A MODEL FOR THE ANALYSIS OF CARRYOVER EFFECTS IN A THREE-PERIOD, THREE-TREATMENT CROSSOVER TRIAL FOR ASTHMA PATIENTS

A Thesis in
Public Health Sciences
by
Ruchita Dholakia

© 2013 Ruchita Dholakia

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Master of Science
August 2013
The thesis of Ruchita Dholakia was reviewed and approved* by the following

David Mauger
Professor, Division of Biostatistics and Bioinformatics
Public Health Sciences

Jeff Yanosky
Assistant Professor, Division of Epidemiology
Public Health Sciences

Vernon Chinchilli
Professor of Public Health Sciences
Chair of Public Health Sciences

*Signatures are on file in the Graduate School Office
Abstract

Background: A crossover trial design is a repeated measures assignment such that every subject is assigned different treatments in different time periods. This design is more efficient as each subject acts as his own control and one major criticism of this trial is a carryover effect. A carryover effect is defined as the effect of the treatment from the previous time period on the response at the current time period. A crossover trial design is appropriate when the carryover effect can be assumed to be negligible or the trial has incorporated a sufficient duration of washout period to reduce any carryover bias. This study evaluates the presence of a carryover effect as well as the difference in residual effects for different treatments in a three x three crossover trial with a pseudo-washout period. The trial extended over forty-eight weeks with sixteen weeks in each treatment arm. The first four weeks of each sixteen week period was considered as washout.

Objectives: This study develops a model to analyze the carryover effect and its decay in a three-period crossover trial comparing the responses of asthmatic kids to three add-on therapies (BADGER trial). The primary source of data for this analysis was the daily diary maintained by the subjects in the trial. The study attempts to distinguish the decay pattern for the various therapies to establish the need/appropriateness of the washout period.

Methods: The analysis utilizes data collected for the BADGER trial. The BADGER trial recruited patients, 6 to 17 years of age between March 2007 and July 2008, at Childhood Asthma Research and Education (CARE) Network centers. Analysis of Variance comparing summated outcomes for drug sequences was used for a preliminary test of carryover. Non-parametric and semi-parametric survival analysis was used to compare the carryover effects for different combinations of current and previous treatments. The primary outcome variables used included asthma control days (ACD), symptom free days (SFD), rescue free days (RFD), % change in FEV levels, and proportion of differential responders.
**Results:** A preliminary analysis to check for carryover was not significant (p-value =0.30). A Kaplan-Meier analysis for ACD during the washout period showed a significant carryover for LABA (p-value = 0.02) when it was followed by ICS and for LTRA (p-value = 0.02) when it was followed by LABA. Both of these are true under the assumption that in the first period a subject does not have a residual effect of any previous treatment. As expected, the change in FEV from the baseline to the end of the washout for the various groups was not significant for ICS (p-value = 0.0845), or LABA (p-value = 0.8404), or LTRA (p-value=0.7263). When compared for the relative possibility of a differential response, both the 12-week and 16-week data gave similar results.

**Conclusion:** The trial did not show a strong evidence for a carryover effect, though we have some statistical evidence to show that the carryover to the next period also depends on the treatment in the next period. Since, there is no carryover, the washout period data can be used for analysis with appropriate consideration for the risk of carryover bias.
Contents

List of Tables ...................................................................................................................... vi
List of Figures ..................................................................................................................... vii
Acknowledgements ........................................................................................................... viii
Chapter 1 ............................................................................................................................. 1
Introduction ......................................................................................................................... 1
Chapter 2 ............................................................................................................................. 3
Methods ............................................................................................................................... 3
   2.1 Outcome Measures ...................................................................................................... 3
   2.2 Statistical Analysis ................................................................................................... 4
Chapter 3 ............................................................................................................................. 7
Results .................................................................................................................................. 7
   3.1 Preliminary analysis to determine existence of a carryover effect ......................... 7
   3.2 Time to event modeling using Kaplan - Meier curves ............................................. 7
   3.3 Cox PH model for hazard rates ............................................................................... 20
   3.4 Percentage change in FEV comparison using ANOVA ........................................ 21
   3.5 Differential responders to the three therapies, a comparison of results of 12-week to 16-week period ......................................................................................... 22
Chapter 4 ............................................................................................................................. 26
Discussion ........................................................................................................................... 26
Chapter 5 ............................................................................................................................. 28
Conclusion ......................................................................................................................... 28
Bibliography ...................................................................................................................... 29
List of Tables

