_mu - ur.eligible) / sqrt(pop_sd^2+noise^2) )  # z score calc.

trial <- d[d$eligible==1,]    # select the trial subjects
```

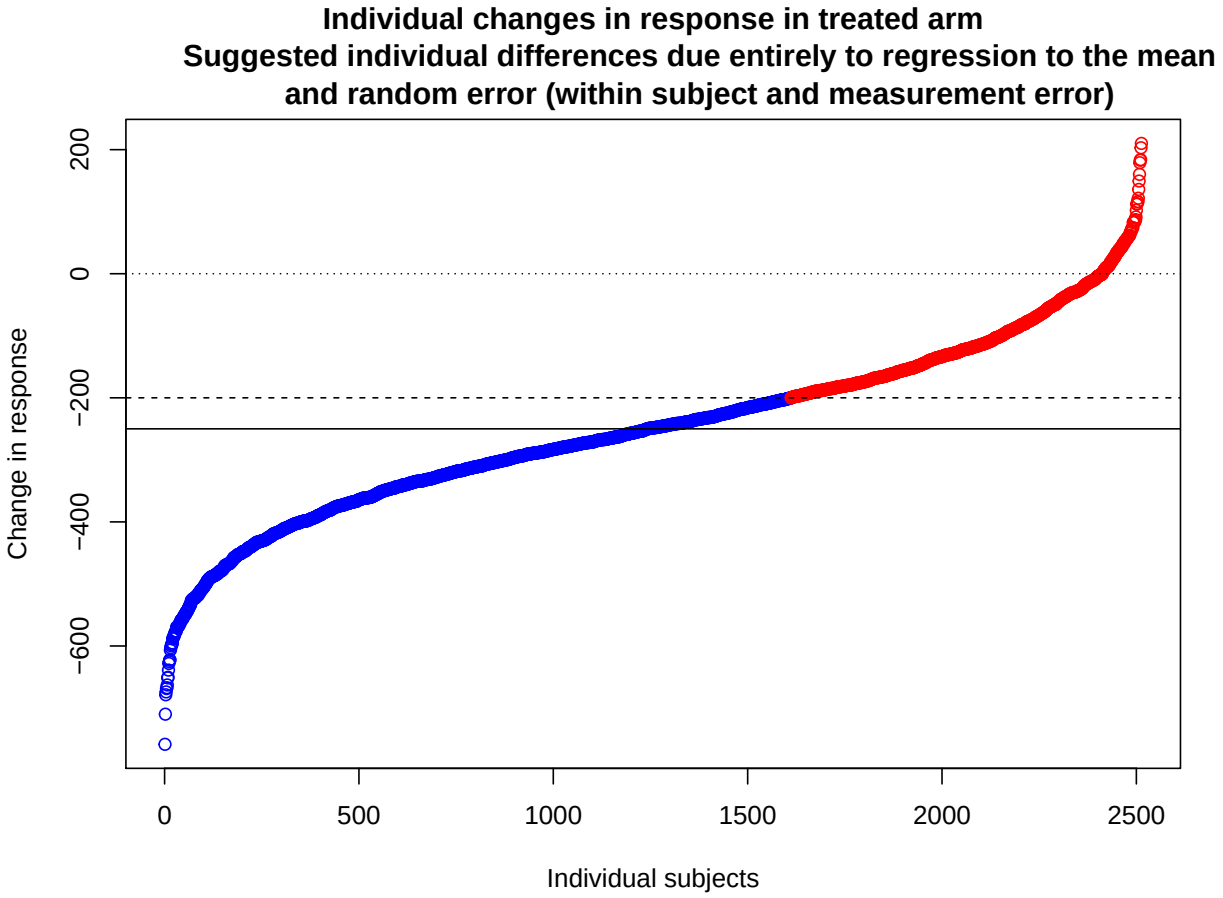
3 First rows of trial data

'y.0true' is the true baseline for each subject. Yet it is not observed, 'y.0observed' is that which is observed and includes measurement error. 'y.1true' is 'y.0true'- 'beta.treatment', the treatment effect , for treated subjects only, otherwise equal to the baseline. But again this is not observed as it is measured with error and the estimate is recorded in variable 'y.1observed'.

y.0true	y.0observed	treat	beta.treatment	y.1true	y.1observed	delta.observed
242.78	297.41	1	-250	-7.22	-102.21	-399.62
-376.58	-405.81	0	-250	-376.58	-431.88	-26.06
