

# STOR 765 REPORT

## Evaluating the Impact of ABM Medication Synchronization Implementation on Medication Adherence

Consultant Client  
Cheng Qiu Alice Cheng

PROFESSOR: Steve Marron

September 19, 2020

### Abstract

The aim of this consulting project is to understand the relationship between Appointment-Based Model (ABM) core component implementation and pharmacy performance on medication adherence. ABM core component implementation is measured using survey data from 68 North Carolina community pharmacies and medication adherence (Proportion of Days Covered (PDC)  $\geq 0.80$ ) is measured using medication adherence rates (percentage of patients at each pharmacy that achieved adherence). Lasso regression and variable significance analysis were run to isolate the important drivers. Many significant drivers with different levels of impacts were found in all five ABM core components.

## 1 Introduction

Medication non-adherence is one of the most expensive avoidable healthcare expenditures in the United States. One strategy used to combat medication non-adherence by the community pharmacies is medication synchronization. Medication synchronization means scheduling medication refills to a common fill date in order to lower the number of pharmacy visits necessary for patients. The appointment based model (ABM) builds upon the basic definition of medical synchronization and adds on individualized pharmacist consultation. The five core components of ABM medication synchronization include the (1) identification and enrollment of patients, (2) medication review and patient assessment, (3) alignment of refills, (4) preparation of medications, and (5) delivery of medications in addition to other services. In this study, the client wants to examine the underlying drivers of medical adherence in community pharmacies; more specifically, the relationship between ABM core component implementation and medication adherence rate in community pharmacies.

## 2 Data Set

The data set was provided by the client via excel spread sheet files. The data set contained 68 samples (community pharmacies), 6 response variables (medication adherence rates) and 70 independent variables (survey question answers). The adherence rates were measured by the system software over a 6 months period before and after the survey for each of the three chronic disease groups (Anti-hypertensive RASA agents, Antihyperlipidemic statins, and Oral hypoglycemic agents). There were a total of 36 questions on the survey; 11 of them were pharmacy demographics questions and the rest were ABM core component implementation assessment questions. The survey answers were well documented and a data dictionary was provided.

Four of the pharmacies were thrown out because they contained null values (Incomplete surveys). Question 7F was also thrown out because every pharmacy gave the exact same answer. The encoded data set was reshaped in R for easier manipulation and modeling. The final data set contained 384 samples (64 [pharmacies]\*3 [diseases]\*2[pre and post]), 1 response variable (adherence rate), and 66 independent variables (e.g. self-assessment tool questions). Three disease subsets were created because client was also curious about the relationship between adherence rate and chronic diseases.

## 3 Exploratory Data Analysis

The purpose of exploratory data analysis is to visualize the data set and discover the underlying quirks. We first plotted the adherence rates using a box and whisker plot (Figure 1). In a box and whisker plot, the box represents the 1st quartile, the median and the 3rd quartile of the data set. The whiskers represent the min and the max. The dots represent all the outliers. We immediately noticed that there are multiple clear outliers. After consulting with the client, the outlier in the Diabetes group was deleted and other outliers were kept.

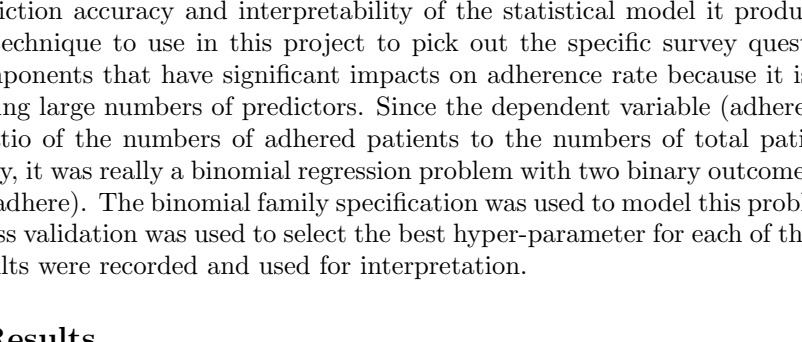


Figure 1: The box plots showed that there are some variations in adherence score among disease types.

We plotted the adherence rates again with the 95% CI for the mean (Figure 2). The blue and red box plots correspond with the pre survey adherence rates and the post survey adherence rates. The line segment inside the box plot represents the 95% CI for the mean. Since the CI overlaps each other, the plots do not show statistically significant improvements in adherence rates post survey.