Table 2.1 Current and previous group combinations used .......................................................... 5
Table 3.1 ANOVA output for the preliminary analysis ................................................................ 7
Table 3.2.1: Log-rank and Wilcoxon tests for ICS as the current treatment (ACD analysis) ........ 8
Table 3.2.1.1 : Log-rank and Wilcoxon tests for ICS_LABA and ICS_None ................................. 9
Table 3.2.1.2 : Log-rank and Wilcoxon tests for ICS_LTRA and ICS_None ............................. 10
Table 3.2.2.1: Log-rank and Wilcoxon tests for LABA as the current treatment ......................... 12
Table 3.2.3.1 : Log-rank and Wilcoxon tests for LABA_LTRA vs. LABA_None ....................... 21
Table 3.2.4.1 : Log-rank and Wilcoxon tests for ICS as the current treatment ......................... 15
Table 3.2.5.1 : Log-rank and Wilcoxon tests for LABA as the current treatment ...................... 16
Table 3.2.6.1 : Log-rank and Wilcoxon tests for LTRA as the current treatment ....................... 17
Table 3.2.7.1 : Log-rank and Wilcoxon tests for ICS as the current treatment ......................... 18
Table 3.2.8.1 : Log-rank and Wilcoxon tests for LABA as the current treatment ...................... 19
Table 3.2.9.1 : Log-rank and Wilcoxon tests for ICS as the current treatment ......................... 20
Table 3.3.1: Hazard rate for ICS_LABA vs. ICS_None .................................................................. 21
Table 3.3.2: Hazard rate for LABA_LTRA vs. LABA_None....................................................... 21
Table 3.4.1 ANOVA output for change in FEV with ICS as current treatment ......................... 21
Table 3.4.2 ANOVA output for change in FEV with LABA as current treatment ..................... 22
Table 3.4.3 ANOVA output for change in FEV with LTRA as current treatment ...................... 22
List of Figures

Figure 3.2.1.1: KM curve comparison for ICS as the current treatment (ACD analysis) .............. 8
Figure 3.2.1.2: KM curve comparison for ICS_LABA and ICS_None........................................ 9
Figure 3.2.1.3: KM curve comparison for ICS_LTRA and ICS_None...................................... 10
Figure 3.2.2.1 : KM curve comparison for LABA as current treatment (ACD analysis).............. 11
Figure 3.2.2.2: KM curve comparison for LABA_LTRA and LABA_ICS .............................. 12
Table 3.2.2.2 : Log-rank and Wilcoxon tests for ICS_LTRA and ICS_None.............................. 13
Figure 3.2.3.1: KM curve comparison for LTRA as the current treatment (ACD analysis) ........ 14
Figure 3.2.4.1: KM curve comparison for ICS as the current treatment (SFD analysis)............. 15
Figure 3.2.5.1 KM curve for LABA as the current treatment (SFD analysis)............................. 16
Figure 3.2.6.1 KM curve for LTRA as the current treatment (SFD analysis) ............................... 17
Figure 3.2.7.1 KM curve for ICS as the current treatment (RFD analysis) ............................... 18
Figure 3.2.8.1 KM curve for LABA as the current treatment (RFD analysis)............................ 19
Figure 3.2.9.1 KM curve for LTRA as the current treatment (RFD analysis) ............................. 20
Figure 3.5.1 Differential response analysis for ICS and LABA for 12 and 16 week data .......... 23
Figure 3.5.2 Differential response analysis for ICS and LTRA for 12 and 16 week data .......... 24
Figure 3.5.3 Differential response analysis for LABA and LTRA for 12 and 16 week data ......... 25
Acknowledgements

I would like to express my deep appreciation to my thesis advisor, Dr. David Mauger, who continually challenged and supported me throughout this research. He allowed me to use his extensive knowledge of statistical analysis of crossover data, and without his guidance and help this thesis would not have been possible. I would also like to thank my committee members, Dr. Jeff Yanosky and Dr. Vernon Chinchilli. A special thanks to Dr. Li Wang, for her support and guidance and to Dr. Douglas Leslie for his mentorship. Lastly, unending thanks to my family, without them, this would not have been possible.
Chapter 1

Introduction

In a crossover trial design, all the treatments are given to all the patients, but not all at the same time. Each treatment is giving at a different time and the order is random. Each patient is randomized to a sequence of treatments. A crossover trial is characterized by the crossover of patients from one treatment to another during the course of the trial.

The main advantage of this repeated-measures design is that each patient acts as his own control. A crossover design is more efficient than a parallel design as it can give greater power for the same number of subjects.

When applied incorrectly, a crossover design may lead to incorrect results, mainly because of the residual treatment effect. One key criticism of the crossover design is the carryover effect. A carryover effect is defined as the effect of the treatment from the previous time period on the response at the current time period. It is primarily due to this reason that FDA views crossover designs unfavorably.\(^1\) Quite a few authors have discussed the issue of carryover effects in two-period crossover trials\(^2\)\(^-\)\(^4\). They have pointed out that unless carryover effects are negligible, it is better to employ a parallel design, or, in case of a crossover design, the results of only the first period should be used. However, Willan and Peter (1986)\(^5\) have argued that for a parallel design to be preferable, the carryover effect should be substantial which is unlikely to exist in most cases. In cases even where the carryover effect is insignificant, it is shown to interact with time.\(^6\) To account for a carryover effect, many crossover trials incorporate a washout (no treatments received) period between treatments. This period helps to account for any changes brought about in the subject due to the previous treatment. A well-designed crossover trial will include a washout period and always
include baseline measurements of important variables. However, it is sometimes the case that a true washout period is unethical because it leaves a sick patient untreated. In such cases, one approach is to employ as pseudo-washout during which time an active treatment is given, but the data during that time is not used in the analysis. In crossover trials, baseline data measurements become important to improve treatment effect estimation.\textsuperscript{7}

Crossover trials are appropriate when they are used for treatments that alleviate the condition rather than completely cure the patient.\textsuperscript{8} Thus, these trials are most appropriate for chronic conditions like asthma where the disease returns after the patient is taken off the treatment.

