AFRL-SA-WP-TR-2017-0005

Genetically Guided Statin Therapy

Deepak Voora, MD
Duke University

March 2017

Final Report
for January 2013 to July 2016

Air Force Research Laboratory
711th Human Performance Wing
U.S. Air Force School of Aerospace Medicine
Aeromedical Research Department
2510 Fifth St., Bldg. 840
Wright-Patterson AFB, OH 45433-7913

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

Qualified requestors may obtain copies of this report from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-SA-WP-TR-2017-0005 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//	//SIGNATURE//
COL NICOLE ARMITAGE	DR. RICHARD A. HERSACK
Chief, En Route Care Research Division	Chair, Aeromedical Research Department

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
1 Mar 2017	Final Technical Report	January 2013 – July 2016

4. TITLE AND SUBTITLE

Genetically Guided Statin Therapy

5a. CONTRACT NUMBER
FA8650-13-2-6375

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Deepak Voora, MD

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Duke University
Office of Research Administration
2200 West Main Street, Suite 710
Durham, NC 22708-4677

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
USAF School of Aerospace Medicine
Aeromedical Research Dept/FHE
2510 Fifth St., Bldg. 840
Wright-Patterson AFB, OH 45433-7913

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-SA-WP-TR-2017-0005

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES
Cleared, 88PA, Case # 2017-1319, 28 Mar 2017.

14. ABSTRACT

Statins are important and effective medications to lower cholesterol and prevent cardiovascular disease. Long-term adherence is a challenge, due, in part, to statin intolerance due to musculoskeletal side effects. In objective #1, we randomly assigned 159 primary care patients with statin intolerance not prescribed statins to either genotype-guided statin therapy (GGST) or usual care. GGST patients and their providers received *5 genetic risk information about statin side effects and genotype-specific statin prescription recommendations. Patients in the usual care arm received general information regarding statin risk and prescriptions. The outcomes measured at 3 and 8 months were statin reinitiation, low-density lipoprotein cholesterol (LDLc), and statin adherence using the validated Morisky Medication Adherence Scale. In primary care patients with statin intolerance, *SLCO1B1* GGST is effective in lowering LDLc compared to usual care (131.9±42.0 vs. 144.4±43.0, p=0.041). GGST, however, did not improve statin adherence in those who reinitiated statin therapy. In objective #2, we used electronic medical record billing, laboratory, and pharmacy data to build and validate a predictive model that identifies patients who will be nonadherent to statin therapy. The overall accuracy of this model as assessed by the area under the receiver operating characteristics curve was 0.81.

15. SUBJECT TERMS
Statins, cholesterol, LDL, cardiovascular disease, genetic-informed strategy, statin prescription, statin adherence

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Cole Hutchison
U	U	U	SAR	34	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.

TABLE OF CONTENTS

Section **Page**

TABLE OF CONTENTS (concluded)

Section **Page**

LIST OF FIGURES

Page

LIST OF TABLES

Page

1.0 SUMMARY

Statins are well established for lowering cholesterol and preventing cardiovascular disease. High rates of statin nonadherence are a concerning phenomenon that are linked to poorer cardiovascular outcomes and greater cost. Statin-induced side effects are a leading cause of statin nonadherence. A genetic variant (*5) in the hepatic uptake transporter gene, *SLCO1B1*, has been linked with side effects and premature statin discontinuation. Delivery of *SLCO1B1*5* risk information and tailored statin therapy may improve adherence and achieve greater cholesterol lowering. We describe a randomized controlled study design testing the hypothesis that *SLCO1B1* genotype-guided statin therapy can improve statin adherence and low-density lipoprotein levels in patients with a prior history of statin intolerance compared to a strategy that does not incorporate genetic testing and the development of a predictive model of statin non-adherence that is applicable to military beneficiaries.

