

Chapter 4: Worked Example

Jules Lanari-Collard

2024-08-09

Standard Crossover Design

```
data <- read_xlsx("../data/CrossOverData2.xlsx") %>%
  mutate(across(Subject:treat, as_factor)) %>%
  select(-Pre)

data.wide <- data %>%
  pivot_wider(id_cols = c(Subject, Sequence),
              names_from = Period, values_from = Post,
              names_prefix = "Period")

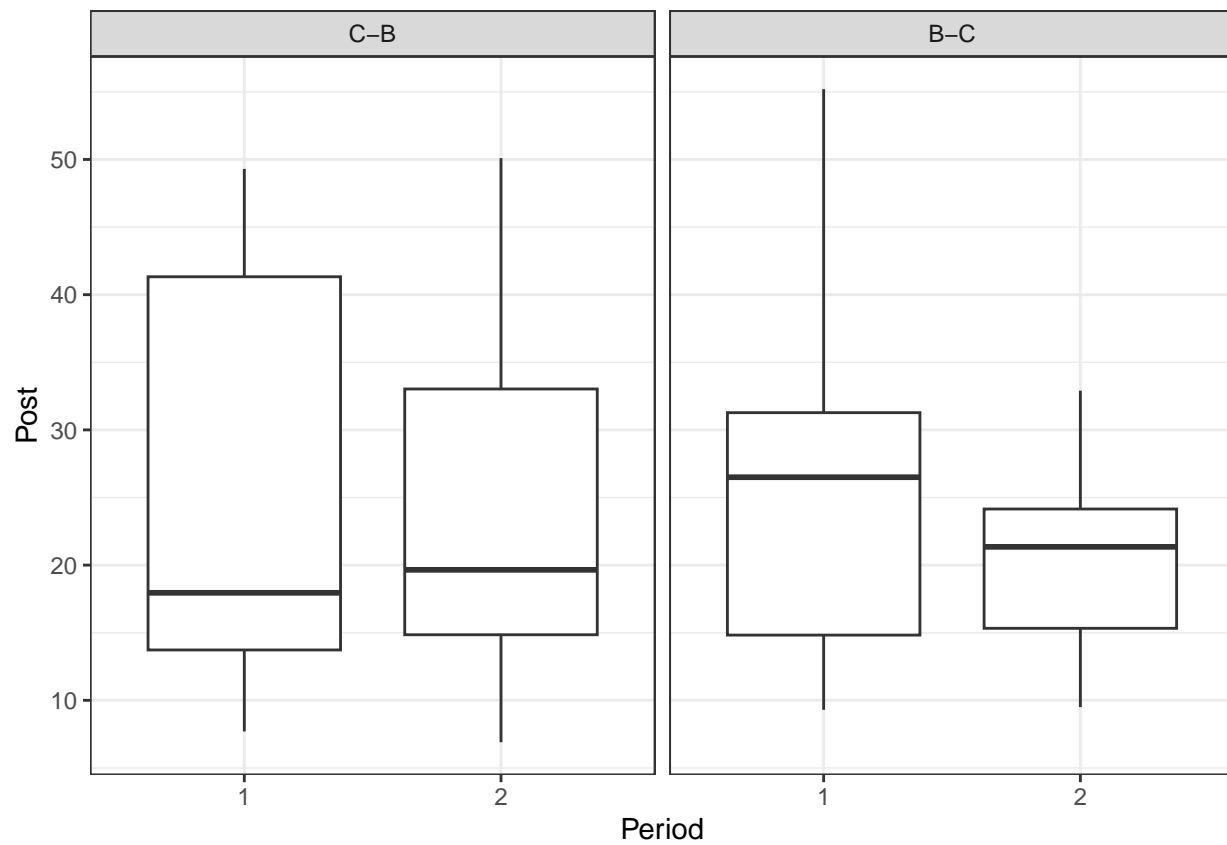
means <- data.wide %>%
  group_by(Sequence) %>%