Asthma is a chronic condition and it affects about 25.9 million people of which roughly 7.1 million are children.\textsuperscript{9} It is a widespread condition and it accounts for about 2 million emergency department visits in one year\textsuperscript{10} and greater than 15 million outpatient visits\textsuperscript{11,12}. It is one of the most visible chronic conditions seen in children. On an average, 1 out of every 10 school-aged kids suffers from asthma\textsuperscript{13}. In order to get an idea of how serious the condition is, as per an estimate 10.5 million school-days are missed annually due to asthma.\textsuperscript{14}

This study employs data from a clinical trial comparing the best step-up therapy for asthma. Data from the Best ADd-on Therapy Giving Effective Responses (BADGER) trial\textsuperscript{15} was used for the analysis. The main objective of the trial was to compare the response of participants on low dose inhaled corticosteroids (ICS) to either doubling the dose of ICS or adding a long acting beta agonist (LABA) or adding a leukotriene receptor antagonist (LTRA). This was a double blind 3-period, 3-treatment crossover trial. After the run-in period, the trial lasted for 48 weeks with 16 weeks in each treatment arm. The first four weeks of each treatment arm was considered the washout period, in the sense that the data from the first four weeks was not considered for analysis. The trial concluded that LABA step-up was significantly more likely to provide the best response than either ICS or LTRA step-up.
Chapter 2

Methods

Data from the Best ADd-on Therapy Giving Effective Responses (BADGER) trial was used for the analysis. The trial had 182 participants of which 157 completed the three treatment arms. The data for the patients was collected at baseline, at monthly visits, and also in a daily diary to be maintained by patients. The self-reported diary contained data for rescue albuterol use, AM and PM peak flows, wheezing and coughing symptoms, unscheduled provider visits for asthma, and school/work absenteeism due to asthma. This daily diary data formed the primary data for analysis. If the diary did not have information for four of the above mentioned fields, then the data for that day was not used for analysis. In addition, the data from the monthly visit file was also used. Further refinements included excluding patients that had more than 10 missing daily diary days and more than two missing monthly visits. A final sample of 167 patients was used for analysis. Overall, I had eligible datapoints for 98% of the days during the washout period.

2.1 Outcome Measures

The outcome measures used for the analysis included Asthma Control Days (ACD), Symptom Free Day (SFD), Rescue Free Day (RFD), and proportion of differential responders during the washout period for each treatment transition. Another outcome variable used was the change in Forced Expiratory Volume (FEV) from the end of the run-in period to the end of the washout period.

An ACD was defined as a day in the daily diary without nighttime albuterol rescue use, morning and evening peak flows greater than 80% of the reference value, wheezing or coughing symptoms, unscheduled asthma-related visits to a healthcare provider, or absence from school or work due to
asthma. An RFD is a day without rescue albuterol use, either at night or before exercise. An SFD is a day without coughing or wheezing symptoms, no use of rescue albuterol in the night, no unscheduled visits to a provider for asthma, and no absenteeism from school/work due to asthma.

Annualized Asthma Control Days (AACD) was used in the preliminary analysis for determining carryover effect. AACD was calculated as the sum of ACDs in the period divided by the days in the period and multiplied by 365.

A subject was identified as a differential responder in a two treatment comparison if the number of annualized ACDs (AACD) differed by more than 31 days or if the percentage change in FEV from the baseline to the end of the treatment period was greater than 5%. This analysis was carried out only on the data for the washout period.

2.2 Statistical Analysis

The main aim of the thesis was to find the existence of carryover effect in the 3X3 crossover trial and model its decay. The preliminary analysis focused on the hypothesis that the carryover effect was similar for each drug sequence expanding the model proposed by Grizzle (1965)\textsuperscript{16,17} for a 2X2 crossover trial.

\[ W_{ijk} = \mu + b_{ij} + p_k + \phi_m + \lambda_m + \varepsilon_{ijk} \]

Where \( i = \) sequence, \( j = \) patient, \( k = \) period and \( m = \) treatment

\( \mu \) : overall mean

\( b_{ij} \) : effect of jth patient with ith sequence , \( \sim N(0, \sigma_b^2) \)

\( p_k \) : effect of kth period \( \phi \)

\( m \) : direct effect of mth drug treatment

\( \lambda_m \) : carryover effect of mth drug sequence
\( \varepsilon_{ijk} \) : random error, \( \sim N(0, \sigma^2) \)

Based on this, the null hypothesis being tested was:

\( H_0: \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 \) (based on the 6 randomized drug sequences)

The preliminary outcome (W) used was the sum of AACDs for each drug sequence was and the sequences were compared by means of analysis of variance. Pairwise comparisons between the drug sequences would be used in case the results of the analysis of variance were significant.