2.0 OBJECTIVE 1: GENETICALLY GUIDED STATIN CLINICAL TRIAL

2.1 Introduction

Atherosclerotic cardiovascular disease accounts for substantial morbidity and mortality worldwide, and the role of statin therapy in cholesterol lowering and cardiovascular risk reduction has been well established. High rates of statin nonadherence are a concerning clinical phenomenon because nonadherence has been linked to poorer cardiovascular outcomes and greater mortality and cost. A genetic variant in a statin uptake transporter gene, *SLCO1B1*, has been linked with increased side effects and premature statin discontinuation. Knowledge of the *5 variant can enable tailored statin therapy in an effort to improve adherence and achieve greater cholesterol lowering. Here we describe a randomized controlled study design that tests the hypothesis that genotype-guided statin therapy (GGST) can improve statin adherence and low-density lipoprotein cholesterol (LDLc) levels in patients with a prior history of statin intolerance.

2.2 Methods

This was a two-arm, unblinded, randomized controlled trial comparing genetically guided therapy to usual care. Duke University, North Carolina, was the primary coordinating center. Primary care sites selected for patient enrollment within the Duke University Health System were part of the Duke University Primary Care Research Consortium. The other primary care site was the David Grant USAF Medical Center (DGMC), Travis Air Force Base, California. The trial stratified patients who were nonadherent to statins due to prior side effects in a 1:1 fashion by *SLCO1B1*5* genotype and site. The study was approved by the Duke and the U.S. Air Force Institutional Review Boards. A full description of the trial methodology is the subject of a publication in press at Pharmacogenomics titled “Rationale and Design of the *SLCO1B1* Genotype Guided Statin Therapy Trial.” The trial was registered in clinicaltrials.gov (NCT01894230).

2.3 Patient Population

Patients considered eligible for this study were current patients (i.e., seen within the last year) within the Duke University Primary Care Research Consortium or DGMC, age >18 years, and current nonutilizers of statin therapy due to discontinuation (per the patient or their provider) due to suspected side effects. An active email account and computer access for completion of online surveys and the ability to provide informed consent were also requirements for eligibility.

2.3.1 Patient Exclusion Criteria. Patients were excluded from the study if they had a history of rhabdomyolysis or creatinine kinase elevation greater than 10 times the upper limit of normal or unexplained elevation of hepatic enzymes (aspartate transaminase or alanine transaminase greater than three times the upper limit of normal) while on any statin therapy. Patients who previously used more than four statins were not included in the study. Patients with daily grapefruit juice usage of 1 quart per day on average were excluded due to the potential inhibition of CYP450 3A4 affecting statin metabolism [1]. Expected long-term use (>3 months) of drugs known to interfere with statin metabolism or disposition at the time of randomization was a criterion for exclusion as well. Short-term use of less than 14 days was permitted. Participants of other drug research studies within 30 days of this trial were also excluded. Statin use within the last 6 weeks was also an exclusion criterion.

2.3.2 Sample Size and Power. To achieve 95% power in our study using a Student's t-test, we calculated the required number of participants to be 150 to detect a significant difference in the primary outcome (a 1.0-point difference in the Morisky Medication Adherence Scale [MMAS]). Considering an expected dropout rate of 10%, the total sample size was calculated to be 167.

2.3.3 Recruitment and Randomization. Potential subjects were identified either through physician referral, patient self-referral, or electronic medical record chart review. The Duke Enterprise Data Unified Content Explorer, an informatics tool for screening patients, was used at the Duke sites participating in the trial by searching for a prior history of statin allergy or intolerance. A letter was mailed by the study coordinator (signed by the subjects' primary care physician) to the subjects identified through chart review. A follow-up phone call by the study coordinator was made to assess subject interest at 2 weeks, and consent and baseline visits were scheduled. For self-referred or physician-referred subjects, follow-up consent and baseline visits were scheduled in a similar fashion after a screen to assess eligibility.

Randomization was conducted by the Data Manager, who was not involved in any part of the data analysis or subject recruitment or contact. Since the *SLCO1B1*5* genotype was expected to have a detectable impact on study outcomes, we performed stratified randomization based on *SLCO1B1*5* genotype (carriers vs. non-carriers) and clinic site. The results of the genotype testing were sent to the Data Manager by email and were entered into an electronic database along with recruiting site information. At this point, randomization took place; when completed, the research coordinator assessed that all baseline surveys were complete and no prohibited medications as described above had been prescribed in the meantime (through electronic medical record [EMR] review). Randomization messages were sent to the providers and subjects via email in both arms, but the usual care arm did not receive genotyping results at this point. They were provided the genotype results at the 8-month period marking the completion of the study.