-54.79	22.84	1	-250	-304.79	-257.88	-280.72
200.57	307.47	0	-250	200.57	458.68	151.21
125.52	67.51	1	-250	-124.48	-68.11	-135.62
-188.88	3.22	0	-250	-188.88	-122.90	-126.11
120.26	268.11	1	-250	-129.74	-66.51	-334.63
-48.49	-273.99	0	-250	-48.49	-140.53	133.45
-337.31	-181.27	1	-250	-587.31	-485.80	-304.53
-143.42	-262.07	0	-250	-143.42	-133.98	128.09
-233.26	-52.74	1	-250	-483.26	-452.02	-399.29
90.52	99.20	0	-250	90.52	176.74	77.54
407.52	433.77	0	-250	407.52	429.13	-4.64
194.98	259.15	0	-250	194.98	198.46	-60.69
-281.82	-333.91	0	-250	-281.82	-276.42	57.49
-231.65	-369.79	1	-250	-481.65	-540.12	-170.33
190.95	262.64	0	-250	190.95	308.13	45.49
173.81	235.08	1	-250	-76.19	-18.84	-253.93
203.52	296.22	1	-250	-46.48	-144.68	-440.91
-212.60	-371.30	0	-250	-212.60	-150.15	221.15

4 Focus on the intervention arm only - not recommended!

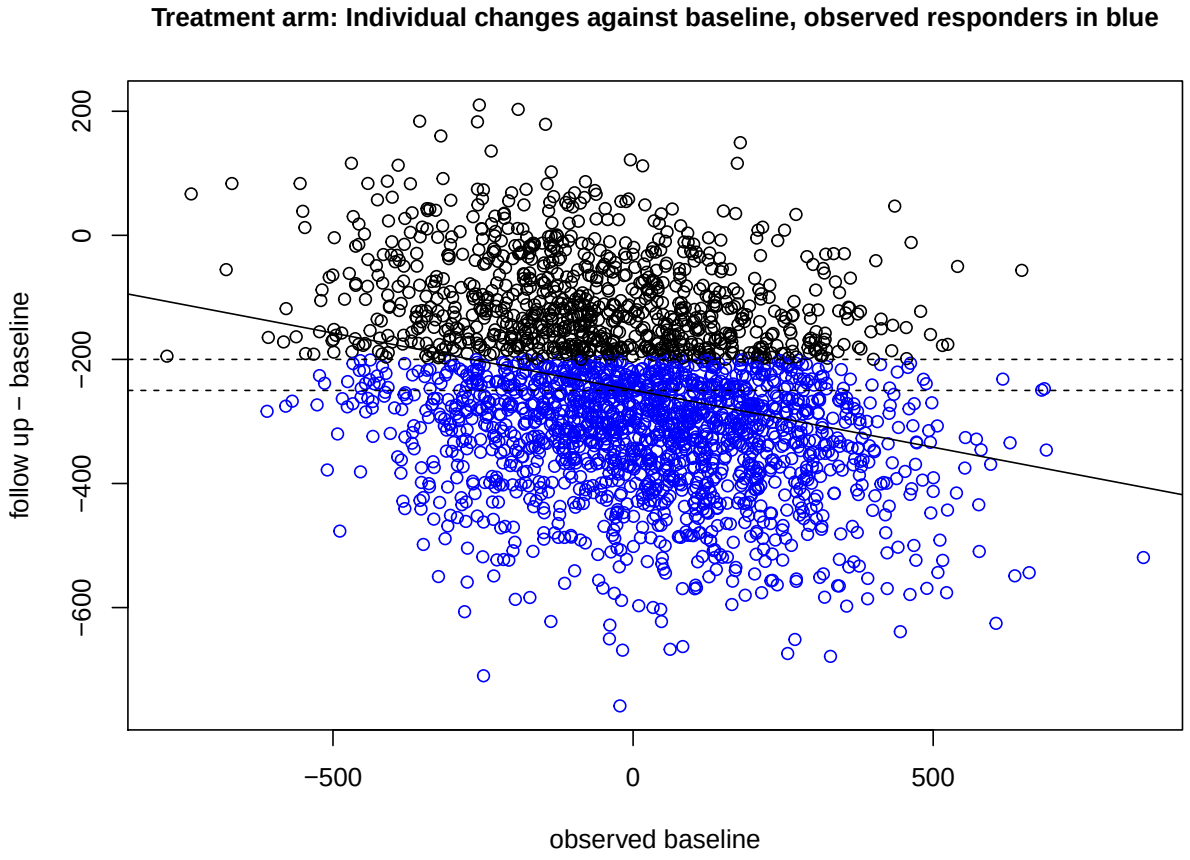
The subjects in blue were observed to respond only, those in red observed not to respond.



No of red points 902 in treatment arm. So 0.3589335 is the proportion that fail to show clinically relevant difference. In the control arm (similar plot to treatment arm plot not shown) 214 observed responders. So 0.0860474 is the proportion that show clinically relevant difference in the control arm.

5 Treatment arm only

Observed responders in blue. But **EVERYBODY** responded to the drug **EQUALLY** ! Apparent individual difference is due **ENTIRELY** to random within subject error, measurement error and regression to the mean.

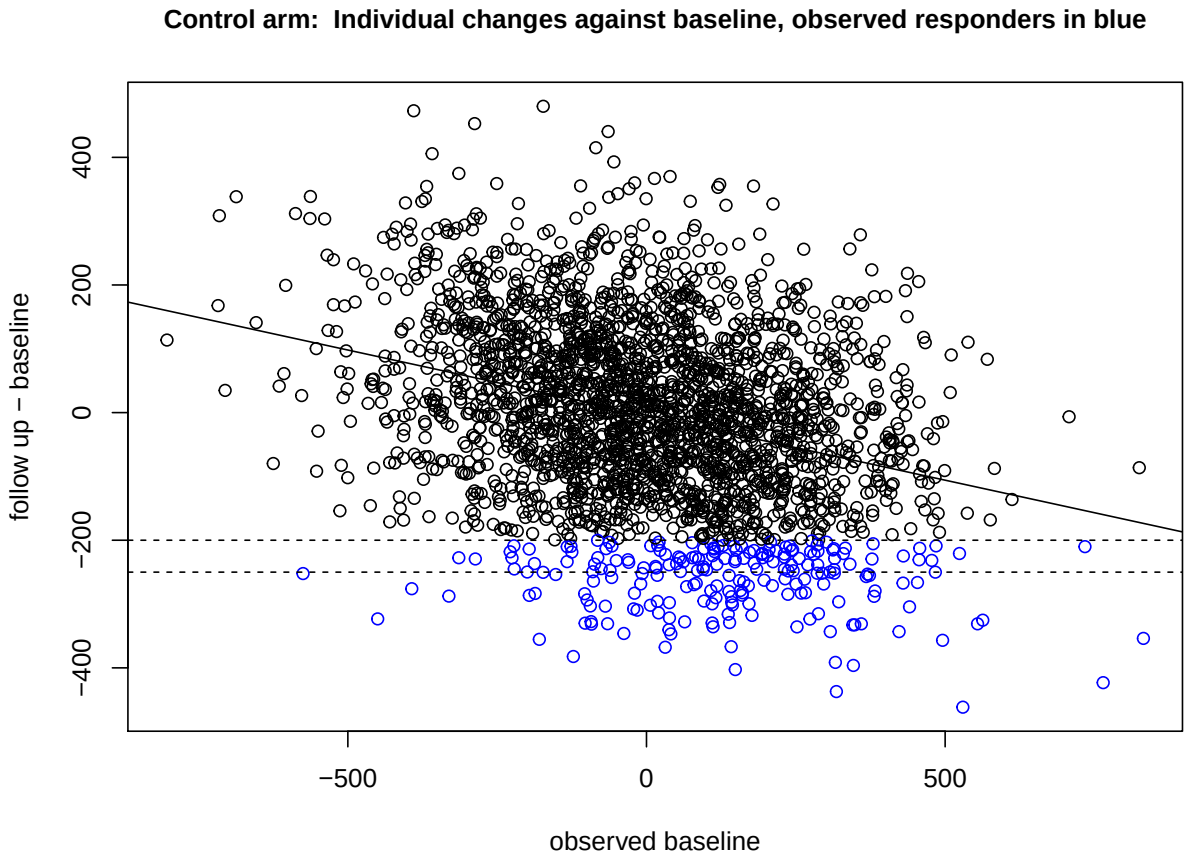