  summarise(Period1 = mean(Period1),
            Period2 = mean(Period2))

medians <- data.wide %>%
  group_by(Sequence) %>%
  summarise(Period1 = median(Period1),
            Period2 = median(Period2))

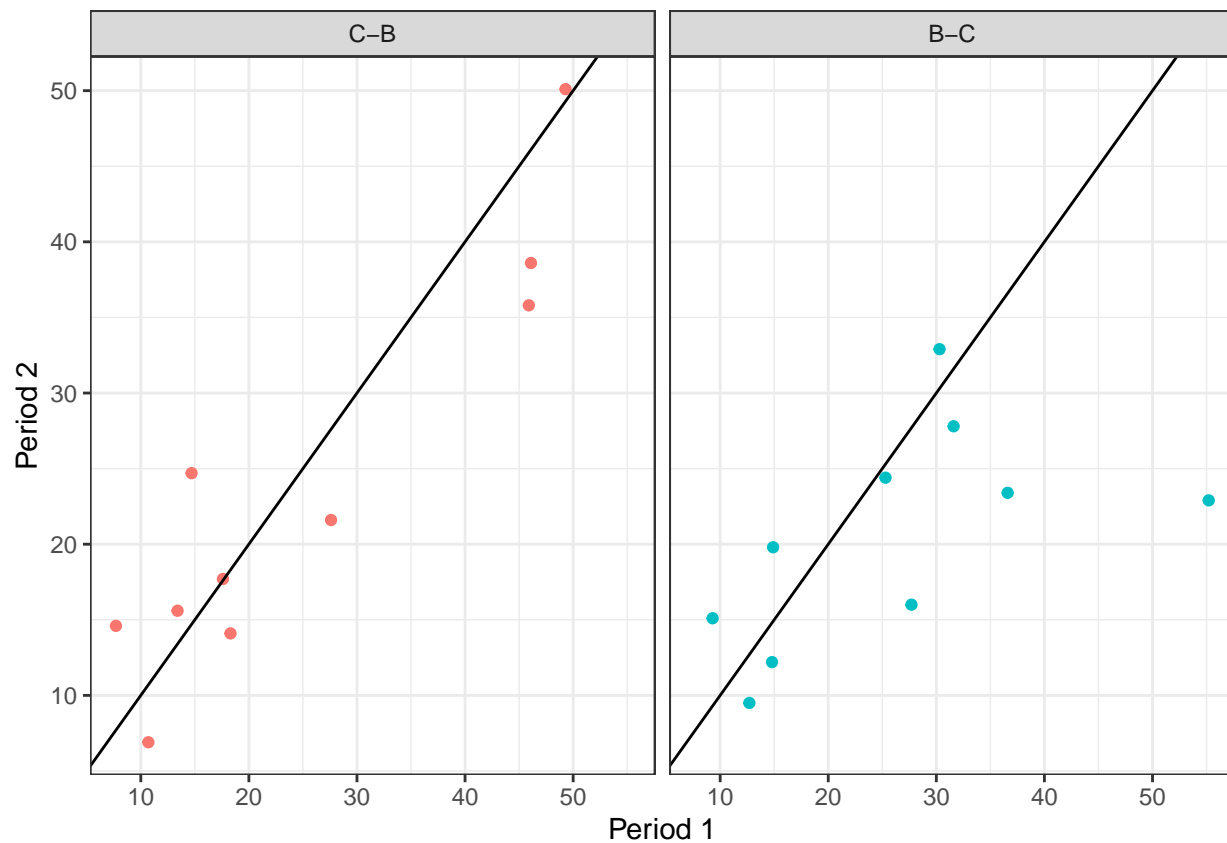
medians.long <- medians %>%
  rename(`1` = "Period1", `2` = "Period2") %>%
  pivot_longer(cols = c(`1`, `2`), names_to = "Period", values_to = "Value",
               names_transform = as_factor)
```

