Design and Analysis of 2x2 Cross-Over Trials with Continuous Data

Jules Lanari-Collard

MATH 410 Final Report

McGill University

Montréal, Québec, Canada

August 12, 2024

Supervised by Dr. José Correa

Abstract

In this project report, the design and analysis of 2x2 cross-over trials on continuous data is discussed and carried out. 2x2 cross-over trials are a type of clinical trial whereby subjects are assigned to a sequence of two treatments (either AB or BA, for example). Chapter 1 outlines the 2x2 cross-over design, and discusses its theoretical and practical benefits and drawbacks. Chapter 2 demonstrates how cross-over trial data can be summarised and visualised, and chapter 3 addresses different techniques used to analyse cross-over data, and puts them into practice with *R* code examples [1].

[1] Full source code is available in appendices A and B, and at the project's Github repository `https://github.com/velominatus/Analysis-of-Crossover-Trials`

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Cross-over Trials

1.1 The Cross-over Design

A cross-over trial is defined as a "trial in which subjects are given sequences of treatments with the object of studying differences between individual treatments" [1]. This differs from the traditional parallel group trial, where subjects typically only receive one treatment throughout the trial.

The most common and simple cross-over design is the 2x2 cross-over trial, whereby there are two treatments and two treatment groups. In this design, for example comparing two treatments A and B, subjects receive either A or B in the first period, and then are 'crossed over' to the other treatment in the second period. A and B could either be two active treatments, or one is a control treatment (or placebo).

Cross-over designs are typically used for testing treatments for ongoing or chronic diseases, where "there is no question of curing the underlying problem which has caused the illness but a hope of moderating its effects through treatment" [1]. Examples of such conditions include asthma, rheumatism, epilepsy and migraines. Due to time constraints and the risk of drop-out or carryover, cross-over designs are more suitable to single-dose trials, as opposed to trials involving multiple doses over a period of time [1].

1.1.1 Randomisation

A unique characteristic of the cross-over design is that only the *order* of treatments is randomised, which has the following consequences [2]:

- The validity of treatment comparison does not depend on randomisation.
- Randomisation does not guarantee an unbiased comparison of treatments.
- Treatment groups differ with respect to their recent exposure to potentially effective treatments.

In conclusion, the primary issue with the cross-over design is that the comparability of treatments is not guaranteed by the structure of the trial alone, but instead depends on the treatments themselves [2].

1.2 Advantages

The principal advantage of a cross-over design is that it can lead to significant savings in resources [1].

Firstly, when compared to a parallel group trial, a cross-over design only requires half as many subjects to obtain the same number of observations per treatment.

Secondly, the data can be interpreted in terms of subject-level *difference to control*, eliminating between-patient variation [1], which is usually greater than within-patient variation [2]. This reduces the number of observations (and hence subjects) required for the same precision in estimation. Furthermore, within-subject responses to treatments are usually positively correlated [2], reducing the variance of the estimated treatment difference (from control) and increasing efficiency, as demonstrated below:

Noticing that:

$$cov(\bar{Y}_A, \bar{Y}_B) = \rho_{AB} sd(\bar{Y}_A) sd(\bar{Y}_B) = \rho_{AB} \cdot \frac{\sigma}{\sqrt{n}} \cdot \frac{\sigma}{\sqrt{n}} = \rho_{AB} \frac{\sigma^2}{n}$$

Assuming constant variance in observations between patients and no period or carryover effects, we then have:

$$\begin{aligned} var(\hat{\delta}_{AB}) &= \frac{\sigma^2}{n} + \frac{\sigma^2}{n} - 2cov(\bar{Y}_A, \bar{Y}_B) \\ &= 2\frac{\sigma^2}{n}(1 - \rho_{AB}) \end{aligned}$$

Where σ^2 is the observation variance, $\bar{Y}_A$ and $\bar{Y}_B$ are the average observations for treatments A and B respectively, ρ_{AB} is the within-subject response correlation, and $\hat{\delta}_{AB} := \bar{Y}_A - \bar{Y}_B$.

Given that $\rho_{AB} = 0$ in parallel group trials, positive correlation of within-subject responses will make a crossover trial more efficient [2].

An oft-overlooked benefit to the crossover trial is improved recruitment and reduced spill-over rates [2]. Since all subjects are guaranteed to receive each treatment at least once, it can be easier to recruit participants. Spill-over is a significant risk for parallel group trials, where if a treatment is known to be desirable (e.g. exercise), subjects in the control group voluntarily take the treatment. Cross-over trials are evidently less at risk of this, since all subjects know they will receive the treatment.

1.3 Disadvantages

Many of the aforementioned benefits of the crossover design come with drawbacks. For example, the longer trial and multiple treatments could be seen instead as an added inconvenience to patients, and analysis of results is more difficult and complex (particularly when there are more than two treatments or groups). Also, the design cannot be used for infectious diseases, where either a significant deterioration or improvement in condition can occur during treatment, introducing bias into the subsequent treatment period. Patient

drop-out is also very harmful to cross-over trials, since observations for all treatments are required to analyse the within-patient differences for each patient. In comparison, a parallel group trial can still recover some information after a patient drops out [1].

There is also risk of *period-by-treatment interaction* complicating analysis, where the effect of the treatment is not constant over time [1]. In other words, when the period in which the treatment is administered affects the effectiveness of the treatment.

1.3.1 Carryover

The most common example of a period-by-treatment interaction is carryover, defined as "the persistence of a treatment applied in one period in a subsequent period of treatment" [1]. In a cross-over design, this occurs when at the beginning of the second treatment period, patients are not in the state they would have been in, had they not received treatment in the first period. This causes the effect of one treatment to be misinterpreted as the effect of both treatments combined, introducing bias to the treatment effect estimates.

Carryover is both difficult to test for and then to adjust for. Tests for carryover effects are difficult to interpret independently of the treatment effect, and including carryover parameters in the model introduces additional uncertainty and requires additional (usually unreasonable) assumptions [1].

1.3.2 Wash-out Period

Senn [1] proposes that the most efficient way to deal with carryover is introducing a *wash-out period* to the experiment design. A wash-out period is "a period in a trial during which the effect of a treatment given previously is believed to disappear" [1]. Effectively, time is allowed after the first treatment to allow any lingering effects to 'wash out'. After the wash-out period, we can assume that all measurements taken are no longer affected by the previous treatment, with the significant consequence being that *all conclusions become conditional on the absence of a carryover effect.*

Chapter 2

Summary & Visualisation of Cross-over Data

This chapter will follow example data from Jones and Kenward [3], on a study testing efficacy of a drug (A) and a placebo (B) on patients with *chronic obstructive pulmonary disease* (COPD). Due to the nature of the condition, the observation measurement used was *peak expiratory flow rate* (PEFR), defined as "a simple measure of the maximal flow rate that can be achieved during forceful expiration following full inspiration" [4]. The study employed a 2x2 cross-over design on 56 patients, where each patient measured their PEFR 3 times per day during the treatment periods, recording the highest value. A sub-sample of the data is outlined in table 2.1.

Table 2.1: Subsample of COPD Trial Data (PEFR in L/min)

Sequence	Subject	Subject Label	Period 1	Period 2
AB	1	7	121.90	116.67
AB	2	8	218.50	200.50
AB	3	9	235.00	217.14
AB	4	13	250.00	196.43
AB	5	14	186.19	185.50

2.1 Summary Tables

When the outcome data is continuous, a simple summary table showing a measure of central tendency and a measure of variability, grouped by sequence and period, suffices to give an overview of the data. If the data are normal, the mean and standard deviation should be used, and the median and interquartile range (IQR) for non-normal data [5]. Table 2.2 demonstrates the recommended format with the COPD trial data.

Table 2.2: Summary Table for COPD Trial Data

		Overall		Period 1		Period 2	
	Subjects	Mean	SD	Mean	SD	Mean	SD
AB	27	242.52	81.53	245.84	82.78	239.20	81.70
BA	29	223.08	72.99	215.99	72.63	230.16	73.94
Total	56	232.45	77.49	230.38	78.43	234.52	77.20

2.2 Summary Plots

Before undertaking any analysis, it is important to gain an understanding of the data and preliminary results. This section demonstrates the plots suggested by Jones and Kenward [3] for cross-over trials[1] [2].