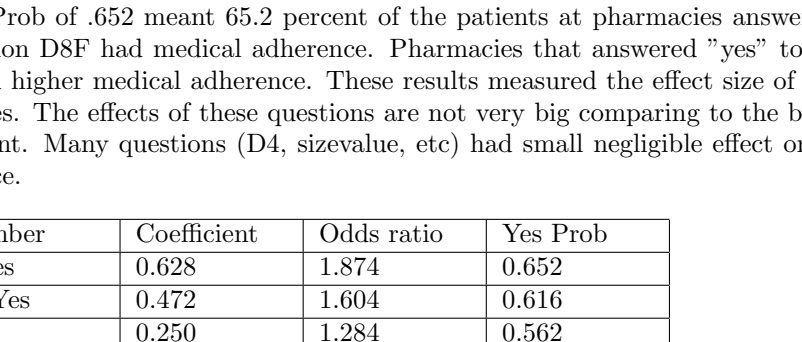


Figure 2: Box plots with 95 percent CI for the mean showed small improvements in adherence rate post survey. Its statistical significance will be studied in depth in Section 5.

## 4 Lasso Regression

### 4.1 Introductions

Lasso (least absolute shrinkage and selection operator) regression is a machine learning technique that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the statistical model it produces. It is a good technique to use in this project to pick out the specific survey questions and core components that have significant impacts on adherence rate because it is efficient in handling large numbers of predictors. Since the dependent variable (adherence rate) was a ratio of the numbers of adhered patients to the numbers of total patients in a pharmacy, it was really a binomial regression problem with two binary outcomes (adhere and not adhere). The binomial family specification was used to model this problem. Five folds cross validation was used to select the best hyper-parameter for each of the models. The results were recorded and used for interpretation.

### 4.2 Results

The report will focus on the results from the comprehensive model including all three diseases. Out of the 66 independent variables, 48 of them were selected by the lasso regression. The results were divided into 2 tables (positive impacts and negative impacts) and they were listed in the order of coefficient magnitude.

The coefficient of each question selected gave the approximate magnitude and direction of the impact. The coefficients were then converted into odds ratios and probabilities for easier interpretation. Take question D8FYes as an example, Odds ratio of 1.874 meant patients at pharmacies that answered "yes" to the question have an 87.4 percent increase in the odds of adhering comparing to patients at pharmacies that answered "no" to the question. In other words, the odds of patients adhering vs patients not adhering at pharmacies that answered "yes" was 1.874 : 1.

Yes Prob of .652 meant 65.2 percent of the patients at pharmacies answered "yes" to question D8F had medical adherence. Pharmacies that answered "yes" to question D8F had higher medical adherence. These results measured the effect size of the questionnaires. The effects of these questions are not very big comparing to the baseline of 50 percent. Many questions (D4, sizevalue, etc) had small negligible effect on medical adherence.

Q. Number	Coefficient	Odds ratio	Yes Prob
D8F-Yes	0.628	1.874	0.652
Q18C-Yes	0.472	1.604	0.616
Q2-3	0.250	1.284	0.562
Q4-3	0.212	1.236	0.553
Q23F-Yes	0.189	1.208	0.547
Q12A-Yes	0.173	1.189	0.543
Q7E-Yes	0.168	1.182	0.542
Q16-5	0.159	1.172	0.540
Q1-Yes	0.152	1.164	0.538
Q5-1	0.148	1.160	0.537
D8A-Yes	0.146	1.158	0.537
D3-3	0.144	1.155	0.536
DiseaseRASA	0.137	1.147	0.534
Q23DYes	0.124	1.132	0.531
D8BYes	0.121	1.129	0.530
D531-50	0.120	1.127	0.530
Q18HYes	0.112	1.118	0.528
D516-30	0.087	1.091	0.522
Q7D-Yes	0.082	1.086	0.521
Q17-1	0.075	1.078	0.519
Q22-Yes	0.068	1.070	0.517
D6-Yes	0.066	1.068	0.517
D10-Yes	0.057	1.059	0.514
D8D-Yes	0.043	1.044	0.511
Q3-1	0.042	1.043	0.510
Q16-7	0.034	1.035	0.509
Q19-Yes	0.028	1.028	0.507
Q24-3	0.018	1.018	0.504
Q20-Yes	0.001	1.001	0.500
Q23B-Yes	0.001	1.001	0.500
sizevalue	0.000	1.000	0.500
D4	0.000	1.000	0.500

Table 1: Questions with positive impacts on adherence rate.