Summarising the Data

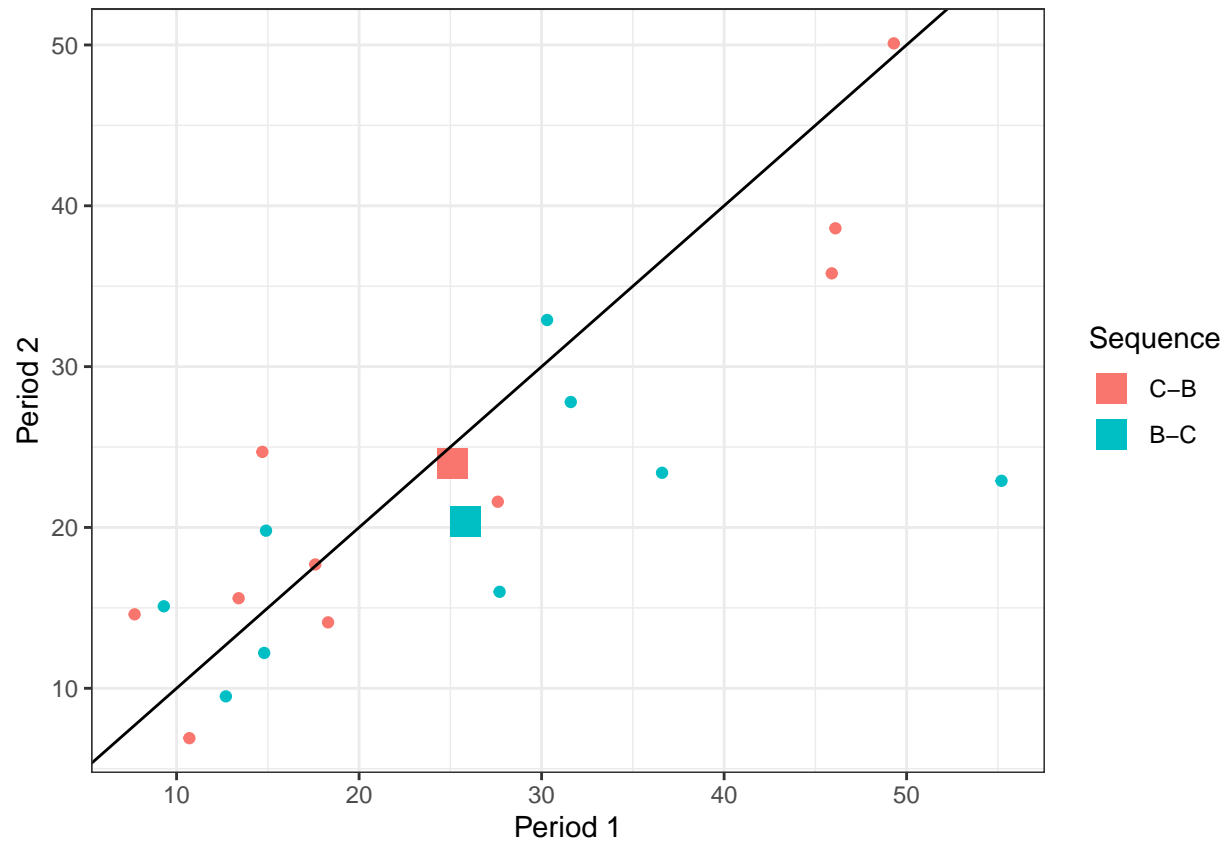
```
ggplot(data = data, mapping = aes(x = Period, y = Post)) +
  facet_wrap(~Sequence) +
  theme_bw() +
  geom_boxplot()
```