An overview of the trial data can be obtained using a simple boxplot, separated by sequence and period, as shown in figure 2.1. A subject-profile plot, demonstrated in figure 2.2, provides a subject-level summary of the data. These can even be combined to give a comprehensive summary of the data, as shown in figure 2.3.

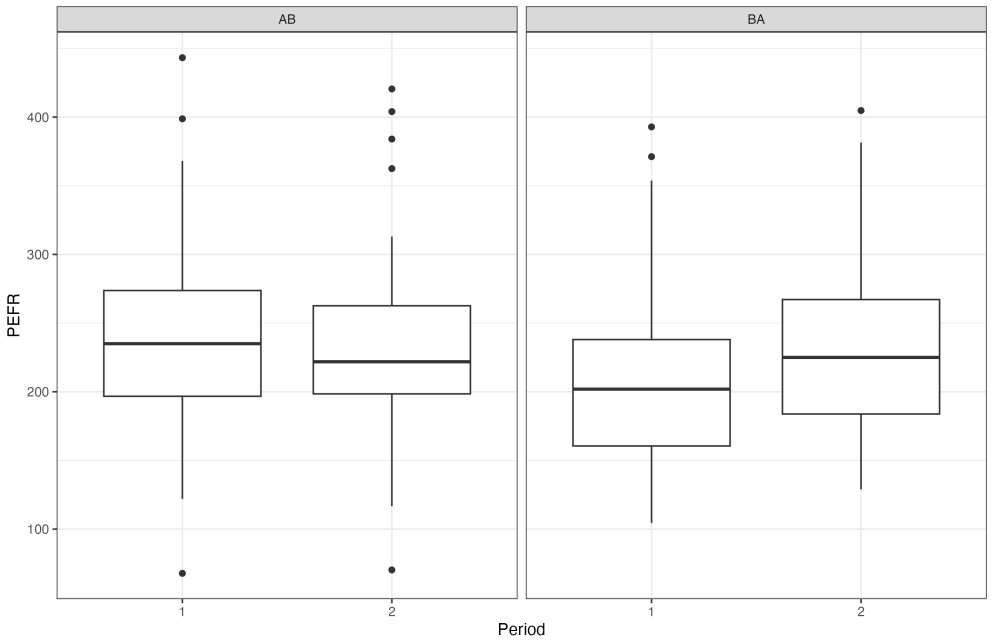

Figure 2.1: Boxplot

[1] These plots apply to any continuous 2x2 cross-over trial data; plots shown in this chapter use the COPD trial data for demonstration purposes.

[2] All plots shown were generated using the `ggplot2` package in R.

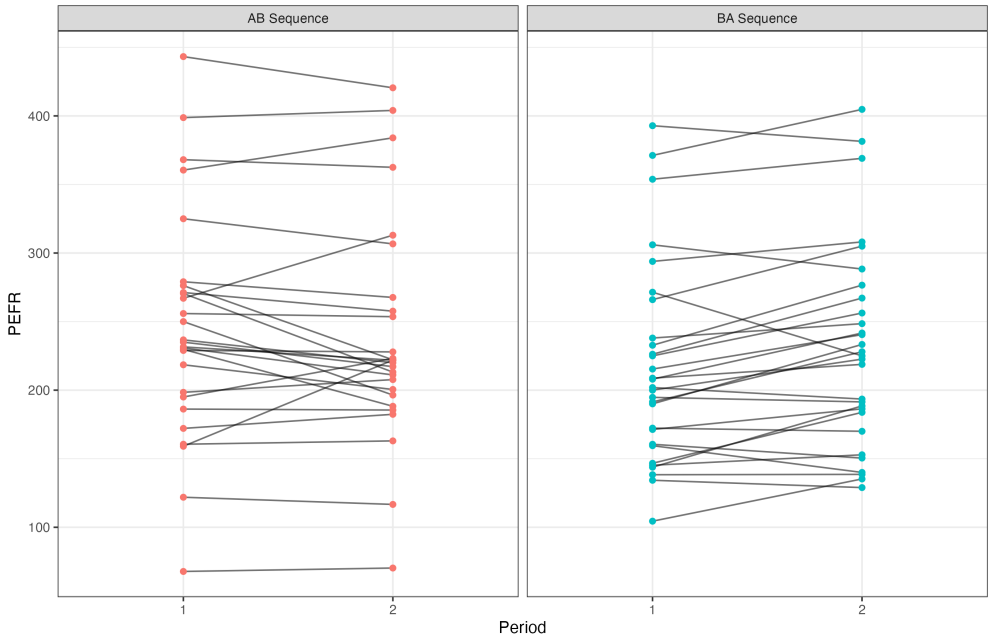

Figure 2.2: Subject-Profile Plot

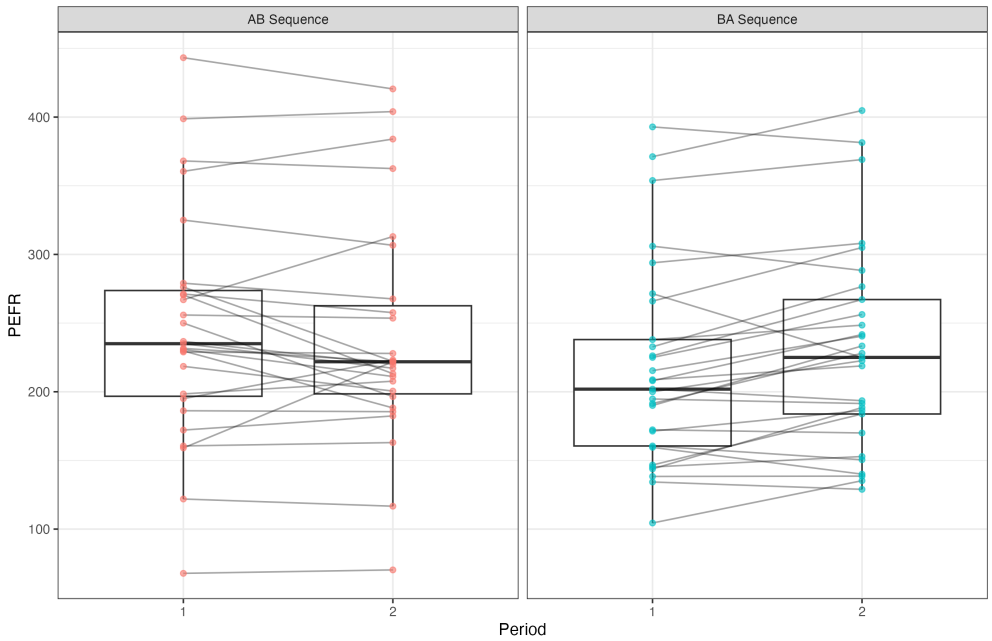

Figure 2.3: Paired Boxplot

To begin to understand potential treatment effects, we can use a simple scatter plot of observations in each period, separated by group, as demonstrated in figure 2.4. Also included is a straight line with slope 1 and intercept 0; if the treatments were identical we would expect all observations to lie along this line. Using this plot we can verify within-patient correlation, and also notice the spread of points across the line, indicating between-patient variation.