While Grizzle’s approach provided an intuitive approach to estimation of initial carryover, it cannot be used to arrive at final conclusions. Arriving at non-significant results for the preliminary test for carryover is not the same as concluding it is zero.\(^1\)

To further probe if the carryover effect is indeed zero or non-existent and if each treatment showed similar carryover effects irrespective of the treatment following it, we made 9 groups specifying the current_previous treatment combinations. These unique group combinations are created under the assumption that when the subject starts the trial in period 1, they have no carryover effect from any previous treatment.

<table>
<thead>
<tr>
<th>Current_previous</th>
<th>ICS_None</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICS_LABA</td>
<td></td>
</tr>
<tr>
<td>ICS_LTRA</td>
<td></td>
</tr>
<tr>
<td>LABA_None</td>
<td></td>
</tr>
<tr>
<td>LABA_ICS</td>
<td></td>
</tr>
<tr>
<td>LABA_LTRA</td>
<td></td>
</tr>
<tr>
<td>LTRA_None</td>
<td></td>
</tr>
<tr>
<td>LTRA_ICS</td>
<td></td>
</tr>
<tr>
<td>LTRA_LABA</td>
<td></td>
</tr>
</tbody>
</table>
The time to the first adverse event related to asthma for the current_previous treatment combinations was compared by means of the Kaplan Meier curves and the Log-rank and Wilcoxon tests. An adverse event was defined as either a non-Asthma Control Day (non ACD), or a non-Symptom Free Day (non SFD), or a non-Rescue Free Day (non RFD). A semi-parametric Cox proportional hazards model was also used to compare the groups.

The percentage change in FEV comparison was also carried out for the same current_previous combinations. The current_previous combinations were compared by means of an analysis of variance. An individual was identified as a differential responder if their response to two treatments differed by more than 31 annualized non-asthma control days or their percentage change in FEV from the baseline was greater than 0.05. To compare the treatments based on proportion of differential responders, a risk ratio model was used. This helped determine the probability of favorable responders to one treatment as compared to the other. All data analysis was performed using SAS version 9.3.
Chapter 3

Results

3.1 Preliminary analysis to determine existence of a carryover effect

The preliminary analysis results for determining if the carryover effect was equal for all the six drug sequences is shown in table 3.1. The results of the ANOVA are insignificant at a significance level of 5% with F value equal to 1.22 (p-value = 0.2998). Thus we fail to reject the null hypothesis and accordingly fail to conclude that the carryover effect is equal for each drug sequence.

Table 3.1 ANOVA output for the preliminary analysis

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>387775.43</td>
<td>77555.09</td>
<td>1.22</td>
<td>0.2998</td>
</tr>
<tr>
<td>Error</td>
<td>167</td>
<td>11088237.37</td>
<td>63361.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>172</td>
<td>11476012.80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The insignificant results for the analysis do not mean the carryover effect is zero and hence we next move onto a Kaplan Meier curve comparison of the groups.

3.2 Time to event modeling using Kaplan - Meier curves

Kaplan-Meier is a non-parametric estimation of survival rates and some of the key assumptions regarding the validity of the p-value are as follows:

I. The observations and independent and identically distributed.

II. The censoring distribution is independent of the treatment group. This assumption, however, is not valid in our case since our earlier data cleaning process took care of the censored observations.
3.2.1 Time to event for first non-Asthma control day (ACD) for ICS as the current treatment

For patients who were currently on inhaled corticosteroids (ICS), there were three possibilities, the subjects either had:

i. ICS as the current treatment and LABA as the previous treatment (long acting beta agonist)

ii. ICS as the current treatment and LTRA as the previous treatment (leukotriene receptor antagonist).

iii. ICS as the current treatment and no previous treatment (i.e., ICS as the treatment in period 1).

Figure 3.2.1.1: KM curve comparison for ICS as the current treatment (ACD analysis)

Table 3.2.1.1: Log-rank and Wilcoxon tests for ICS as the current treatment (ACD analysis)

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>5.8115</td>
<td>2</td>
<td>0.0547</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>6.4685</td>
<td>2</td>
<td>0.0394</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>10.1858</td>
<td>2</td>
<td>0.0061</td>
</tr>
</tbody>
</table>
The KM curve comparison from figure 3.2.1.1 shows the curves to be similar in shape and the curve for the group ICS_None never overlaps or interests the other two curves. Also, all the three treatment combinations seem to be similar after about 3 weeks in the washout period. The Log-rank test for equality between curves is marginally insignificant (p-value = 0.055) and the Wilcoxon test is marginally significant (p-value = 0.039). This points to the fact observed in the KM curve analysis that the groups are more different at the beginning rather than towards the end.