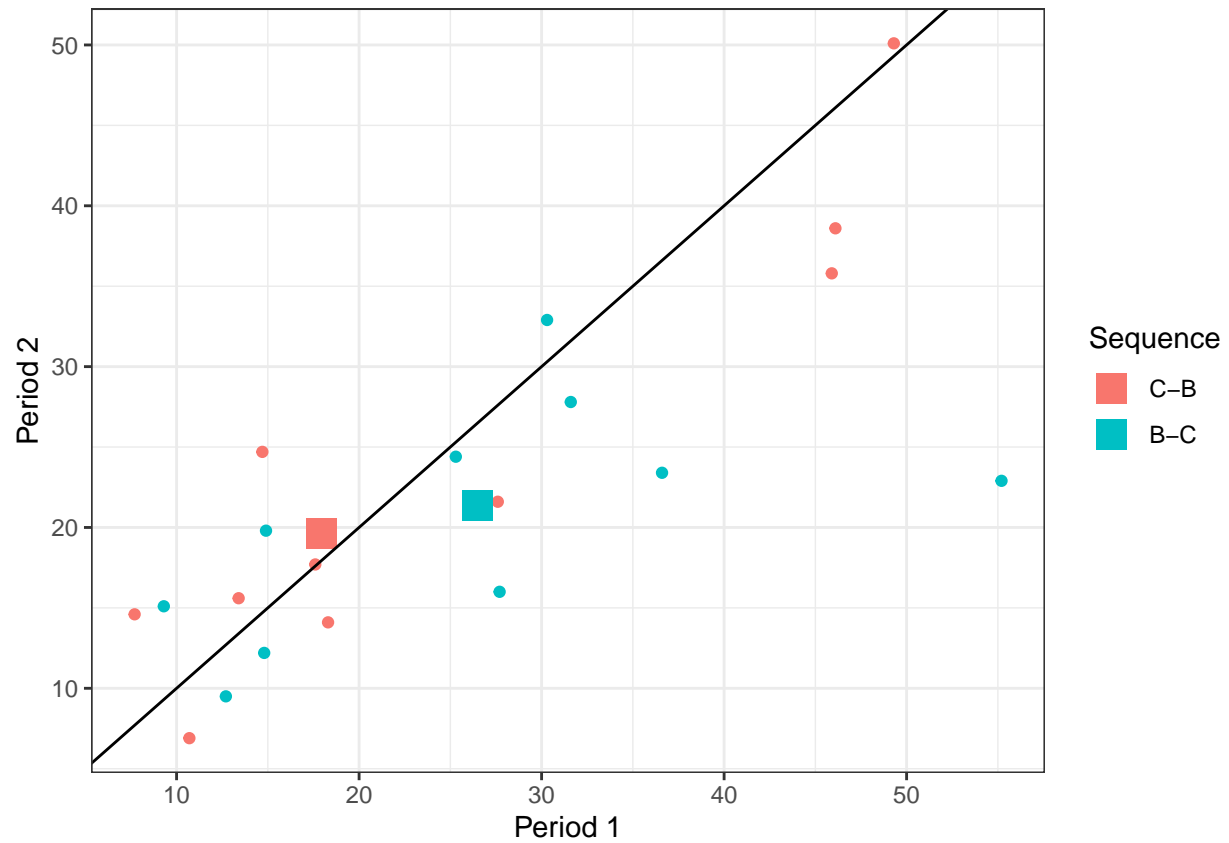
```
ggplot(data = data.wide, mapping = aes(x = Period1, y = Period2, col = Sequence)) +
  facet_wrap(~Sequence) +
  geom_point(show.legend = FALSE) +
  geom_abline(slope=1, intercept=0) +
  labs(x="Period 1", y="Period 2") +
  theme_bw()
```