2.3.4 Survey Instruments. The survey instruments administered to the patients in this study are described below. The primary outcome of the study was the eight-item MMAS, which includes eight yes/no items that are summed to create an overall adherence score ranging from 0 to 8, with higher scores indicating better adherence. This survey has been validated and applied to evaluate LDLc goal achievement in clinical practice previously and is a quick, patient-reported measure of adherence that is easy to implement in the outpatient setting [2].

To assess patients' beliefs about medications, which may play a critical role in their adherence behavior, the Beliefs about Medications Questionnaire (BMQ) was administered with the score ranging from 5 to 25 generated from the sum of five questions [3]. This questionnaire assesses beliefs regarding necessity of and concerns with disease-specific medications.

The Medication Possession Ratio (MPR) is defined based on the number of statin refills over time from randomization to end of follow-up. The sum of the number of pills dispensed for each statin prescription during the follow-up period divided by the sum of the actual number of days of follow-up (date of 8-month survey minus the date of randomization) is the MPR.

The Brief Pain Inventory (BPI) is a self-administered questionnaire that was initially developed to assess pain in cancer patients but has subsequently been validated as a form of assessing noncancer pain in patients [4].

The SF-12 is a widely validated 12-item questionnaire designed to generate a physical and mental health summary for application to large populations [5].

2.3.5 Timeline and Follow-Up. At the baseline visit, height, weight, blood work for fasting lipids, and genotype assessment were collected. Surveys were administered to identify self-reported subject demographics such as age, sex, race, ethnicity, education, smoking, and alcohol use. The MMAS and BMQ were administered at baseline and at 3- and 8-month follow-up visits. Usual physical activity was measured with the Stanford Brief Activity Survey at baseline and at the 8-month follow-up, and the SF-12 health survey assessed quality of life at baseline and at the 3- and 8-month follow-up visits. A diet history and the BPI were also taken at this visit. The BPI was repeated at the 3- and 8-month follow-up visits.

The study followed the patients for an 8-month time period during which study outcomes were assessed at various time points. At 3 months, surveys and fasting lipid blood work were collected. New statin prescription and medication use records were collected from the pharmacy or the EMR. At 4 months and 6 months, the EMRs were reviewed for statin utilization and medication use records from the pharmacy or EMR were collected. At 8 months, this review was repeated and a final set of surveys and fasting lipid blood work was collected. Subjects who were randomized to the usual care arm received their genotyping results after completion of all study requirements at the 8-month period along with a copy to their primary care provider. Figure 1 shows the timeline of the study.

2.3.6 Study Outcomes. The primary outcome of this study is medication adherence, assessed by the MMAS tool. The survey was administered at baseline and at 3 and 8 months of follow-up. The primary hypothesis is that GGST leads to greater adherence, corresponding to a higher MMAS score. The secondary outcomes to be studied are (1) LDLc at 3 months and 8 months, (2) MPR at 8 months, (3) number of new statin prescriptions, and (4) patient reported quality of life, physical activity, perceptions regarding statin therapy, and pain as assessed by the survey instruments described above.