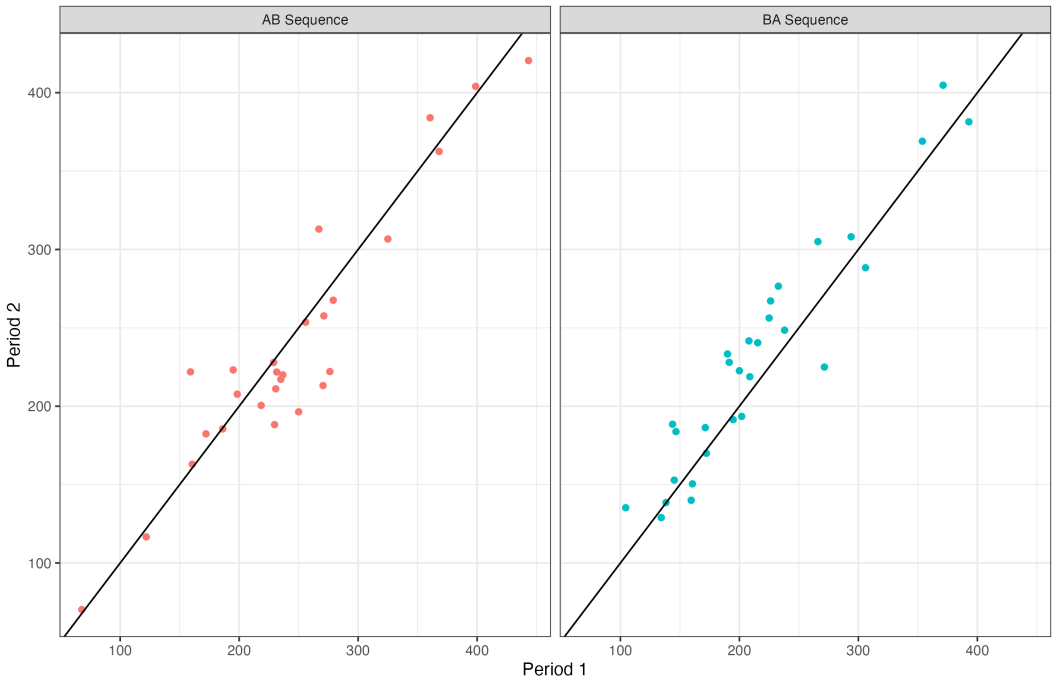

Figure 2.4: Period 2 vs Period 1 Plots

Additionally, we can overlay these two plots, distinguishing the groups by colour and including the centroids for each group, as demonstrated in figure 2.5. This allows us to make some preliminary conclusions:

- Both groups appearing symmetrical about the diagonal line suggests absence of a

period effect.

- Centroids appearing on opposite sides of the line, with some vertical separation, suggests there exists a direct treatment effect.

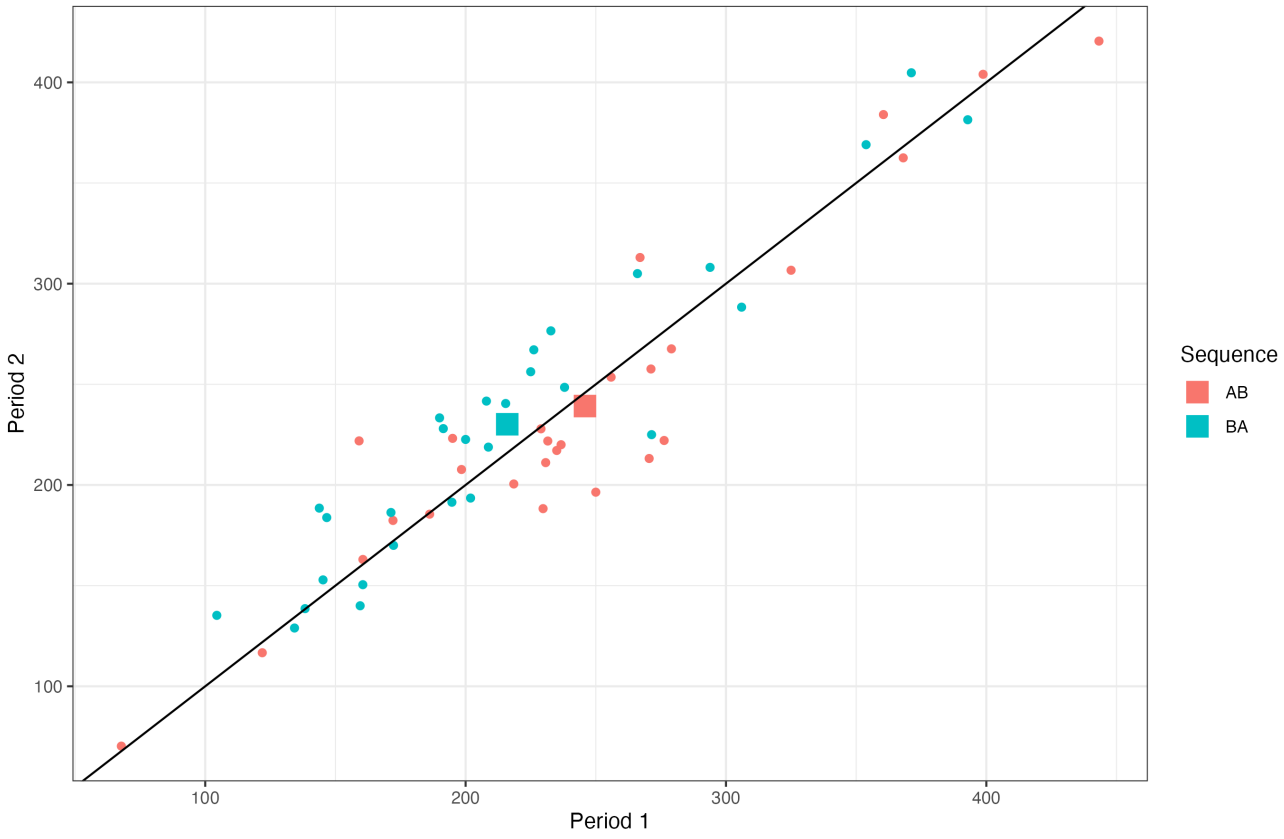

Figure 2.5: Period 2 vs Period 1 with Centroids

Finally, the summary table 2.2 can be visualised using a groups-by-periods plot, as shown in figure 2.6.

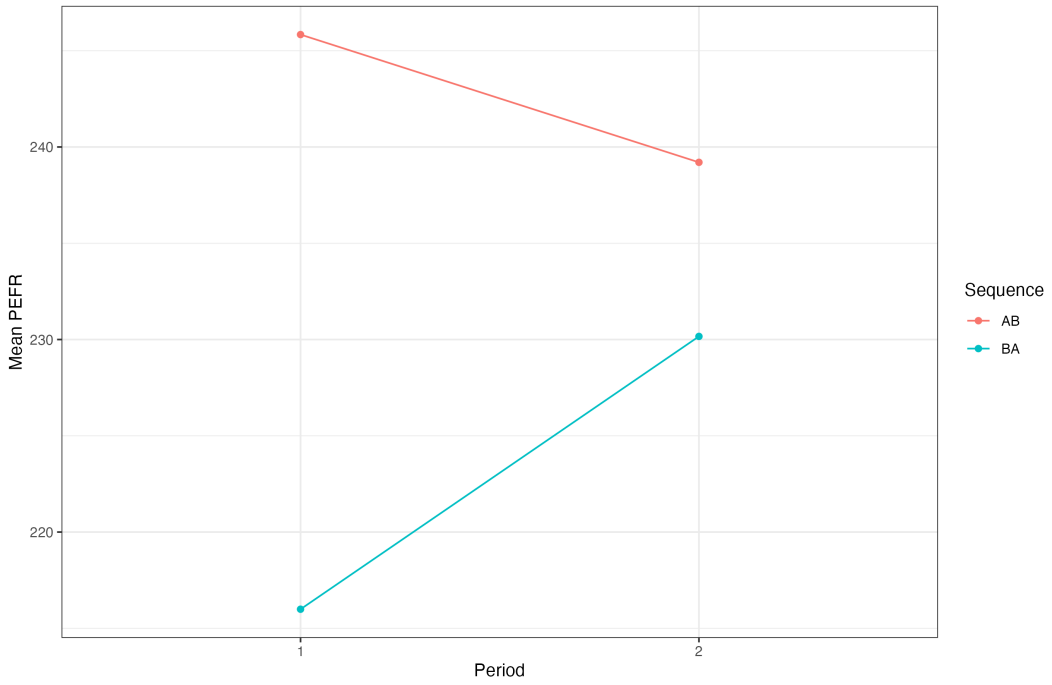

Figure 2.6: Groups-by-Periods Plot

Chapter 3

Analysis of 2x2 Cross-over Trials

This chapter uses data from Senn and Auclair [6] as an example for the analysis presented[1]. These data also consist of measurements of PEFR, this time on children suffering from athsma. The trial was a 2x2 cross-over design, comparing the effectiveness of two drugs; salbutamol (Sal) and formoterol (For). 13 patients were randomised into either a For-Sal sequence or a Sal-For sequence. A subsample of the data is presented in table 3.1[2].

