Analysis of base-line follow-up data

Bendix Carstensen Steno Diabetes Center
Gentofte, Denmark
http://BendixCarstensen.com

SDC

August 2015

http://BendixCarstensen.com/SDC/LEAD/Vickers

Outline

- Observations / measurements for each individual, i:

Outline

- Observations / measurements for each individual, i:
  - Baseline y_{0i}

Outline

- Observations / measurements for each individual, i:
  - Baseline y_{0i}
  - Follow-up y_{1i}

Outline

- Observations / measurements for each individual, i:
  - Baseline y_{0i}
  - Follow-up y_{1i}
  - treatment group

Outline

- Observations / measurements for each individual, i:
  - Baseline y_{0i}
  - Follow-up y_{1i}
  - treatment group
  - covariates

Outline

- Observations / measurements for each individual, i:
  - Baseline y_{0i}
  - Follow-up y_{1i}
  - treatment group
  - covariates
- Topic of interest:

Outline

- Observations / measurements for each individual, i:
  - Baseline y_{0i}
  - Follow-up y_{1i}
  - treatment group
  - covariates
- Topic of interest:
  - How much is the change from baseline to follow-up

Outline

- Observations / measurements for each individual, i:
  - Baseline y_{0i}
  - Follow-up y_{1i}
  - treatment group
  - covariates
- Topic of interest:
  - How much is the change from baseline to follow-up
  - How much does this depend on treatment / covariates

The change from baseline to follow-up

- Baseline is subject to random error

The change from baseline to follow-up

- Baseline is subject to random error
- If the random error at baseline is large **positive**:

The change from baseline to follow-up

- Baseline is subject to random error
- If the random error at baseline is large **positive**:
    - baseline measurment "artificially" **large**

The change from baseline to follow-up

- Baseline is subject to random error
- If the random error at baseline is large **positive**:
  - baseline measurment “artificially” **large**
  - **change** to follow-up **smaller**

The change from baseline to follow-up

- Baseline is subject to random error
- If the random error at baseline is large **positive**:
  - baseline measurment “artificially” **large**
  - **change** to follow-up **smaller**
- If the random error at baseline is large **negative**:

The change from baseline to follow-up

- Baseline is subject to random error
- If the random error at baseline is large **positive**:
  - baseline measurment "artificially" **large**
  - **change** to follow-up **smaller**
- If the random error at baseline is large **negative**:
  - baseline measurment "artificially" **small**

The change from baseline to follow-up

- Baseline is subject to random error
- If the random error at baseline is large **positive**:
  - baseline measurment "artificially" **large**
  - **change** to follow-up **smaller**
- If the random error at baseline is large **negative**:
  - baseline measurment "artificially" **small**
  - **change** to follow-up **larger**

The change from baseline to follow-up

- Baseline is subject to random error
- If the random error at baseline is large **positive**:
  - baseline measurment “artificially” **large**
  - **change** to follow-up **smaller**
- If the random error at baseline is large **negative**:
  - baseline measurment “artificially” **small**
  - **change** to follow-up **larger**
- $\Rightarrow$ the change depends on the baseline **measurement**.

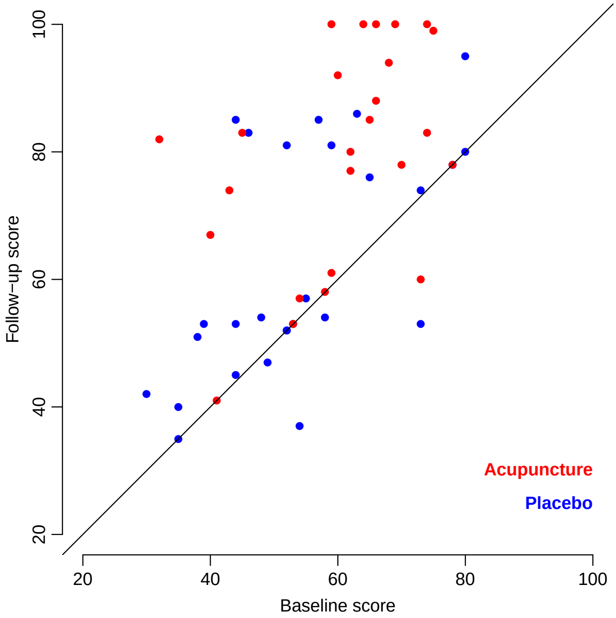
Example from Vickers et al.

```
> library( Epi )
> library( foreign )
> acp <- read.dta( "./data/sportsmen.dta" )[,-4]
> names( acp ) <- c("bl","fu","gr")
> acp$gr <- factor( acp$gr, labels=c("Placebo","Acupuncture") )
> str( acp )
'data.frame': 54 obs. of  3 variables:
 $ bl: num  59 53 46 38 52 63 30 73 44 48 ...
 $ fu: num  81 53 83 51 81 86 42 74 45 54 ...
 $ gr: Factor w/ 2 levels "Placebo","Acupuncture": 1 1 1 1 1 1 1 1 1 1 ...
> head( acp )
  bl fu      gr
1 59 81 Placebo
2 53 53 Placebo
3 46 83 Placebo
4 38 51 Placebo
5 52 81 Placebo
6 63 86 Placebo
```