To further look at a pairwise comparison between the groups, we repeat the KM curve analysis for i) ICS_None and ICS_LABA and, ii) ICS_None and ICS_LTRA. The results are shown in Figures 3.2.1.2 & 3.2.1.3 and Table 3.2.1.2 and 3.2.1.3

Figure 3.2.1.2: KM curve comparison for ICS_LABA and ICS_None
Table 3.2.1.2: Log-rank and Wilcoxon tests for ICS_LABA and ICS_None

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>2.5372</td>
<td>1</td>
<td>0.1112</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>2.8196</td>
<td>1</td>
<td>0.0931</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>5.7237</td>
<td>1</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

Figure 3.2.1.3: KM curve comparison for ICS_LTRA and ICS_None

Table 3.2.1.3: Log-rank and Wilcoxon tests for ICS_LTRA and ICS_None

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>5.5182</td>
<td>1</td>
<td>0.0188</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>6.3271</td>
<td>1</td>
<td>0.0119</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>9.7237</td>
<td>1</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

Kaplan-Meier curves show a statistically significant difference between the ICS_None and ICS_LABA group (p-value = 0.019 for log-rank test and p-value = 0.012 for Wilcoxon). Kaplan-
Meier curves showed no significant difference between the ICS_None and ICS_LTRA group (p-value = 0.111 for log-rank test and p-value = 0.093 for Wilcoxon). This points to the possibility of a LABA residual when ICS is the current treatment.

3.2.2 Time to event for first non-Asthma control day (ACD) for LABA as the current treatment

For patients who were currently on long acting beta agonist (LABA), there were three possibilities, the subjects either had:

i. LABA as the current treatment and ICS as the previous treatment (inhaled corticosteroids).

ii. LABA as the current treatment and LTRA as the previous treatment (leukotriene receptor antagonist).

iii. LABA as the current treatment and no previous treatment (i.e., LABA as the treatment in period 1).

Figure 3.2.2.1 : KM curve comparison for LABA as current treatment (ACD analysis)
The KM curve comparison from figure 3.2.2.1 shows the curve for the group LABA_None and LABA_ICS to be similar and overlapping. However, the curve for the LABA_LTRA group seems to be distinctly different. The Log-rank test for equality between curves is marginally insignificant (p-value = 0.064) and the Wilcoxon test is marginally significant (p-value = 0.045) as shown in table 3.2.2.1. The specific comparison between LABA_None and LABA_LTRA is highly significant (p-value = 0.017 for log-rank and 0.022 for Wilcoxon).

Table 3.2.2.1: Log-rank and Wilcoxon tests for LABA as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>5.4874</td>
<td>2</td>
<td>0.0643</td>
<td></td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>6.2086</td>
<td>2</td>
<td>0.0449</td>
<td></td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>6.8926</td>
<td>2</td>
<td>0.0319</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.2.2.2: KM curve comparison for LABA_LTRA and LABA_ICS
### Table 3.2.2.2: Log-rank and Wilcoxon tests for ICS_LTRA and ICS_None

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>5.6979</td>
<td>1</td>
<td>0.0170</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>5.2491</td>
<td>1</td>
<td>0.0220</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>6.5359</td>
<td>1</td>
<td>0.0106</td>
</tr>
</tbody>
</table>

#### 3.2.3 Time to event for first non Asthma control day (ACD) for LTRA as the current treatment

For patients who were currently on leukotriene receptor antagonist (LTRA) there were three possibilities, the subjects either had:

1. LTRA as the current treatment and LABA as the previous treatment (long acting beta agonist)
2. LTRA as the current treatment and ICS as the previous treatment
3. LTRA as the current treatment and no previous treatment (i.e., LTRA as the treatment in period 1).

The results of the KM curve are shown in figure 3.2.3.1
Figure 3.2.3.1: KM curve comparison for LTRA as the current treatment (ACD analysis)

As figure 3.2.3.1 and table 3.2.3.1 show, Kaplan-Meier curves showed no significant difference between the groups when LTRA was the current treatment (p-value = 0.78 for log-rank and p-value = 0.50 for Wilcoxon).

Table 3.2.3.1: Log-rank and Wilcoxon tests for LTRA as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>0.4954</td>
<td>2</td>
<td></td>
<td>0.7806</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>1.3805</td>
<td>2</td>
<td></td>
<td>0.5014</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>1.4150</td>
<td>2</td>
<td></td>
<td>0.4929</td>
</tr>
</tbody>
</table>

3.2.4 Time to event for first non symptom free day (SFD) for ICS as the current treatment

The KM curve results were marginally insignificant for the groups when time to first non SFD was modeled for the ICS group as seen in figure 3.2.4.1. The p-value is marginally insignificant for log-rank test (p-value = 0.07) and marginally significant for Wilcoxon test (p-value = 0.04) as seen in Table 3.2.4.1. A pairwise comparison shows the KM curve results to be significant for ICS_LABA
and ICS_None (p-value = 0.048 for log-rank and p-value = 0.017 for Wilcoxon). This means, for SFD as well, an evidence of a LABA carryover is seen when ICS is the current treatment.