[1] All the analysis shown in the *in practice* sections was done in the *R* programming language, with the `tidyverse` package loaded. Any additional packages required are indicated where applicable.

[2] For analysis, cross-over data should usually be in 'long' format, as in table 3.1.

Table 3.1: Subsample of For/Sal Data (L/min)

Sequence	Subject	Period	Treat	PEF
For-Sal	1	1	For	310
For-Sal	1	2	Sal	270
Sal-For	2	1	Sal	370
Sal-For	2	2	For	385
Sal-For	3	1	Sal	310
Sal-For	3	2	For	400

3.1 Matched-Pairs *t*-Test

The most straightforward analysis for a 2x2 cross-over trial uses a matched-pairs t-test, which takes advantage of the paired structure of the data (every observation in period 1 can be matched to the subject's corresponding observation in period 2) [1]. Note that carrying out the alternative, a simple two-sample t-test, would give the same result as if the experiment had been carried out as a parallel group trial, as the paired structure of the data would be ignored. First we will introduce some basic notation:

- n: number of subjects
- y_{A_k}, y_{B_k}: measured outcomes for subject k ($k = 1, \cdots, n$) after treatments A and B respectively.
- d_k: cross-over difference (i.e. difference between treatments) for subject k ($k = 1, \cdots, n$), defined as $d_k := y_{A_k} - y_{B_k}$.

And then the mean cross-over difference $\bar{d} := \frac{1}{n}\sum_{i=1}^{n} d_k$ is an estimator for the true treatment effect (or difference) τ. The matched-pairs t-test then simply involves testing the hypothesis $H_0 : \tau = 0$ against $H_1 : \tau \neq 0$, by computing a t-statistic and comparing it against the corresponding critical value [1]:

$$\begin{aligned} \bar{d} &:= \frac{1}{n}\sum_{i=1}^{n} d_k \\ s &= \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_k - \bar{d})^2} \\ \widehat{se(\bar{d})} &= \frac{s}{\sqrt{n}} \\ \implies \frac{\bar{d}}{\widehat{se(\bar{d})}} &\sim t_{n-1} \end{aligned}$$

In Practice

Provided the data is in 'long' format, as in table 3.1, we can easily carry out a paired t-test using the `rstatix::t_test()` function. The result of the test on the PEFR data is outlined in 3.2, showing a statistically significant difference between the For and Sal treatments.

```
library(rstatix)
data %>%
    arrange(Subject, Treat) %>% # sort for correct pairings
    t_test(PEF ~ Treat, paired = TRUE) # Table 3.2
```

Table 3.2: Matched-Pairs t-Test on For/Sal Data

	Group 1	Group 2	n1	n2	t-statistic	df	p
PEF	For	Sal	13	13	4.031	12	0.002

3.1.1 Assumptions

This method involves two important assumptions:

- Cross-over differences d_k are independently and randomly distributed about the true treatment effect τ.
- Cross-over differences d_k are distributed approximately normally.

The most important assumption is the first one, and there are a number of situations which could cause it to be violated:

Period effect

A period effect is present when there is a trend over time affecting the experiment as a whole. Under unbalanced allocation between treatment groups, this introduces bias to the treatment effect estimates. More generally, this can also cause us to overestimate the influence of random variation, as systematic changes between periods would be incorrectly "ascribed to random variation" [1].

Sequence Effect

Results may be affected by the sequence that subjects are assigned to. However, under successful randomisation, we can assume that there is no sequence effect [7] (recalling that in a cross-over design, only the *sequence* of treatments is randomised). Note that this assumption cannot be tested statistically [7].

Period-by-treatment interaction & carryover

Period-by-treatment interaction occurs when the effect of the treatment varies with the period in which it is given, which could affect the cross-over difference observations. As discussed in chapter 1, presence of a carryover effect introduces significant bias to the cross-over differences.

Patient-by-treatment interaction

Patient-by-treatment interaction occurs when the effect of the treatment varies from patient to patient. This effect increases variability and causes results to be difficult to interpret, and requires a more complex trial design (administering the same treatment multiple times) to investigate further [1].

Patient-by-period interaction

Further complication occurs if patients are subject to period effects which are different for each patient, known as a patient-by-period interaction. Once again this adds to the variance and difficulty of interpretation.

Consequences

Patient-by-period and patient-by-treatment interactions do not necessarily threaten the validity of the experiment, but they do increase variance and reduce ease of interpretation. As discussed above, period-by-treatment interactions (including carry-over) are best addressed by careful trial design, but are also the fundamental drawbacks to the cross-over design [1]. If present, they seriously threaten the validity of subsequent analysis. Period and sequence effects, whilst not ideal, can be adjusted for using more complex models, which are demonstrated later in this chapter.

3.2 Linear Models

The issues with the matched-pairs t-test, and the cross-over design as a whole, are better addressed within the framework of a linear model. This allows us to control for period and sequence effects, whilst also taking into account the within-subject differences. It also allows scope to include other variables of interest as covariates, such as sex and age.

The general framework for a cross-over trial linear model is shown in equation 3.1; some

decisions need to be made on how to treat each individual variable to then implement regression.

$$Y_{ijk} = \mu + \tau_{jk} + \pi_j + \lambda_{j-1,k} + \gamma_k + s_{ik} + \epsilon_{ijk} \tag{3.1}$$

where Y_{ijk} is the response of the ith subject in the kth sequence at the jth period. The model then includes the direct treatment effect τ_{jk}, period effect π_j, carryover effect $\lambda_{j-1,k}$, sequence effect γ_k and subject effects s_{ik}. The errors ϵ_{ijk} are i.i.d. with mean 0 and variance σ_s^2 [7].

The direct treatment effect, period effect and sequence effect should all be considered as fixed effects [7]. As mentioned in the preceding discussion about carryover effect, it should be mitigated at the experiment design stage, as its presence seriously threatens the validity of the trial and post-hoc analysis. Furthermore, in the 2x2 cross-over design specifically, the carryover effect is inherent in the sequence [7]. Hence in the regression models below, no carryover term is included, but instead we include a term for the sequence effect.

If we assume the subject effects s_{ik} to be unknown fixed parameters, the regression will only use information from within-subject comparisons. ANOVA and regression in this case become excessively complicated. Alternatively, when subject totals are judged to contain useful information, the subject effects should be treated as random effects as part of a random intercepts mixed model [3].

3.2.1 Mixed Model for Cross-over Design

The more intuitive approach for 2x2 cross-over trials, judging between-subject differences to be relevant, is a random intercepts mixed model [3]. In a random intercepts mixed model, treating subject effects as random effects, we include a term for the unique effect of each subject, so that we have subject-specific intercepts [8]. In this way, within-subject information is incorporated into the model. The corresponding regression is demonstrated in equation 3.2, and can be estimated using the Restricted Maximum Likelihood (REML) method [3].

$$Y = \beta_0 + \beta_1 T + \beta_2 P + \beta_3 Se + \beta_{subject} + \epsilon \tag{3.2}$$

where Y is the outcome variable, T is the treatment administered, P is the period, Se is the sequence, and $\beta_{subject}$ the subject-specific effect. Notice that β_1 represents the direct treatment effect.

Additional subject-level covariates, such as sex and age, can be added as covariates in the same way as for a simple linear regression.

In Practice

The `lme4` package provides functionality to easily implement mixed models in R, and the `lmerTest` calculates p-values for such models. For the random-intercept model specifically,

we include the term `(1|Subject)` to specify that the intercept should vary by subject [8].

```
library(lme4)

library(lmerTest)

mixed.model <- lmer(PEF~Treat+Period+Sequence + (1 | Subject),

                    data = data)
```

3.2.2 Verifying Assumptions

Before analysing model results, it is important to verify that some key assumptions hold, to ensure validity of the fitted model. The following assumptions must hold for either the simple regression or the mixed model to be valid for a 2x2 cross-over trial:

- Linearity
- Homoscedasticity
- Normality of residuals

Linearity is not a concern as the independent variable (treatment) only takes 2 values. Homoscedasticity and normality of residuals are best verified graphically using a scatter plot of residuals against fitted values and a Q-Q plot, respectively.