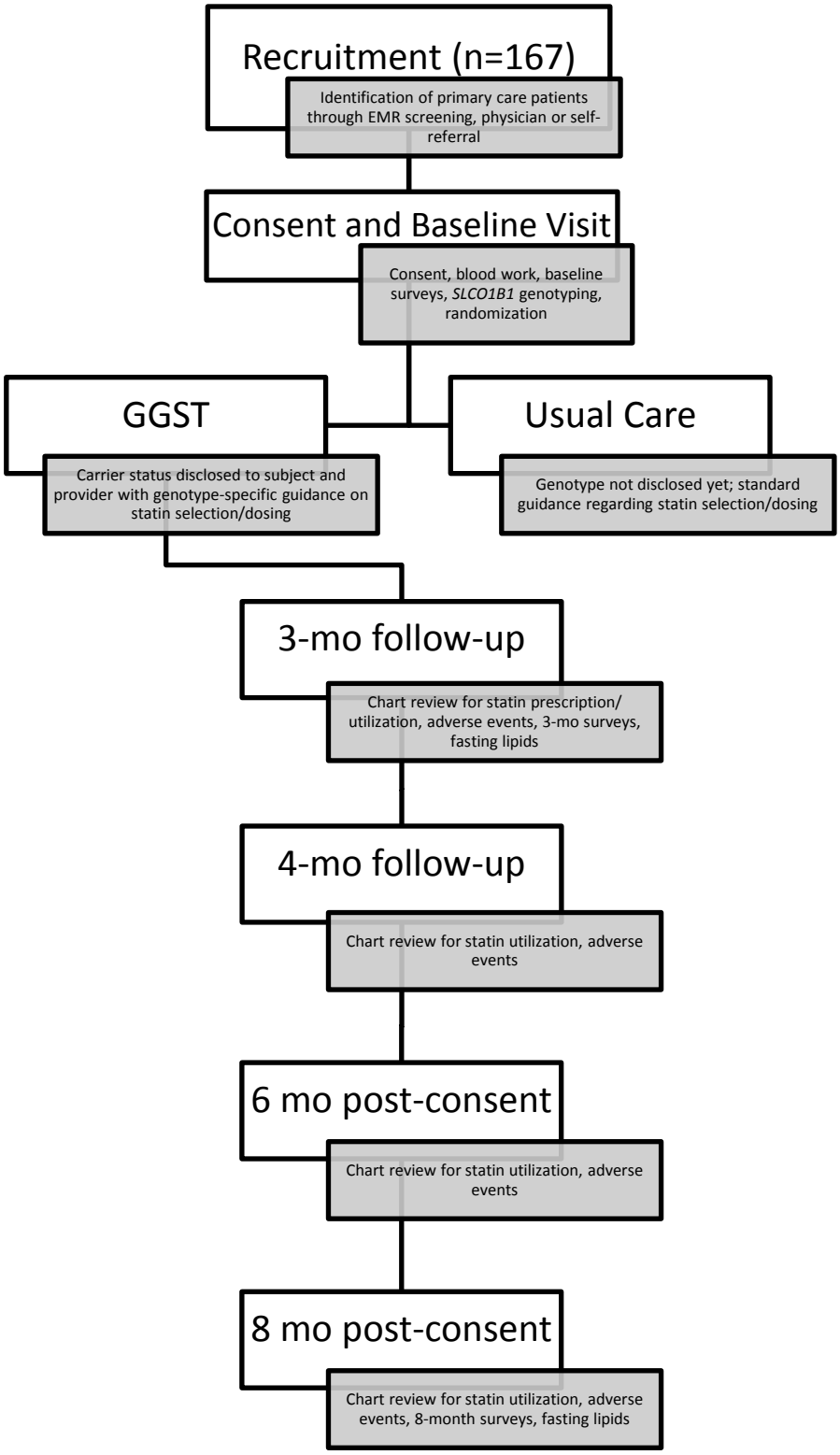

Figure 1. Schedule of study from recruitment to 8-month follow-up.

2.4 Data Collection and Analysis

De-identified subject data were entered by site coordinators into a secure electronic database (REDCap), which is HIPAA [Health Insurance Portability and Accountability Act] compliant. Data were stored on a secure server managed by the Duke University School of Medicine IT Department. Only study coordinators and Data Management Staff were allowed access to the password-protected database. The database was locked once it had been declared complete and accurate. Subject confidentiality was maintained by ensuring that each subject had a unique study identification number to which only key research personnel had access for scheduling follow-up surveys and obtaining laboratory test results. No individual identifiers were published in any of the research publications associated with the study. Data collection was conducted periodically through EMR review at the 3-, 4-, 6-, and 8-month follow-up periods for adverse events and statin prescriptions as described below. Data monitoring visits were completed at the milestones of 5, 100, and 167 patients enrolled. Remote monitoring and quality control activities occurred periodically throughout the study.