Figure 3.2.4.1: KM curve comparison for ICS as the current treatment (SFD analysis)

![KM curve comparison for ICS as the current treatment (SFD analysis)](image)

Table 3.2.4.1 : Log-rank and Wilcoxon tests for ICS as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>5.3103</td>
<td>2</td>
<td>0.0703</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>6.2747</td>
<td>2</td>
<td>0.0434</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>7.9070</td>
<td>2</td>
<td>0.0192</td>
</tr>
</tbody>
</table>

3.2.5 Time to event for first non Symptom free day (SFD) for LABA as the current treatment

As the KM curves in figure 3.2.5.1 show, the difference between the groups for SFD when LABA is the current treatment is non-significant. This is verified by the Log-rank and Wilcoxon tests, p-
values of 0.32 and 0.17 respectively. This implies there is no evidence of a carryover effect for this group and when the outcome event modeled is the time to first non-SFD.

Figure 3.2.5.1 KM curve for LABA as the current treatment (SFD analysis)

Table 3.2.5.1 : Log-rank and Wilcoxon tests for LABA as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt;Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>2.2776</td>
<td>2</td>
<td></td>
<td>0.3202</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>3.5549</td>
<td>2</td>
<td></td>
<td>0.1691</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>3.7944</td>
<td>2</td>
<td></td>
<td>0.1500</td>
</tr>
</tbody>
</table>

3.2.6 Time to event for first non Symptom free day (SFD) for LTRA as the current treatment

As the KM curves in figure 3.2.6.1 show, the difference between the groups for SFD when LTRA is the current treatment is non-significant. This is verified by the Log-rank and Wilcoxon tests, p-values of 0.269 and 0.096 respectively. This implies there is no evidence of a carryover effect for this group and when the outcome event modeled is the time to first non-SFD.
Figure 3.2.6.1 KM curve for LTRA as the current treatment (SFD analysis)

Table 3.2.6.1: Log-rank and Wilcoxon tests for LTRA as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>2.6231</td>
<td>2</td>
<td>0.2694</td>
<td></td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>4.6863</td>
<td>2</td>
<td>0.0960</td>
<td></td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>5.5042</td>
<td>2</td>
<td>0.0638</td>
<td></td>
</tr>
</tbody>
</table>

3.2.7 Time to event for first non- Rescue free day (RFD) for ICS as the current treatment

As the KM curves in figure 3.2.7.1 show, the difference between the groups for RFD when ICS is the current treatment is non-significant. This is verified by the Log-rank and Wilcoxon tests, p-values of 0.55 and 0.49 respectively. This implies there is no evidence of a carryover effect for this group and when the outcome event modeled is the time to first non-RFD.
Figure 3.2.7.1 KM curve for ICS as the current treatment (RFD analysis)

Table 3.2.7.1: Log-rank and Wilcoxon tests for ICS as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>1.1972</td>
<td>2</td>
<td></td>
<td>0.5496</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>1.4232</td>
<td>2</td>
<td></td>
<td>0.4909</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>1.8146</td>
<td>2</td>
<td></td>
<td>0.4036</td>
</tr>
</tbody>
</table>

3.2.8 Time to event for first non-Rescue free day (RFD) for LABA as the current treatment
As the KM curves in figure 3.2.8.1 show, the difference between the groups for RFD when LABA is the current treatment is non-significant. This is verified by the Log-rank and Wilcoxon tests, p-values of 0.33 and 0.71 respectively. This implies there is no evidence of a carryover effect for this group and when the outcome event modeled is the time to first non-RFD.
Figure 3.2.8.1 KM curve for LABA as the current treatment (RFD analysis)

![KM curve for LABA as the current treatment (RFD analysis)](image)

Table 3.2.8.1: Log-rank and Wilcoxon tests for LABA as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt; Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>2.2163</td>
<td>2</td>
<td>0.3302</td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>0.6892</td>
<td>2</td>
<td>0.7085</td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>2.0778</td>
<td>2</td>
<td>0.3538</td>
</tr>
</tbody>
</table>

3.2.9 Time to event for first non-Rescue free day (RFD) for LTRA as the current treatment

As the KM curves in figure 3.2.9.1 show, the difference between the groups for RFD when LABA is the current treatment is non-significant. This is verified by the Log-rank and Wilcoxon tests, p-values of 0.59 and 0.56 respectively. This implies there is no evidence of a carryover effect for this group and when the outcome event modeled is the time to first non-RFD.
Figure 3.2.9.1 KM curve for LTRA as the current treatment (RFD analysis)

![KM curve for LTRA as the current treatment (RFD analysis)](image)