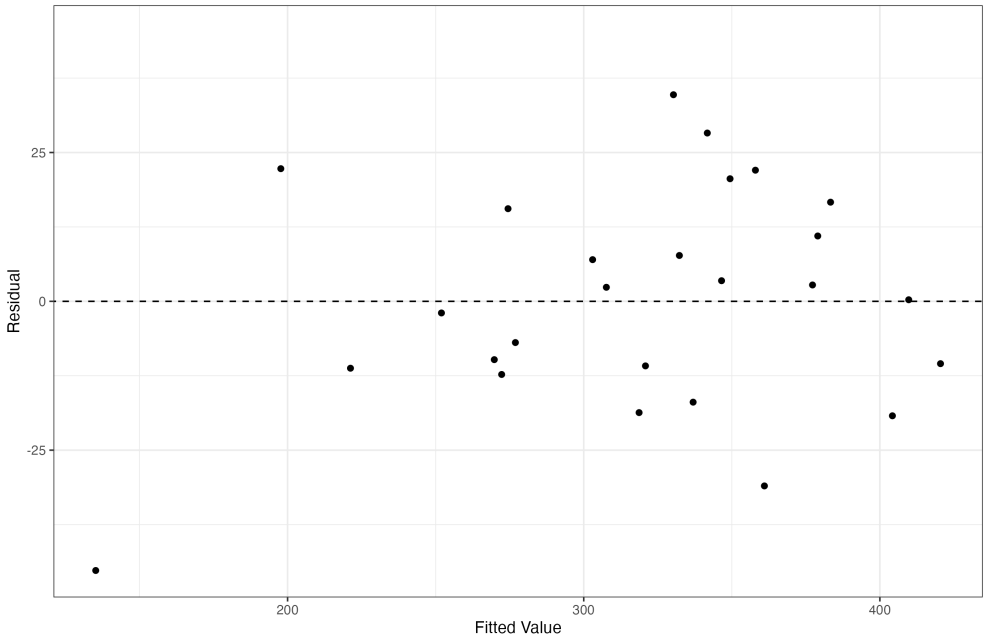

Figure 3.1: Plot to verify homoscedasticity

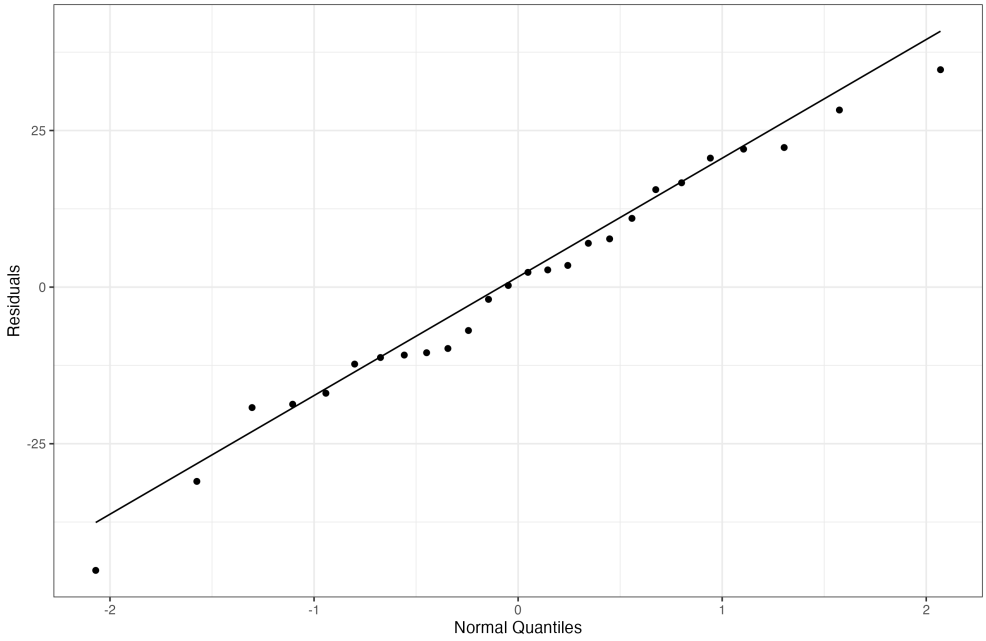

Figure 3.2: Q-Q Plot to verify normality of residuals

In Practice

The `broom.mixed::augment()` function extends the original dataframe with detailed information from the model, including residuals and fitted values. These data can then be used to construct the plots demonstrated in figures 3.1 and 3.2.

```
library(broom.mixed)

mixed.model.metrics <- mixed.model %>% augment()
```

For these data, figures 3.1 and 3.2 provide no evidence to suggest that either assumption has been violated.

3.2.3 Analysing Results

Whichever model is chosen, the estimated coefficients and their corresponding p-values should be reported, as in table 3.3. We can then use the p-values to determine the existence of a treatment effect (or a difference between treatments). It is also useful to report the adjusted (LS) means (see table 3.4) for each treatment, which provide an insightful summary of the differences between treatments, and hence the magnitude of the treatment effect (if any).

In Practice

The `broom.mixed::tidy()` function can be used to obtain the model coefficients, std. errors, p-values etc. shown in table 3.3:

Table 3.3: Mixed Model Estimates for For/Sal Data

	Estimate	Std. Error	df	t	p
(Intercept)	337.14	28.28	11.92	12.57	<0.01
TreatSal	-46.61	10.78	-4.32	11.00	<0.01
Period2	15.89	10.78	1.47	11.00	0.17
SequenceSal-For	-7.20	40.20	-0.18	11.00	0.86

Table 3.4: LS Means for Mixed Model on For/Sal Data

Sequence	Difference	Adj. Mean	SE	df	Lower CI	Upper CI
For	.	341.49	20.81	12.57	284.06	398.92
Sal	.	294.88	20.81	12.57	237.45	352.31
.	For - Sal	46.61	10.78	11.00	16.22	77.00

```
library(broom.mixed)

tidy(mixed.model) # Table 3.3
```

We see that there is a statistically significant difference between the treatments, at the $< 1\%$ significance level.

The package `emmeans` can calculate adjusted means for mixed models; we specify `~ Treat` as an argument to separate the adjusted means by treatment.

```
library(emmeans)

emm <- emmeans(mixed.model, pairwise ~ Treat)

emm %>% rbind(emm$contrasts) # Table 3.4
```

3.3 Adjusting for Baseline Values

In pre-post designs, analysis of covariance (ANCOVA), whereby the pre-test (baseline) measurement is included as a covariate, is the most efficient method of analysis [9]. For 2x2 cross-over trials, we can further improve model efficiency by taking baseline measurements into account (when available). This requires that baseline measurements be taken for each subject prior to each treatment period, so that a total of 4 measurements per subject are taken over the course of the trial.

3.3.1 Insect- and Beef-Derived Proteins Study

This section uses data from a cross-over study by Dai, Lov, Martin-Arrowsmith et al. [10]. The trial measured various outcomes after ingesting a protein-based beverage, comparing between beverages with beef- or insect-derived proteins. 20 subjects were randomised into either a cricket-beef (C-B) sequence or a beef-cricket (B-C) sequence. For this example, we use the data collected on the subjects' insulin levels. For each treatment period, insulin levels were measured just before ingestion (Pre), and 300 minutes after ingestion (Post), resulting in 4 measurements per subject. A subsample of the data is shown in table 3.5.

Table 3.5: Subsample of Protein Data

		Period 1		Period 2	
Subject	Sequence	Pre	Post	Pre	Post
1	C-B	20.3	18.3	14.6	14.1
2	C-B	11.1	14.7	11.1	24.7
4	C-B	38.3	45.9	51.6	35.8
5	C-B	40.3	46.1	51.8	38.6
7	C-B	24.6	17.6	31.3	17.7

3.3.2 Tables and Plots

We can expand the summary table to include the pre-treatment measurements as well, as shown in table 3.6.

The boxplot can also be extended to show the pre-treatment measurements, as shown in figure 3.3.