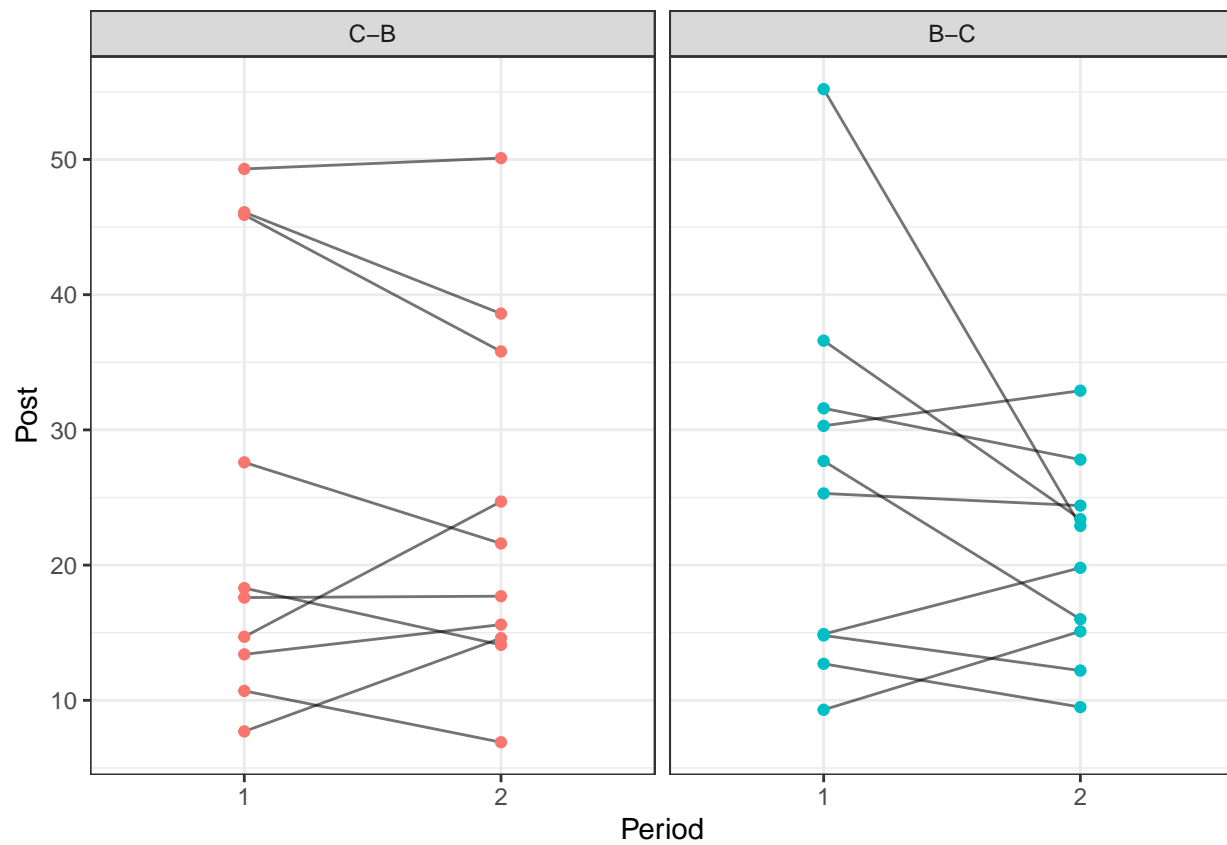
```
ggplot(data = data.wide, mapping = aes(x = Period1, y = Period2, col = Sequence)) +
  geom_point(show.legend = FALSE) +
  geom_point(data = means,
             mapping = aes(x = Period1, y = Period2),
             shape = "square", size = 5) +
  geom_abline(slope=1, intercept=0) +
  labs(x="Period 1", y="Period 2") +
  theme_bw()
```



```
ggplot(data = data.wide, mapping = aes(x = Period1, y = Period2, col = Sequence)) +
  geom_point(show.legend = FALSE) +
  geom_point(data = medians,
            mapping = aes(x = Period1, y = Period2),
            shape = "square", size = 5) +
  geom_abline(slope=1, intercept=0) +
  labs(x="Period 1", y="Period 2") +
  theme_bw()
```