Table 3.2.9.1: Log-rank and Wilcoxon tests for ICS as the current treatment

<table>
<thead>
<tr>
<th>Test of Equality over Strata</th>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr&gt;Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-Rank</td>
<td>1.0653</td>
<td>2</td>
<td>0.5870</td>
<td></td>
</tr>
<tr>
<td>Wilcoxon</td>
<td>1.1680</td>
<td>2</td>
<td>0.5577</td>
<td></td>
</tr>
<tr>
<td>-2Log(LR)</td>
<td>0.8760</td>
<td>2</td>
<td>0.6453</td>
<td></td>
</tr>
</tbody>
</table>

3.3 Cox PH model for hazard rates

The Cox PH model was repeated for the combinations were the KM curves showed a significant difference. Table 3.3.1 gives the output of the PH model for the comparison of ICS_LABA vs. ICS_None. We see that the hazard for a non-ACD reduces by 33% when ICS is preceded by LABA as against ICS being the treatment in the first period (p-value = 0.047).
Table 3.3.1: Hazard rate for ICS_LABA vs. ICS_None

<table>
<thead>
<tr>
<th>Parameter</th>
<th>D F</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
<th>Hazard Ratio</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>group_name</td>
<td>ICS LABA</td>
<td>-0.40152</td>
<td>0.20210</td>
<td>3.9469</td>
<td>0.0470</td>
<td>0.669</td>
<td>group_name ICS LABA</td>
</tr>
</tbody>
</table>

Table 3.3.2 gives the hazard rates for a model comparing LABA_LTRA to LABA_None. The results shows that the hazard(risk) for a non-ACD on cases where LABA was preceded by LTRA was significantly lower than when LABA was preceded by no treatment in the trial(LABA was treatment in period 1) as indicated by the hazard ratio of 0.669 (p-value = 0.047).

Table 3.3.2: Hazard rate for LABA_LTRA vs. LABA_None

<table>
<thead>
<tr>
<th>Parameter</th>
<th>D F</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
<th>Hazard Ratio</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>group_name</td>
<td>LABA LTRA</td>
<td>-0.40323</td>
<td>0.19931</td>
<td>4.0929</td>
<td>0.0431</td>
<td>0.668</td>
<td>group_name LABALTRA</td>
</tr>
</tbody>
</table>

3.4 Percentage change in FEV comparison using ANOVA

The analysis of variance was used to compare the change in forced expiratory volume (FEV) from the baseline for the current_previous treatment combinations for each current treatment.

Table 3.4.1 shows the output for the FEV comparison when ICS was the current treatment. The results are not statistically significant (p-value = 0.08) indicating there is no difference between the groups with regards to change in FEV.

Table 3.4.1 ANOVA output for change in FEV with ICS as current treatment

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Anova SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>group_name</td>
<td>2</td>
<td>0.05915803</td>
<td>0.02957902</td>
<td>2.51</td>
<td>0.0845</td>
</tr>
</tbody>
</table>
Table 3.4.2 shows the output for the FEV comparison when LABA was the current treatment. The result is highly insignificant (p-value = 0.84) indicating there is no difference between the groups with regards to change in FEV.

Table 3.4.2 ANOVA output for change in FEV with LABA as current treatment

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Anova SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>group_name</td>
<td>2</td>
<td>0.00514818</td>
<td>0.00257409</td>
<td>0.17</td>
<td>0.8404</td>
</tr>
</tbody>
</table>

Table 3.4.3 shows the output for the FEV comparison when LTRA was the current treatment. The result is highly insignificant (p-value = 0.73) indicating there is no difference between the groups with regards to change in FEV.

Table 3.4.3 ANOVA output for change in FEV with LTRA as current treatment

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Anova SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>group_name</td>
<td>2</td>
<td>0.00736851</td>
<td>0.00368425</td>
<td>0.32</td>
<td>0.7263</td>
</tr>
</tbody>
</table>

All the three ANOVA comparisons actually confirm the trial design that we had expected no change in FEV at the end of the washout period.

3.5 Differential responders to the three therapies, a comparison of results of 12-week to 16-week period

A differential response comparison of ICS to LABA for the 12 week period (accounting for washout) shows respondents behaving more favorably to LABA than to ICS. The probability of a favorable response to LABA is 2.2 times as compared to ICS (p-value = 0.008). Refer Figure 3.54.1. Similarly, a comparison of ICS to LABA for the 16 week period (ignoring carryover) shows respondents behaving more favorably to LABA than to ICS. The probability of a favorable response
to LABA is 2.5 times as compared to ICS (p-value = 0.002). Thus, the results for the 12-week and 16-week data are similar with CI overlap.