3.3.3 Mixed Model with Baselines

After comparison of many different models, Mehrotra [11] concludes that the most efficient method to interpret baseline measurements is to include *within-subject difference in baselines*

Table 3.6: Summary Table for Protein Data (with Baselines)

		Overall				Period 1				Period 2			
		Pre		Post		Pre		Post		Pre		Post	
Sequence	Subjects	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

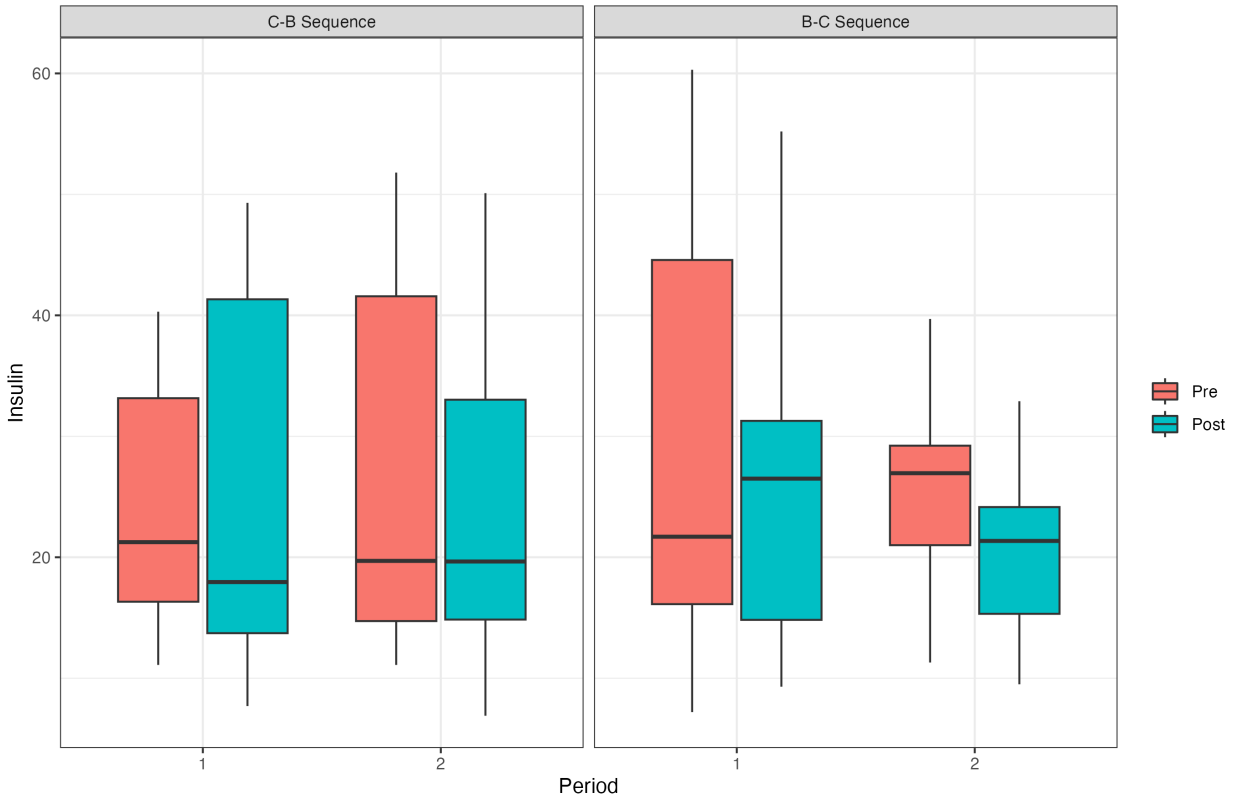

Figure 3.3: Boxplot with Baseline Measurements for Protein Data

as a covariate, interacting with the period effect.

As in section 3.2.1, a random intercepts mixed model, treating subject effects as random, remains a suitable method to account for both between- and within-subject differences. The new mixed model, including within-subject differences in baseline as a covariate, is shown in equation 3.3.

$$Y = \beta_0 + \beta_1 T + \beta_2 P \cdot X_{diff} + \beta_3 P + \beta_4 X_{diff} + \beta_5 Se + \beta_{subject} + \epsilon \tag{3.3}$$

Where $X_{diff} := X_1 - X_2$, with X_1 and X_2 being the baseline measurements before treatment periods 1 and 2 respectively.

Table 3.7: Subsample of Protein Data (Long Format with Baseline Differences)

Subject	Sequence	Period	Treatment	Pre	Post	Pre_diff
1	C-B	1	CRICKET	20.3	18.3	5.7
1	C-B	2	BEEF	14.6	14.1	5.7
2	C-B	1	CRICKET	11.1	14.7	0.0
2	C-B	2	BEEF	11.1	24.7	0.0
4	C-B	1	CRICKET	38.3	45.9	-13.3

In Practice

As in section 3.2.1, first the data must be pivoted into long format, this time also with a column for within-subject baseline differences, as in table 3.7 (see appendix for more detailed code).

Now we implement the model using `lme4::lmer()`:

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

At this point, the assumptions should be verified as in section 3.2.2. The resulting estimates for the protein trial data are shown in table 3.8, with the adjusted means shown in table 3.9. We notice that there is no statistically significant difference between the cricket- and beef-derived proteins, in terms of ensuing insulin levels.

Table 3.8: Mixed Model on Protein Data Estimates with Baseline Interaction

	Estimate	Std. Error	df	t	p
(Intercept)	25.30	4.23	5.97	20.55	<0.01
TreatmentBEEF	0.71	1.90	0.37	17.00	0.71
Period2	-3.24	1.81	-1.79	17.00	0.09
Pre_diff	0.05	0.26	0.19	20.55	0.85
SequenceB-C	-0.35	5.84	-0.06	17.00	0.95
Period2:Pre_diff	-0.41	0.16	-2.50	17.00	0.02

Table 3.9: LS Means for Mixed Model on Protein Data

Sequence	Difference	Adj. Mean	SE	df	Lower CI	Upper CI
CRICKET	.	23.48	2.94	20.89	15.82	31.14
BEEF	.	24.19	2.94	20.89	16.53	31.85
.	CRICKET - BEEF	-0.71	1.90	17.00	-5.75	4.33

Bibliography

[1] S. Senn, *Cross-over Trials in Clinical Research*. John Wiley & Sons, 2nd ed., 2002.

[2] S. Piantadosi, *Clinical Trials: A Methodologic Perspective*. John Wiley & Sons, 2nd ed., 2005.

[3] B. Jones and M. G. Kenward, *Design and Analysis of Cross-Over Trials*. Chapman & Hall, 2nd ed., 2003.

[4] B. W. DeVrieze, P. Modi, and A. O. Giwa, "Peak flow rate measurement," 31.07.2023.

[5] T. R. Vetter, "Descriptive statistics: Reporting the answers to the 5 basic questions of who, what, why, when, where, and a sixth, so what?," *Anesthesia & Analgesia*, vol. 125, pp. 1797–1802, 2017.

[6] S. Senn and P. Auclair, "The graphical representation of clinical trials with particular reference to measurements over time," *Statistics in Medicine*, vol. 9, pp. 1287–1302, 1990.

[7] C.-Y. Lim and J. In, "Considerations for crossover design in clinical study," *Korean Journal of Anesthesiology*, vol. 74, pp. 293–299, 2021.

[8] M. Clark, "Mixed models with r."

[9] F. Wan, "Statistical analysis of two arm randomized pre-post designs with one post-treatment measurement," *BMC Medical Research Methodology*, vol. 21, 2021.

[10] J. Dai, J. Lov, P. W. Martin-Arrowsmith, A. Gritsas, and T. A. Churchward-Venne, "The acute effects of insect vs. beef-derived protein on postprandial plasma aminoacidemia, appetite hormones, appetite sensations, and energy intake in healthy young men," *European Journal of Clinical Nutrition*, vol. 76, pp. 1548–1556, 2022.

[11] D. V. Mehrotra, "A recommended analysis for 2x2 crossover trials with baseline measurements," *Pharmaceutical Statistics*, vol. 13, pp. 376–387, 2014.