Example data from Vickers *et al.*:

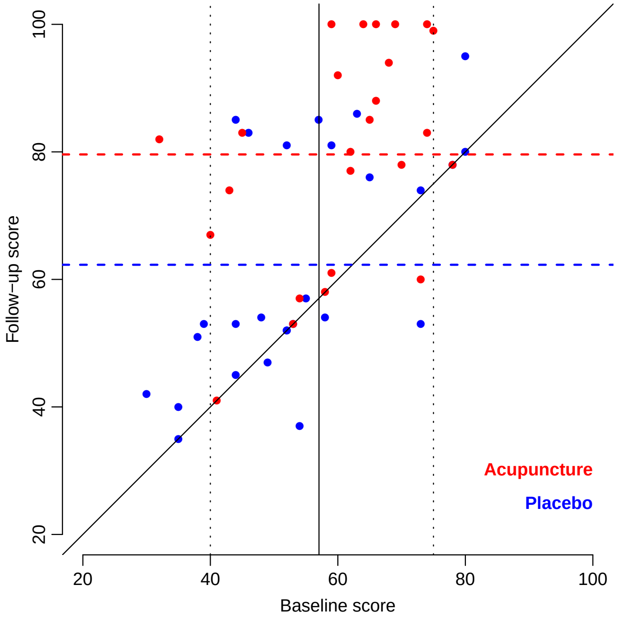
Randomization to acupuncture / placebo

Outcome: Pain/function rating of shoulder pain (0–100).



Follow-up analysis

If the study is randomized, analysis of follow-up is in principle unbiased, because baseline distribution is the same in randomizatin groups.



Analysis of follow-up

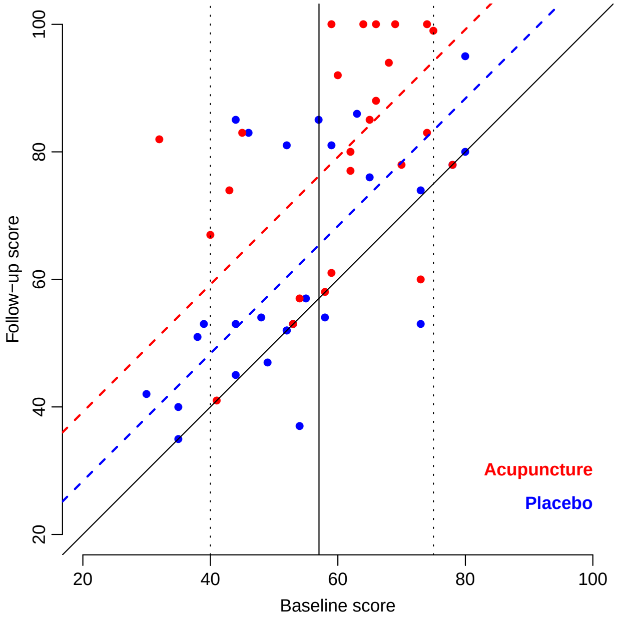
```
> # Follow-up
> fu <- with( acp, tapply( fu   , gr, mean ) )
> c( fu, diff( fu ) )
    Placebo Acupuncture Acupuncture
    62.2963     79.6000     17.3037
> mf <- lm( fu ~ gr, data=acp )
> round( ci.lin( mf ), 3 )
              Estimate StdErr      z P   2.5%  97.5%
(Intercept)     62.296  3.378 18.440 0 55.675 68.918
grAcupuncture   17.304  4.872  3.551 0  7.754 26.853
```

Analysis of change scores

$y_1 - y_0$

If not randomized this is also biased by baseline differences

The change scores are found as the distance to the 45° line.