**Figure 3.5.1 Differential response analysis for ICS and LABA for 12 and 16 week data**

We repeat a similar analysis for ICS to LTRA comparison. The comparison of ICS to LTRA for the 12 week period (accounting for washout) shows respondents behaving more favorably to LTRA than to ICS. The probability of a favorable response to ICS is 0.8 times as compared to LTRA but the results are inconclusive (p-value = 0.5). Refer Figure 3.5.2. Similarly, a comparison of ICS to LTRA for the 16 week period (ignoring carryover) shows the probability of a favorable response to LTRA and ICS is the same (p-value = 1). Thus, the results for the 12-week and 16-week data are similar with CI overlap.
Figure 3.5.2 Differential response analysis for ICS and LTRA for 12 and 16 week data

Now, we look at the comparison of LTRA to LABA for the 12 week period (accounting for washout) shows respondents behaving more favorably to LABA than to LTRA. The probability of a favorable response to LABA is 2.8 times as compared to ICS (p-value <0.0001). Refer Figure 3.5.3. Similarly, a comparison of LTRA to LABA for the 16 week period (ignoring carryover) shows respondents behaving more favorably to LABA than to LTRA. The probability of a favorable response to LABA is 2.6 times as compared to ICS (p-value = 0.0014). Thus, the results for the 12-week and 16-week data are similar with CI overlap.
Figure 3.5.3 Differential response analysis for LABA and LTRA for 12 and 16 week data

Relative probability : 2.8  
(p-value = <0.0001)  

Relative probability : 2.6  
(p-value = 0.0014)
Chapter 4

Discussion

Our study of the daily diary maintained by patients in the BADGER trial showed no significant evidence of a carryover effect that lasted for the duration of the 4-week pseudo washout period that the original trial had incorporated. However, it did show trends of a selective carryover effect for specific current-previous treatment combinations. The initial preliminary analysis for carryover was performed using stage 1 of Grizzle’s procedure. Grizzle’s test has drawn a lot of review from other researchers like Jones and Lewis\textsuperscript{19}, and Brown\textsuperscript{3}. Most of them have inferred that the test is not the best and accurate way of measuring carryover. However, still Brown\textsuperscript{3} recommends using at least stage 1 of the procedure for preliminary analysis. We used the preliminary analysis as an additional measure to probe for the existence of a carryover effect. Alternate 3-period crossover designs have been proposed\textsuperscript{20} that facilitate the examination of a carryover effect by using differences between the subjects rather than sum as was used in the Grizzle procedure. However, this approach is designed for 2-period, 3-treatment trials which is not applicable in our case.

Although the preliminary analysis for carryover was insignificant, the KM curve comparison showed a statistically significant evidence of a LABA carryover when it was followed by ICS and a LTRA carryover when it was followed by LABA. Other studies\textsuperscript{21} have used KM curves to compare treatment groups but none have grouped them as current_previous combinations. This study uses data from only one trial and, thus the finding must be considered preliminary, these findings suggest that it would be worthwhile to look for carryover effect in the context of the Treatment1*Treatment2 interaction. Our findings suggest that the carryover effect of each treatment might be felt differently depending on the treatment following it.
It is recommended that a crossover design be used only when it is known that carryover effects will not be substantial enough to bias the results. Researchers prefer erring on the side of caution and usually incorporate a washout period in a crossover trial unless it has been proven that a carryover effect does not exist. However, as our study shows that even for specific treatments like LABA that showed a carryover effect, the effects of the residual did not last the entire duration of the washout period. Hence, maybe it was worthwhile not to incorporate the entire 4 weeks as washout period.

The original trial had compared the proportion of patients with a favorable response when comparing two treatments and hence when we repeated a similar approach with the data for the entire treatment period (16 week), the 12 week and 16 week data did not show any substantial difference. We found no difference between differential response that can be attributed to the inclusion or exclusion of a washout period. Whether the inclusion of a washout period is worth losing data that could have otherwise increased statistical power, will have to be an objective call best determined by the possible risks associated with the carryover bias.

It is fair to ask if this approach of comparing carryover effects for current_previous treatment combinations by means of Kaplan-Meier curves could be practically applied for analyzing carryover effect in other crossover trials. I have tried to restrict the analysis of treatment interaction to strictly statistical methods without bringing in any pharmacological considerations for the treatment interaction. One limitation of this approach is the assumption that a subject in period 1 is considered to be coming off no treatment. This approach was used in the absence of relevant data or a measure to incorporate the same.

These findings suggest that this KM analysis approach to identify carryover effects may indeed be viable. However, an extensive analysis of similar multi-period, multi-treatment studies is necessary to arrive at any conclusion.
Chapter 5

Conclusion

The analysis showed a carryover effect for LABA when it was followed by ICS and for LTRA when it was followed by LABA. We can conclude that ICS has little or no carryover effect and the effects of LTRA are independent of the preceding treatment. As the data for differential response days shows, the relative probabilities did not differ much between the entire trial period and only the last 12 weeks (all-washout).

We can conclude that the trial did not show a substantial amount of differential carryover effect, though some treatments showed a selective lingering effect. This means we could have used the entire 48 week data or shortened the trial by 12 weeks, saving time and money. However, any such decision would have to appropriately account for the risk of carryover bias.
Bibliography