Appendix A

R Code for Chapter 2

```
library(tidyverse)

library(readxl)

library(kableExtra)


# Data Import

data <- read_xlsx("data/DataExample_Ch2_JK.xlsx")

data <- data %>%

  arrange(Group, Subject) %>%

  mutate(Sequence = if_else(Group == 1, "AB", "BA")) %>%

  mutate(across(c(Group, Sequence, Subject, Subject_Label), as_factor))
```

```
data.long <- data %>% # Long format needed for summaries
  rename(`1` = "Period1", `2` = "Period2") %>%
  pivot_longer(cols = c(`1`, `2`), names_to = "Period", values_to = "PEFR",
               names_transform = list(Period = as_factor)) %>%
  mutate(Treatment = if_else((Sequence == "AB" & Period == 1) |
                               (Sequence == "BA" & Period == 2), "A", "B")) %>%
  mutate(across(Treatment, as_factor))


# Table 2.1 - Subsample
sink("report/tables/ch2/pefrDataSubsample.tex", type="output")
data %>%
  select(Sequence, Subject, Subject_Label, Period1, Period2) %>%
  arrange(Sequence, Subject) %>%
  rename(`Subject Label` = "Subject_Label", `Period 1` = "Period1",
         `Period 2` = "Period2") %>%
  head(5) %>%
  kbl(format = "latex",
      caption = "Subsample of COPD Trial Data (PEFR in L/min)",
      label = "pefrDataSubsample",
      booktabs = TRUE,
```

```
      digits = 2)
sink()


# Constructing Summary Table
summary.period.sequence <- data %>% # Summary by period & sequence
  group_by(Sequence) %>%
  summarise(Subjects = n(),
            `Mean PEFR Period 1` = round(mean(Period1), 2),
            `SD PEFR Period 1` = round(sd(Period1), 2),
            `Mean PEFR Period 2` = round(mean(Period2), 2),
            `SD PEFR Period 2` = round(sd(Period2), 2))


summary.sequence <- data.long %>% # Summary by sequence
  group_by(Sequence) %>%
  summarise(`Overall Mean PEFR` = round(mean(PEFR), 2),
            `Overall SD PEFR` = round(sd(PEFR), 2))


summary.period <- data.long %>% # Summary by period
  group_by(Period) %>%
  summarise(ov_mean = round(mean(PEFR), 2),
```

```
            ov_sd = round(sd(PEFR), 2)) %>%
  pivot_wider(names_from = Period,
              values_from = c(ov_mean, ov_sd)) %>%
  rename(`Mean PEFR Period 1` = "ov_mean_1",
         `SD PEFR Period 1` = "ov_sd_1",
         `Mean PEFR Period 2` = "ov_mean_2",
         `SD PEFR Period 2` = "ov_sd_2") %>%
  mutate(Sequence = "Total", Subjects = 56,
         `Overall Mean PEFR` = round(mean(data.long$PEFR), 2),
         `Overall SD PEFR` = round(sd(data.long$PEFR), 2),
         .before = `Mean PEFR Period 1`)

summary.table <- inner_join(summary.period.sequence, summary.sequence,
                            by = "Sequence") %>%
  bind_rows(summary.period) %>%
  select(Sequence, Subjects,
         colnames(summary.sequence),
         colnames(summary.period.sequence))

# Table 2.2 - Summary
```

```
sink("report/tables/ch2/pefrDataSummary.tex", type="output")

summary.table %>%

  kbl(format="latex",

      caption = "Summary Table for COPD Trial Data",

      label = "pefrDataSummary",

      booktabs = TRUE,

      digits = 2,

      col.names = c("", "Subjects", rep(c("Mean", "SD"), 3)),

      linesep = "") %>%

  add_header_above(c(" "=2, "Overall"=2, "Period 1"=2, "Period 2"=2)) %>%

  row_spec(2, hline_after = TRUE) %>%

  column_spec(2, border_right = TRUE)

sink()


# Figure 2.1 - Boxplot

ggplot(data = data.long, mapping = aes(x=Period, y = PEFR)) +

  geom_boxplot() +

  facet_wrap(~Sequence) +

  theme_bw()

ggsave("report/figures/ch2/boxplot.png")
```

```r
# Figure 2.2 - Subject-Profiles Plot
ggplot(data = data.long, mapping = aes(x=Period, y=PEFR)) +
  geom_line(aes(group = Subject_Label), alpha = 0.55) +
  geom_point(aes(colour = Sequence), show.legend = FALSE) +
  facet_wrap(~Sequence) +
  theme_bw()
ggsave("report/figures/ch2/subjectProfilesPlot.png")

# Figure 2.3 - Paired boxplot
ggplot(data = data.long, mapping = aes(x=Period, y=PEFR)) +
  facet_wrap(~Sequence) +
  geom_boxplot(outliers = FALSE) +
  geom_line(aes(group = Subject_Label), alpha = 0.35) +
  geom_point(alpha=0.6, mapping = aes(col = Sequence), show.legend = FALSE) +
  theme_bw()
ggsave("report/figures/ch2/pairedBoxplot.png")

# Figure 2.4 - Period 1 vs Period 2 Plot
ggplot(data = data, mapping = aes(x=Period1, y=Period2, colour=Sequence)) +
```

```
  geom_point(show.legend = FALSE) +
  geom_abline(slope=1,intercept=0) +
  facet_wrap(~Sequence) +
  labs(x = "Period 1", y = "Period 2") +
  theme_bw()
ggsave("report/figures/ch2/periodsPlot.png")


# Figure 2.5 - Period 1 vs Period 2 with Centroids
means <- data %>% # Calculate centroids
  group_by(Sequence) %>%
  summarise(Period1 = mean(Period1), Period2 = mean(Period2))


ggplot(data = data, mapping = aes(x=Period1, y=Period2, colour=Sequence)) +
  geom_point() +
  geom_point(data = means, size=5, shape="square") +
  geom_abline(slope=1,intercept=0) +
  labs(x = "Period 1", y = "Period 2") +
  theme_bw()
ggsave("report/figures/ch2/centroidsPlot.png")
```

```
# Figure 2.6 Groups-by-Periods Plot

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


ggplot(data = means.long, mapping = aes(x = Period, y = PEFR, colour = Sequence)) +
  geom_line(aes(group = Sequence)) +
  geom_point() +
  labs(y = "Mean PEFR") +
  theme_bw()
ggsave("report/figures/ch2/groupsByPeriodsPlot.png")
```

Appendix B

R Code for Chapter 3

B.1 *t*-Test

```
library(tidyverse)

library(readxl)

library(rstatix)

library(kableExtra)


# Data Import

data <- read_xlsx("data/CrossOverData.xlsx") %>%

  mutate(across(Sequence:Treat, as_factor))
```

```
# Table 3.1 - Subsample

sink("report/tables/ch3/forSalData.tex")

data %>%

  arrange(Subject, Period) %>% head(6) %>%

  kbl(format="latex",

      caption="Subsample of For/Sal Data (L/min)",

      label="forSalData",

      booktabs = TRUE,

      digits=2,

      linesep="",

      positioning="bht")

sink()


# Matched-pairs t-test

data %>%

  arrange(Subject, Treat) %>%

  t_test(PEF ~ Treat, paired = TRUE)


# Table 3.2 - t-test

sink("report/tables/ch3/tTest.tex")
```

```
data %>%
  arrange(Subject, Treat) %>%
  t_test(PEF ~ Treat, paired = TRUE) %>%
  kbl(format="latex",
      caption="Matched-Pairs t-Test on For/Sal Data",
      label="tTest",
      booktabs = TRUE,
      digits = 3,
      col.names = c("", "Group 1", "Group 2", "n1", "n2", "t-statistic", "df", "p"))
sink()
```