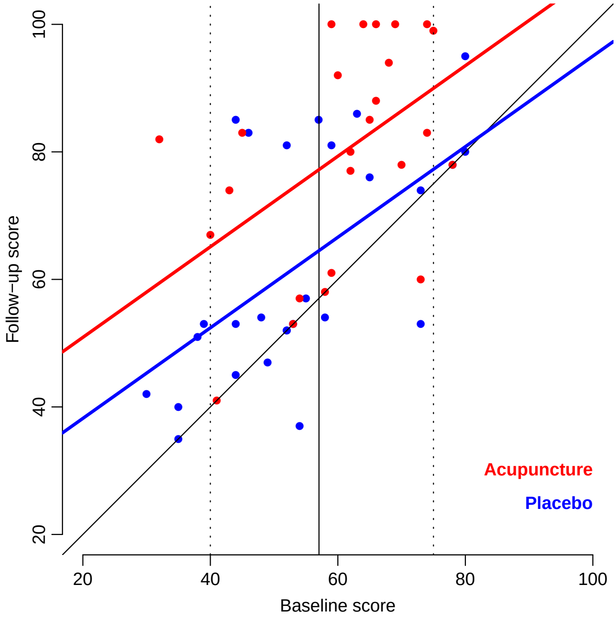
Analysis of change scores

```
> df <- with( acp, tapply( fu-bl, gr, mean ) )
> c( df, diff( df ) )
    Placebo Acupuncture Acupuncture
    8.37037    19.20000    10.82963
> md <- lm( fu-bl ~ gr, data=acp )
> round( ci.lin( md ), 3 )
              Estimate StdErr     z     P  2.5%  97.5%
(Intercept)       8.37  2.948 2.839 0.005 2.592 14.148
grAcupuncture    10.83  4.252 2.547 0.011 2.497 19.163
```

Conditioning on baseline $y_1 | y_0$

Accounts for possible imbalances in baseline distribution.

Separates treatment effect and baseline effect on outcome.



Conditioning on baseline

```
> mc <- lm( fu ~ bl + gr, data=acp )
> round( ci.lin( mc ), 4 )
               Estimate StdErr      z      P   2.5%   97.5%
(Intercept)     23.9973 9.1092 2.6344 0.0084 6.1435 41.8511
bl               0.7102 0.1602 4.4323 0.0000 0.3962  1.0243
grAcupuncture   12.7057 4.2857 2.9647 0.0030 4.3059 21.1056
```

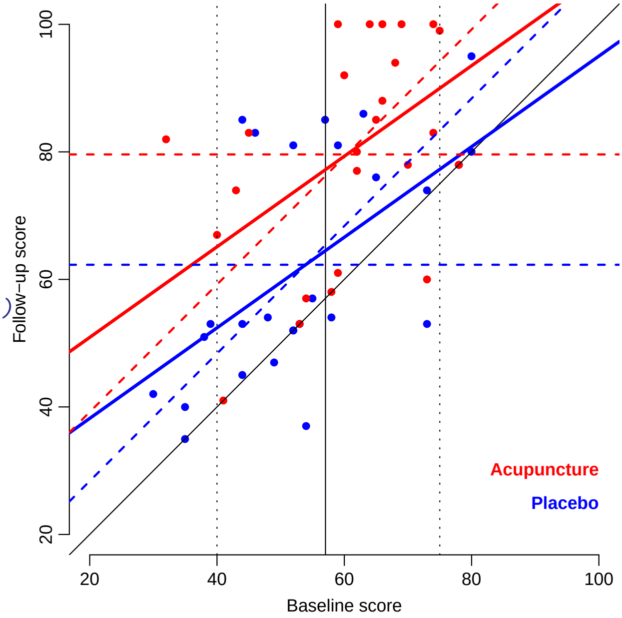
- $y_{i1} = M + By_{i0} + D_g$
- treatment effect (D_g) is 12.7 points:
  - change in placebo:
    $M + (B-1)y_{i0} = 23.997 - 0.290 \times y_{0i}$
  - change in acupuncture:
    $M + (B-1)y_{i0} + D_g = 23.997 - 0.290 \times y_{0i} + 12.706$
- change from baseline to FU depend on baseline
- treatment effect is **difference** in changes