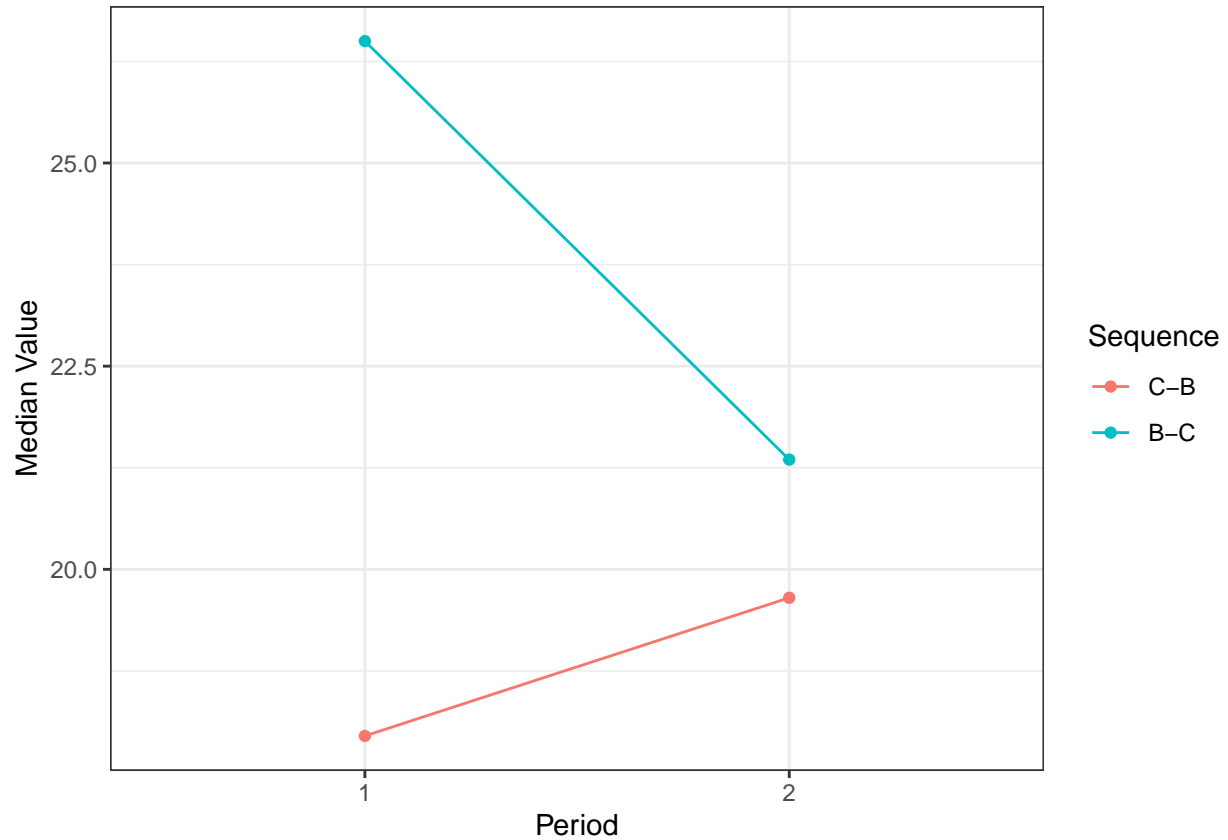
```
ggplot(data = data,  
       mapping = aes(x = Period, y = Post)) +  
  facet_wrap(~Sequence) +  
  geom_line(aes(group = Subject), alpha = 0.55) +  
  geom_point(aes(col = Sequence), show.legend = FALSE) +  
  theme_bw()
```



```
ggplot(data = medians.long, mapping = aes(x = Period, y = Value, colour = Sequence)) +
  geom_line(aes(group = Sequence)) +
  geom_point() +
  labs(y = "Median Value") +
  theme_bw()
```

Table 1: Data with Baseline Measurements

Subject	Sequence	Period	treat	Pre	Post
1	C-B	1	CRICKET	20.3	18.3
1	C-B	2	BEEF	14.6	14.1
2	C-B	1	CRICKET	11.1	14.7
2	C-B	2	BEEF	11.1	24.7
4	C-B	1	CRICKET	38.3	45.9