B.2 Mixed Model

```
library(tidyverse)
library(readxl)
library(kableExtra)
library(lme4)
library(lmerTest)
library(emmeans)
library(broom.mixed)
library(ggpmisc)
```

```
library(rstatix)


# Data Import

data <- read_xlsx("data/CrossOverData.xlsx") %>%

  mutate(across(Sequence:Treat, as_factor))


# Mixed Model

mixed.model <- lmer(PEF~Treat+Period+Sequence + (1 | Subject), data = data)


# Table 3.3 - Mixed Model Estimates

sink("report/tables/ch3/pefrDataEstimates.tex", type="output")

tidy(mixed.model) %>%

  filter(effect == "fixed") %>%

  select(-c(effect, group)) %>%

  kbl(format="latex",

      caption="Mixed Model Estimates for For/Sal Data",

      col.names = c("", "Estimate", "Std. Error", "df", "t", "p"),

      label="pefrDataEstimates",

      booktabs=TRUE,

      digits=2) %>%
```

```
  column_spec(1, border_right = TRUE)
sink()


# LS Means
emm <- emmeans(mixed.model, pairwise ~ Treat)


# Table 3.4 - LS Means
sink("report/tables/ch3/pefrDataMeans.tex", type="output")
emm %>% rbind(emm$contrasts) %>%
  kbl(format="latex",
      caption="LS Means for Mixed Model on For/Sal Data",
      label="pefrDataLSMeans",
      col.names = c("Sequence", "Difference", "Adj. Mean", "SE", "df",
                    "Lower CI", "Upper CI"),
      booktabs = TRUE,
      digits=2) %>%
  column_spec(2, border_right = TRUE)
sink()


# Figure 3.1 - homoscedasticity
```

```r
model.metrics <- mixed.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")
ggsave("report/figures/ch3/homoscedasticity.png")


# Figure 3.2 - Q-Q Plot
ggplot(data = model.metrics, mapping = aes(sample = .resid)) +
  geom_qq() +
  geom_qq_line() +
  labs(x = "Normal Quantiles", y = "Residuals") +
  theme_bw()
ggsave("report/figures/ch3/qqplot.png")
```

B.3 Adjusting for Baseline Values

```
library(tidyverse)

library(readxl)

library(kableExtra)

library(lme4)

library(lmerTest)

library(emmeans)

library(broom.mixed)


# Data Import

data <- read_xlsx("data/CrossOverData2.xlsx") %>%

  mutate(across(Subject:treat, as_factor)) %>%

  rename(Treatment = "treat")


data.wide <- data %>%

  pivot_wider(id_cols = c(Subject, Sequence),

              names_from = Period, values_from = c(Pre, Post)) %>%

  relocate(Post_1, .after = "Pre_1")


data.long <- data %>%
```

```
  pivot_longer(cols = c(Pre,Post),
               names_to = "Measurement",
               values_to = "Insulin",
               names_transform = as_factor)


# Table 3.5 - Subsample
sink("report/tables/ch3/proteinDataSubsample.tex", type="output")
data.wide %>% arrange(Subject) %>%
  head(5) %>%
  kbl(format="latex",
      caption="Subsample of Protein Data",
      label="proteinDataSubsample",
      booktabs=TRUE,
      digits = 1,
      col.names = c("Subject", "Sequence", "Pre", "Post", "Pre", "Post")) %>%
  add_header_above(c(" "=2, "Period 1"=2, "Period 2" = 2)) %>%
  column_spec(2, border_right = TRUE)
sink()


# Constructing Summary Table
```

```
summary.period.sequence <- data.wide %>% # Summary by period & sequence
  group_by(Sequence) %>%
  summarise(Subjects = n(),
            meanPre1 = mean(Pre_1),
            sdPre1 = sd(Pre_1),
            meanPost1 = mean(Post_1),
            sdPost1 = sd(Post_1),
            meanPre2 = mean(Pre_2),
            sdPre2 = sd(Pre_2),
            meanPost2 = mean(Post_2),
            sdPost2 = sd(Post_2))


summary.sequence <- data.long %>% # Summary by sequence
  group_by(Sequence, Measurement) %>%
  summarise(meanOverall = mean(Insulin),
            sdOverall = sd(Insulin)) %>%
  pivot_wider(names_from = Measurement, values_from = c(meanOverall, sdOverall),
              names_sep = "") %>%
  select(Sequence, meanOverallPre, sdOverallPre, meanOverallPost, sdOverallPost)
```

```
summary.period <- data %>% # Summary by period
  group_by(Period) %>%
  summarise(meanPre = mean(Pre),
            sdPre = sd(Pre),
            meanPost = mean(Post),
            sdPost = sd(Post)) %>%
  pivot_wider(names_from = Period, values_from = meanPre:sdPost,
              names_sep = "") %>%
  mutate(Sequence = "Total", Subjects = n_distinct(data$Subject),
         meanOverallPre = mean(data$Pre),
         sdOverallPre = sd(data$Pre),
         meanOverallPost = mean(data$Post),
         sdOverallPost = sd(data$Post),
         .before = meanPre1)


summary.table <- inner_join(summary.sequence, summary.period.sequence,
                            by = "Sequence") %>%
  bind_rows(summary.period) %>%
  relocate(Subjects, .after = Sequence)
```

```
# Table 3.6 - Summary

sink("report/tables/ch3/proteinDataSummary.tex", type="output")

summary.table %>%

  kbl(format="latex",

      caption = "Summary Table for Protein Data (with Baselines)",

      label = "proteinDataSummary",

      booktabs = TRUE,

      col.names = c("Sequence", "Subjects", rep(c("Mean","SD"), 6)),

      digits = 2,

      linesep="") %>%

  add_header_above(c(" "=2,

                     "Pre"=2, "Post"=2,

                     "Pre"=2, "Post"=2,

                     "Pre"=2, "Post"=2)) %>%

  add_header_above(c(" "=2, "Overall"=4, "Period 1"=4, "Period 2"=4)) %>%

  row_spec(2, hline_after = TRUE) %>%

  column_spec(c(2, 6), border_right = TRUE) %>%

  kable_styling(latex_options = "scale_down")

sink()
```

```r
# Figure 3.3 - Boxplot

labels=c(`C-B`="C-B Sequence", `B-C`="B-C Sequence")

ggplot(data = data.long, mapping = aes(x = Period, y = Insulin, fill = Measurement)) +

  facet_wrap(~Sequence, labeller = as_labeller(labels)) +

  geom_boxplot() +

  theme_bw() +

  labs(fill="")

ggsave("report/figures/ch3/proteinBoxplot.png")


# Data long with baselines

data.baselines <- data.wide %>%

  mutate(Pre_diff = Pre_1 - Pre_2) %>%

  rename(`1` = "Post_1", `2` = "Post_2") %>%

  pivot_longer(cols = c(`1`,`2`), names_to = "Period", values_to = "Post",

               names_transform = as_factor) %>%

  select(Subject, Period, Pre_diff) %>%

  inner_join(data, join_by(Subject == Subject, Period == Period)) %>%

  relocate(Sequence, .after = "Subject") %>%

  relocate(Pre_diff, .after = "Post")
```

```
# Table 3.7 - Data Long with Baseline Differences
sink("report/tables/ch3/preDiffSubsample.tex", type="output")
data.baselines %>% arrange(Subject) %>%
  head(5) %>%
  kbl(format="latex",
      caption="Subsample of Protein Data (Long Format with Baseline Differences)",
      label="preDiffSubsample",
      booktabs=TRUE,
      digits = 1,
      position = "hbt")
sink()


# Mixed Model with Baselines
mixed.model.baselines <-
  lmer(Post ~ Treatment + Period * Pre_diff + Sequence + (1|Subject),
                                data = data.baselines)


# Table 3.8 - Mixed Model Estimates with Baselines
sink("report/tables/ch3/proteinDataEstimates.tex", type="output")
tidy(mixed.model.baselines) %>%
```

```
  filter(effect == "fixed") %>%
  select(-c(effect, group)) %>%
  kbl(format="latex",
      caption="Mixed Model on Protein Data Estimates with Baseline Interaction",
      col.names = c("", "Estimate", "Std. Error", "df", "t", "p"),
      label="proteinDataEstimates",
      booktabs=TRUE,
      digits=2,
      linesep = "",
      position="hbt") %>%
  column_spec(1, border_right = TRUE)
sink()


# LS Means
emm <- emmeans(mixed.model.baselines, pairwise ~ Treatment)


# Table 3.9 - LS Means
sink("report/tables/ch3/proteinDataMeans.tex", type="output")
emm %>% rbind(emm$contrasts) %>%
  kbl(format="latex",
```

```
      caption="LS Means for Mixed Model on Protein Data",
      label="proteinDataLSMeans",
      col.names = c("Sequence", "Difference", "Adj. Mean", "SE", "df",
                    "Lower CI", "Upper CI"),
      booktabs = TRUE,
      digits=2,
      position="hbt") %>%
  column_spec(2, border_right = TRUE)
sink()
```