Comparing the three approaches

```
> cmp.cf <- rbind( ci.lin( mf, subset="Acu" ),
+                  ci.lin( md, subset="Acu" ),
+                  ci.lin( mc, subset="Acu" ) )
> rownames( cmp.cf ) <- c("FU","Chg-sc","Cond")
> round( cmp.cf, 4 )
       Estimate StdErr      z      P   2.5%   97.5%
FU      17.3037 4.8723 3.5515 0.0004 7.7542 26.8532
Chg-sc  10.8296 4.2516 2.5472 0.0109 2.4966 19.1627
Cond    12.7057 4.2857 2.9647 0.0030 4.3059 21.1056
```

Comparing the three approaches

```
> round( cmp.cf[,c(1,2,4)], 3 )
       Estimate StdErr     P
FU       17.304  4.872 0.000
Chg-sc   10.830  4.252 0.011
Cond     12.706  4.286 0.003
```



Changes from baseline to FU

```
> ( cf <- coef(mc) )
  (Intercept)              bl grAcupuncture
   23.9973054       0.7102148    12.7057205
> ( mb <- mean( acp$bl ) )
[1] 57.04259
> y0 <- c(40,mb,75)
> p.ch <- cf[1] - (cf[2]-1)*y0
> a.ch <- cf[1] - (cf[2]-1)*y0 + cf[3]
> chg <- cbind( p.ch, a.ch, a.ch-p.ch )
> colnames( chg ) <- c( levels( acp$gr ), "Diff" )
> rownames( chg ) <- round(y0,2)
> round( chg, 2 )
      Placebo Acupuncture  Diff
40      35.59       48.29 12.71
57.04   40.53       53.23 12.71
75      45.73       58.44 12.71
```

Constant treatment effect?

- Depends on baseline & treatment
- ... but treatment **difference** in change does not
- But **does** the treatment effect depend on baseline?

Constant treatment effect?

- Depends on baseline & treatment
- …but treatment **difference** in change does not
- But **does** the treatment effect depend on baseline?

```
> mi <- lm( fu ~ gr + gr:bl, data=acp )
> round( ci.lin( mi ), 4 )
                   Estimate  StdErr      z      P     2.5%   97.5%
(Intercept)         20.3488 11.7437 1.7327 0.0831  -2.6685 43.3661
grAcupuncture       22.1307 19.4070 1.1403 0.2541 -15.9062 60.1677
grPlacebo:bl         0.7779  0.2110 3.6865 0.0002   0.3643  1.1914
grAcupuncture:bl     0.6146  0.2509 2.4498 0.0143   0.1229  1.1063
> anova( mi, mc )
Analysis of Variance Table

Model 1: fu ~ gr + gr:bl
Model 2: fu ~ bl + gr
  Res.Df   RSS Df Sum of Sq      F Pr(>F)
1     48 10942
2     49 10998 -1   -56.565 0.2481 0.6207
```

Constant treatment effect?

- Depends on baseline & treatment
- …but treatment **difference** in change does not
- But **does** the treatment effect depend on baseline?

```
> mi <- lm( fu ~ gr + gr:bl, data=acp )
> round( ci.lin( mi ), 4 )
                   Estimate  StdErr      z      P     2.5%   97.5%
(Intercept)         20.3488 11.7437 1.7327 0.0831  -2.6685 43.3661
grAcupuncture       22.1307 19.4070 1.1403 0.2541 -15.9062 60.1677
grPlacebo:bl         0.7779  0.2110 3.6865 0.0002   0.3643  1.1914
grAcupuncture:bl     0.6146  0.2509 2.4498 0.0143   0.1229  1.1063
> anova( mi, mc )
Analysis of Variance Table

Model 1: fu ~ gr + gr:bl
Model 2: fu ~ bl + gr
  Res.Df   RSS Df Sum of Sq      F Pr(>F)
1     48 10942
2     49 10998 -1   -56.565 0.2481 0.6207
```