```
    Pearson's product-moment correlation

data:  diff and y.0observed
t = -15.274, df = 2511, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3269461 -0.2553828
sample estimates:
       cor
-0.2915724
```

6 Control arm only

Observed responders in blue. But in truth **NO ONE** responded, apparent individual difference is due **ENTIRELY** to random within subject error, measurement error and regression to the mean.




```
    Pearson's product-moment correlation

data:  diff and y.0observed
t = -16.582, df = 2485, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3505910 -0.2798016
sample estimates:
       cor
-0.3156354
```

7 Analyse the trial correctly. Estimate the treatment effect adjusting for baseline

```
Call:
lm(formula = y.1observed ~ y.0observed + treat, data = trial)

Residuals:
    Min      1Q  Median      3Q     Max
-513.14  -89.76    0.48   89.31  450.47

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.957e+00  2.715e+00  -1.458    0.145
y.0observed  8.062e-01  8.609e-03  93.654   <2e-16 ***
treat       -2.456e+02  3.829e+00 -64.145   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 135.4 on 4997 degrees of freedom
Multiple R-squared:  0.7237,    Adjusted R-squared:  0.7236
F-statistic:  6544 on 2 and 4997 DF,  p-value: < 2.2e-16


                   2.5 %       97.5 %
(Intercept)   -9.2789566    1.3647217
y.0observed    0.7893528    0.8231059
treat       -253.1493166 -238.1344275
```

8 Look before leaping

Calculate the difference in SDs of the changes between the intervention and control arms, confidence interval for the sd for changes in each arm

```
      alpha <- 0.05

      x <- trial[trial$treat %in% 0,"delta.observed"]

      lstar <- qchisq(alpha/2, df=  length(x)-1)
      rstar <- qchisq(1-alpha/2, df=  length(x)-1)

      up <- sqrt((length(x)-1)*var(x)/(lstar))
      lo <- sqrt((length(x)-1)*var(x)/(rstar))
      pe <- sqrt(var(x))

      # ctrl arm estimate with 95% CI
      print(c(pe, lo, up), digits=3)
```

```
[1] 142 138 146
```

```
      x1 <- trial[trial$treat %in% 1,"delta.observed"]

      lstar <- qchisq(alpha/2, df=  length(x1)-1)
      rstar <- qchisq(1-alpha/2, df=  length(x1)-1)

      up <- sqrt((length(x1)-1)*var(x1)/(lstar))
      lo <- sqrt((length(x1)-1)*var(x1)/(rstar))
      pe <- sqrt(var(x1))

      # trt arm estimate with 95% CI
      print(c(pe, lo, up), digits=3)
```

```
[1] 142 138 146
```

9 Typical true interindividual variation in response. Adjust for the influence of biological variation and measurement error (removal of noise).

The linear mixed model p-value provides evidence the SD for true interindividual variation is consistent with zero, as it should be, given that the true magnitude of response in the simulation is constant for all subjects randomised to the treated arm. This result provides information that true individual response differences are negligible and analysis of interindividual response is unwarranted.

```
      # True individual response to the intervention
      sqrt(sd(x1)^2-sd(x)^2)  # can be -ve if more var in control group
```

```
[1] NaN
```

```
      # LMM approach
      m1 <- lme(delta.observed~ treat + y.0observed,
          random=~1|treat , data=trial, method="REML",
          weights = varIdent(form = ~1 | treat))

      m0 <-lme(delta.observed~ treat + y.0observed,
          random=~1|treat , data=trial, method="REML")

      print(m1)
```

```
Linear mixed-effects model fit by REML
  Data: trial
  Log-restricted-likelihood: -31633.29
  Fixed: delta.observed ~ treat + y.0observed
 (Intercept)        treat  y.0observed
  -3.9569576 -245.6422839   -0.1938266

Random effects:
 Formula: ~1 | treat
        (Intercept) Residual
StdDev:    7.199333 134.9875

Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | treat
 Parameter estimates:
       0        1
1.000000 1.005658
Number of Observations: 5000
Number of Groups: 2
```

```
      anova(m1,m0) # are the trt ctr interindividual variation in response different?
```

```
   Model df      AIC      BIC    logLik   Test    L.Ratio p-value
m1     1  6 63278.58 63317.68 -31633.29
m0     2  5 63276.66 63309.24 -31633.33 1 vs 2 0.07960173  0.7778
```

```
      c.grp <- m1$sigma
      t.grp <- coef(m1$modelStruct$varStruct, uncons = FALSE)[[1]]*m1$sigma

      # true individual response to the intervention estimate
      sqrt(t.grp^2 - c.grp^2)
```

```
[1] 14.37981
```

```
      # truth
      sd(beta.treatment )
```

```
[1] NA
```

10 References

Based on 'Individual response to treatment: is it a valid assumption?' S Senn BMJ Vol 329 2004

11 Computing Environment

```
R version 3.2.2 (2015-08-14)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 8 x64 (build 9200)

locale:
[1] LC_COLLATE=English_United Kingdom.1252
[2] LC_CTYPE=English_United Kingdom.1252
[3] LC_MONETARY=English_United Kingdom.1252
[4] LC_NUMERIC=C
[5] LC_TIME=English_United Kingdom.1252

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods
[7] base

other attached packages:
[1] nlme_3.1-128 knitr_1.13

loaded via a namespace (and not attached):
 [1] magrittr_1.5    formatR_1.4     tools_3.2.2     htmltools_0.3.5
 [5] yaml_2.1.13     Rcpp_0.12.6     stringi_1.1.1   rmarkdown_1.0
 [9] highr_0.6       grid_3.2.2      stringr_1.0.0   digest_0.6.10
[13] lattice_0.20-33 evaluate_0.9
```

This took 1.69 seconds to execute.