Modelling for Treatment Effect

Using Baseline Measurements

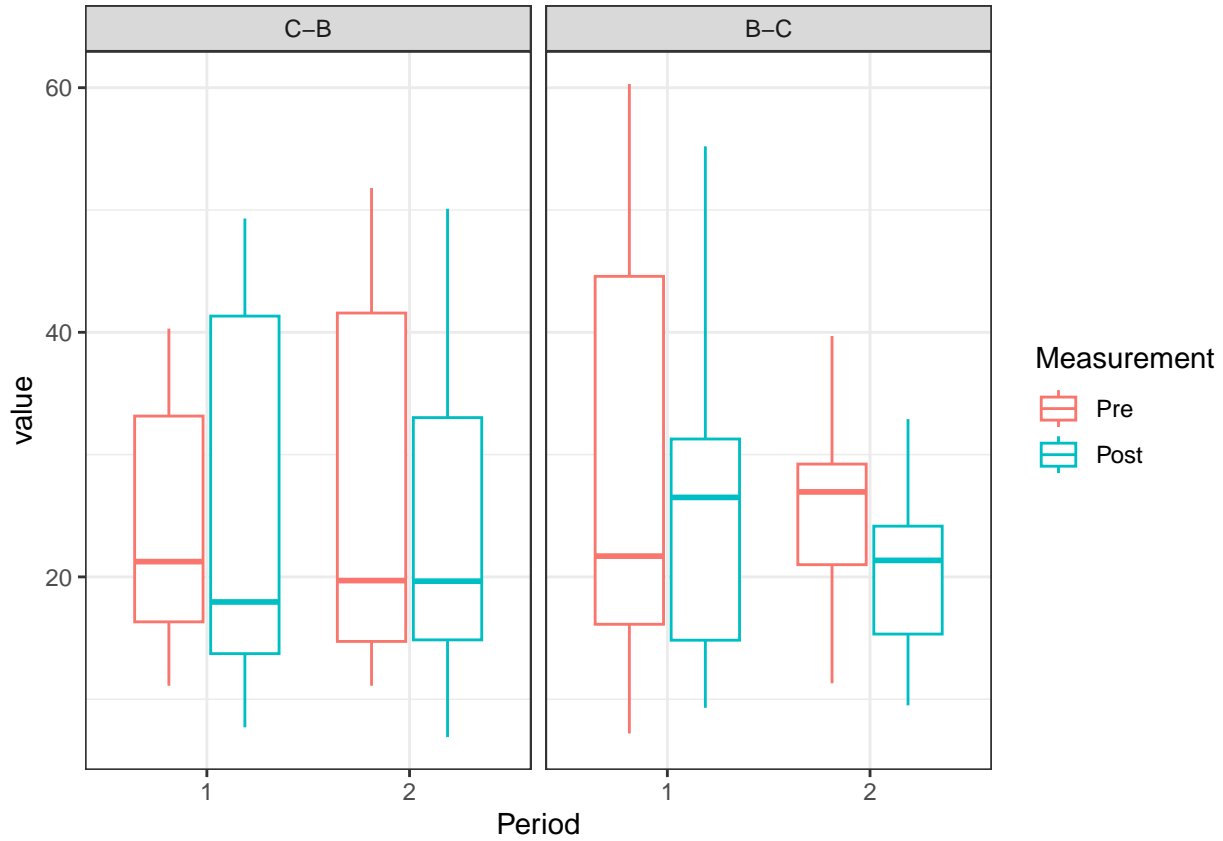
Now we consider the case where we have baseline measurements. A sample of the data is shown in table 1; now there is an extra column with the baseline measurement before each period (Pre).

Our summary table can be expanded to include the baseline measurements, as shown in table 2.

We can also extend the boxplot to incorporate the baseline measurements (see figure ??).

Table 2: Summary Table with Baseline Values (Pre)

Sequence	Subject	Overall				Period 1				Period 2			
		Pre		Post		Pre		Post		Pre		Post	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
C-B	10	25.71	13.53	24.55	14.43	24.03	10.70	25.13	16.07	27.39	16.31	23.97	13.45
B-C	10	27.91	14.39	23.12	11.11	29.73	18.94	25.84	13.83	26.10	8.44	20.40	7.27
Total	20	26.81	13.83	23.84	12.74	26.88	15.25	25.48	14.60	26.75	12.66	22.18	10.68



Modelling with Baseline Measurements

First we construct the ANOVA table, this time including the baseline by period interaction.

```
aov(Post ~ treat + Period * baseline_diff + Sequence + Subject,
     data = data.baselines) %>%
  summary() %>%
  xtable() %>%
  xtable2kable() %>%
  kable_paper()
```