Q. Number	Coefficient	Odd ratio	Probability
D3-5	-0.497	0.608	0.378
D7-Unsure	-0.458	0.633	0.388
Q10-Yes	-0.407	0.666	0.400
Q17-2	-0.399	0.671	0.402
Q18E-Yes	-0.351	0.704	0.413
Q12C-Yes	-0.321	0.725	0.420
D8C-Yes	-0.293	0.746	0.427
Q23G-Yes	-0.290	0.749	0.428
Q18B-Yes	-0.269	0.764	0.433
Q12B-Yes	-0.188	0.829	0.453
Q23A-Yes	-0.170	0.844	0.458
Q13-Yes	-0.150	0.861	0.463
Q5-2	-0.106	0.900	0.474
PrePostPre	-0.095	0.909	0.476
Q24-2	-0.074	0.929	0.482
Q7C-Yes	-0.063	0.939	0.484
Q23E-Yes	-0.061	0.940	0.485
Q16-4	-0.057	0.944	0.486
Q4-1	-0.055	0.946	0.486
D9-Yes	-0.052	0.949	0.487
Q16-3	-0.023	0.978	0.494
DiseaseDiabetes	-0.021	0.979	0.495
Q16-6	-0.019	0.981	0.495
D1	-0.013	0.987	0.497
Q18D-Yes	-0.010	0.990	0.497
Q12-DYes	-0.008	0.992	0.498
Q11-1	-0.003	0.997	0.499
Q25-2	-0.001	0.999	0.500

Table 2: Questions with negative impacts on adherence rate.

### 4.3 Discussion

The lasso model gave us some very interesting results. 48 questions from every core component were chosen, indicating that all five core components had various effects on the medication adherence rate. We will take a closer look at couple questions I thought were interesting. Q10 asked about whether pharmacy requires new prescription to synchronize medication refill and it has a coefficient of -.407; this meant small mundane tasks like getting a new prescription from the doctor had big negative impacts on medication adherence. Q12 asked about factors pharmacy considers when scheduling medication synchronization dates; Q12A (Patient Convenience) has a positive coefficient of .173 and Q12C (Pharmacy Convenience) has a negative coefficient of -.321. Patients are more likely to adhere at pharmacies that prioritize patient convenience and patients are less likely to adhere at pharmacies that prioritize workflow.

## 5 Binomial Regression Variable Significance Analysis

### 5.1 Introduction

Lasso regression in the previous section provided variable selection and provided the effect sizes. In layman's terms, the model picked out survey questions that had an impact on medical adherence and gave the magnitude of each impact. The goal of this section is to find out whether the results were statistically significant. Normally, two way analysis of variance could be deployed here to examine the influences of survey questions on adherence rate. Since adherence rates were already transformed into binary outcomes, two way ANOVA was no longer appropriate. Likelihood ratio tests for generalized linear models (GLM) could be used as an alternate to test the influence of the categorical variables on the binary outcomes. In addition, Benjamin-Hochberg correction was used to adjust the p-values and limit the effects of multiple comparisons. Five out of the total 66 survey questions were ignored because the survey answers are continuous. The analysis was ran 61 times using the 61 eligible survey questions on the main data set.

### 5.2 Results

The analysis sequentially compared the smallest model with the next more complex model by adding one additional variable (Survey Question and Disease) in each step. The comparison was done using the likelihood ratio test and the most parsimonious model is chosen. Take question D3 as example, the models been compared were:

$$Adherence = \beta_0$$

$$Adherence = \beta_0 + \beta_1 * D3$$

$$Adherence = \beta_0 + \beta_1 * D3 + \beta_2 * Disease$$

$$Adherence = \beta_0 + \beta_1 * D3 + \beta_2 * Disease + \beta_3 * (Disease * D3)$$

The three p-values corresponded with the 3 variables and whether they had significant impact on adherence. The table above showed the original p-values and the B-H adjusted p-values for the question term (D3) and the interaction term (D3\*Disease). The Adjusted p-values of 0 and .386 in the first column showed us question D3 had significant influences on adherence but the interaction did not at the level of significance  $\alpha = .05$ .

The results from the Likelihood Ratio Tests are listed below. The Disease variable column is not shown because it is significant with every question.