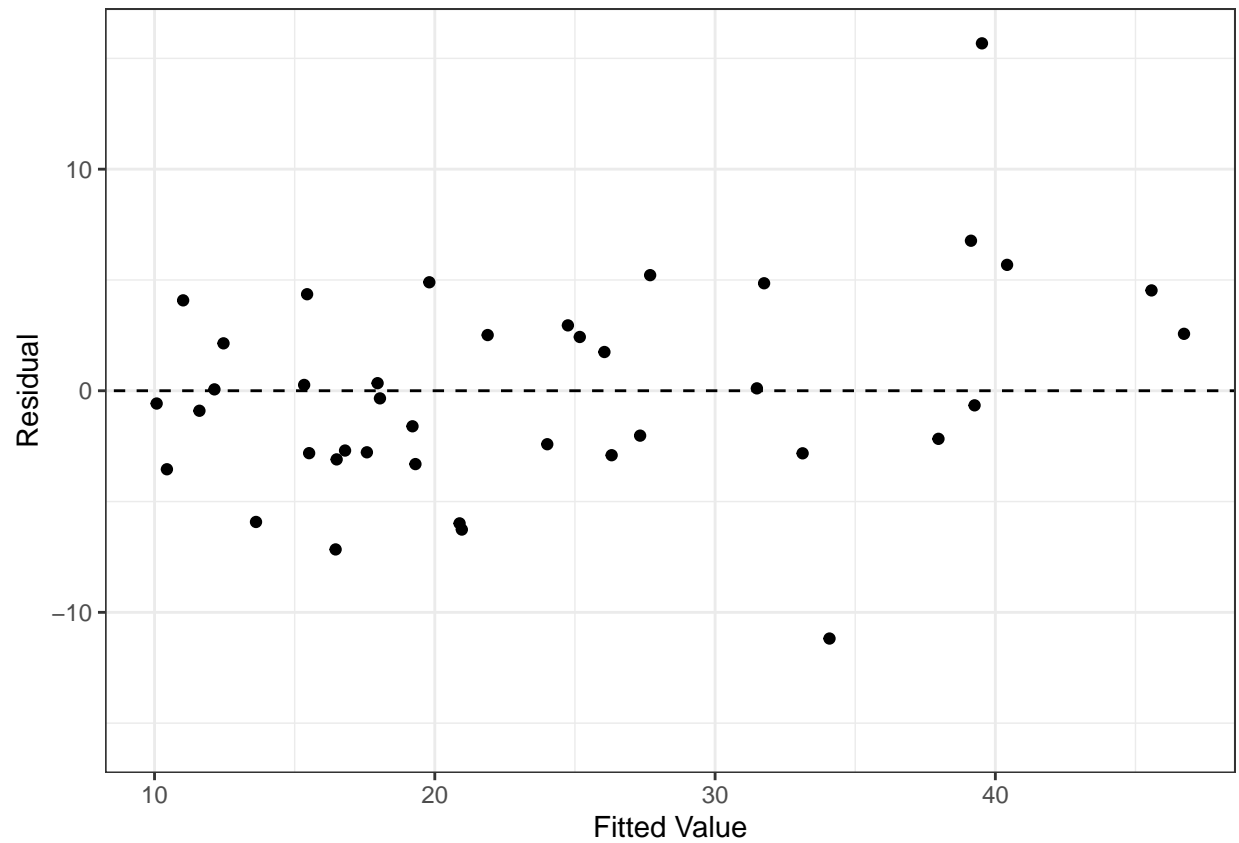

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treat	1	45.80	45.80	1.40	0.2529
Period	1	108.90	108.90	3.33	0.0856
baseline_diff	1	137.19	137.19	4.20	0.0563
Sequence	1	1.08	1.08	0.03	0.8578
Subject	17	5272.19	310.13	9.48	0.0000
Period:baseline_diff	1	204.69	204.69	6.26	0.0229
Residuals	17	555.88	32.70		

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	25.13	4.133396	22.98064	6.0797468	0.0000034
treatBEEF	2.14	2.055567	18.00000	1.0410753	0.3116230
Period2	-3.30	2.055567	18.00000	-1.6053965	0.1258074
SequenceB-C	-1.43	5.472163	18.00000	-0.2613226	0.7968103

```
model <- lmer(Post ~ treat + Period + Sequence + (1|Subject),
              data = data.baselines)
```

```
summary(model)$coefficients %>% kbl() %>% kable_paper()
```

```
model.metrics <- model %>% augment()
ggplot(data = model.metrics, mapping = aes(x = .fitted, y = .resid)) +
  geom_point() +
  geom_abline(slope = 0, linetype = "dashed") +
  scale_y_continuous(limits = symmetric_limits) +
  theme_bw() +
  labs(x = "Fitted Value", y = "Residual")
```



```
ggplot(data = model.metrics, mapping = aes(sample = .resid)) +  
  geom_qq() +  
  geom_qq_line() +  
  labs(x = "Normal Quantiles", y = "Residuals") +  
  theme_bw()
```