Question	Question p-value	BH Adjusted	Interaction term p-value	BH Adjusted	Chosen by Lasso
D3	0.000	0.000	0.076	0.386	Yes
D5	0.000	0.000	0.010	0.168	Yes
D6	0.002	0.004	0.191	0.564	Yes
D7	0.000	0.000	0.152	0.545	Yes
D8A	0.038	0.053	0.157	0.545	Yes
D8B	0.490	0.530	0.530	0.794	Yes
D8C	0.000	0.000	0.459	0.738	Yes
D8D	0.279	0.312	0.007	0.168	Yes
D8E	0.263	0.299	0.380	0.694	No
D8F	0.000	0.000	0.347	0.673	Yes
D9	0.000	0.000	0.422	0.715	Yes
D10	0.000	0.000	0.401	0.696	Yes
D11A	0.006	0.008	0.326	0.655	No
D11B	0.001	0.002	0.003	0.168	No
D11C	0.467	0.514	0.560	0.794	No
Q1	0.001	0.002	0.042	0.338	Yes
Q2	0.002	0.003	0.180	0.564	Yes
Q3	0.001	0.002	0.097	0.401	Yes
Q4	0.000	0.001	0.117	0.453	Yes
Q5	0.000	0.000	0.586	0.794	Yes
Q6	0.870	0.888	0.458	0.738	No
Q7A	0.051	0.065	0.071	0.386	No
Q7B	0.012	0.018	0.650	0.794	No
Q7C	0.178	0.210	0.657	0.794	Yes
Q7D	0.000	0.000	0.683	0.804	Yes
Q7E	0.000	0.000	0.056	0.338	Yes
Q8	0.000	0.000	0.798	0.878	No
Q10	0.000	0.000	0.083	0.391	Yes
Q11	0.001	0.001	0.368	0.694	Yes
Q12A	0.000	0.000	0.645	0.794	Yes
Q12B	0.004	0.007	0.973	0.973	Yes
Q12C	0.001	0.002	0.289	0.655	Yes
Q12D	0.875	0.888	0.166	0.549	Yes
Q13	0.045	0.060	0.615	0.794	Yes
Q14	0.545	0.580	0.321	0.655	No
Q15	0.000	0.000	0.285	0.655	No
Q16	0.000	0.000	0.525	0.794	Yes
Q17	0.000	0.000	0.089	0.391	Yes
Q18A	0.126	0.155	0.025	0.334	No
Q18B	0.000	0.000	0.009	0.168	Yes
Q18C	0.000	0.000	0.056	0.338	Yes
Q18D	0.052	0.065	0.227	0.605	Yes
Q18E	0.000	0.000	0.595	0.794	Yes
Q18F	0.032	0.045	0.055	0.338	No
Q18G	0.000	0.000	0.033	0.338	No
Q18H	0.049	0.065	0.327	0.655	Yes
Q19	0.000	0.000	0.600	0.794	Yes
Q20	0.977	0.977	0.389	0.694	Yes
Q21	0.209	0.241	0.297	0.655	No
Q22	0.001	0.001	0.610	0.794	Yes
Q23A	0.000	0.000	0.874	0.930	Yes
Q23B	0.039	0.053	0.283	0.655	Yes
Q23C	0.000	0.000	0.932	0.973	No
Q23D	0.723	0.758	0.488	0.767	Yes
Q23E	0.000	0.000	0.647	0.794	Yes
Q23F	0.166	0.200	0.695	0.805	Yes
Q23G	0.000	0.000	0.229	0.605	Yes
Q23H	0.002	0.003	0.964	0.973	No
Q24	0.003	0.005	0.953	0.973	Yes
Q25	0.064	0.079	0.197	0.564	Yes
PrePost	0.000	0.000	0.737	0.839	Yes

Table 3: Survey questions GLM likelihood ratio test results  $\alpha = .05$

### 5.3 Discussion

Out of the 45 variables chosen by the lasso model, 12 of them were not significant using the B-H adjusted p-values with  $\alpha = .05$ . Only 7 of them were not significant using the B-H adjusted p-values with a more lenient  $\alpha = .10$ . Overall, lasso did a good job selecting important variables. We could use the result from the previous section with great confidence.

## 6 Conclusion

In this report we examined the responses to a survey measuring local community pharmacies' Appointment-Based Model implementation and it's effect on medication adherence. We used the lasso model to find the important drivers and the likelihood ratio tests ensured their significance. Overall, all five core components had varying effects on medication adherence. We discussed couple questions in detail and saw that small policy changes from the pharmacy could have major impacts on patient adherence.