2.4.1 Statistical Analysis. For the baseline characteristics defined above, frequencies were reported for the categorical data, and continuous data were reported as mean and standard deviation (SD) or median and interquartile range.

The primary outcome is the MMAS. A Poisson log-linear generalized linear model was used with fixed effects for the arm, genotype, and clinic site. Any variables that were imbalanced between the two arms were included as covariates and the results were presented as the additive effect of treatment on the log score with 95% confidence intervals. Diagnostics were used for model checking, including comparisons of the fit of negative binomial models and comparisons of the observed and predicted counts. The following statistical plan was utilized for the secondary outcome measures. For all statistical models, arm, genotype, and site were used as predictors, and baseline values were included as covariates to account for variability unless specified otherwise.

- LDLc at 3 months and 8 months: This continuous outcome was modeled as a linear regression with arm, genotype, and site as predictors. Baseline LDLc was included as a covariate to account for any variability in subject baseline.
- MPR at 8 months: This was modeled as a linear regression.
- Number of new statin prescriptions at 3 months: For this binary outcome, a logistic regression was used to model the results.
- BPI at 3 and 8 months: Pain severity and interference were compared and this was modeled as a linear regression accounting for baseline pain score variability.
- SF-12: 3-month and 8-month scores were compared with baseline scores included as covariate using a linear regression model.
- Physical activity: This was compared at the 8-month follow-up period using baseline physical activity as a covariate, and a proportional odds model was used. If the assumption was not met, a multinomial regression model was used.
- BMQ: This was administered at baseline, 3 months, and 8 months. The final score was modeled with Poisson regression.

A secondary hypothesis was also tested with respect to the interaction between genotype and intervention for the MMAS and LDLc outcomes. A variable was introduced into the model that is the product of these variables and a likelihood-ratio test was used to compare models with and without the interaction. If significant, contrasts were constructed to estimate effects at each level of the interaction. All analysis was conducted in SAS version 9.3 (SAS Institute, Inc., Cary, NC).

The proportion of missing data was summarized at each time point by treatment group, and known reasons for missing data were documented. Reasons for missing data were explored with summary statistics and graphical assessments. For the primary outcome, the number of questions answered were included as an offset variable. For all other outcomes, only complete cases were included in the analytic dataset.

2.4.2 Adverse Event Reporting and Safety Monitoring. In June 2011, the Food and Drug Administration determined this clinical investigation to be a nonsignificant risk device study. Any adverse events that developed during the study protocol were reported by subjects to their primary care physician. The principal investigators and study coordinators were available for questions. The EMR was reviewed for aggregated adverse events and clinical events at the 3-, 4-, 6-, and 8-month time points. Any serious adverse events were immediately reported by the study site coordinator to the site principal investigator (PI) and the coordinating center PI. A research monitor was appointed by the Institutional Review Board to coordinate and review any adverse events and discuss them with the PI at a minimum interval of every 6 months.

2.5 Results and Discussion

2.5.1 Patient Populations. The CONSORT diagram in Figure 2 illustrates the number of patients who were screened for eligibility and agreed to consent to the study. We randomly assigned 159 primary care patients to GGST vs. usual care. All patients (100%) were successfully genotyped for the *SLCO1B1*5* genetic variant, and the median turnaround time was 6 days. The allele frequency for the *5 variant was 25%, which is consistent with the allele frequency for this variant in Caucasians. Follow-up was complete in 147/159 (92%) for the primary outcome.

2.5.2 Baseline Characteristics of the Population. As shown in Table 1, the mean age was 63 years, 57% were female, 41% reported intolerance to > 1 statin, 73% reported myalgia with statins, and 25% were *5 carriers. At baseline, there were some minor differences between the GGST arm and usual care: more women and higher HDL and LDL in the usual care arm.

2.5.3 Primary Outcome. The MMAS was collected at 3 and 8 months in patients who were reinitiated on statin therapy by their providers. We found no differences in the MMAS at 3 or 8 months between the GGST and usual care arms: 3 months: 6.8 (1.5) vs. 6.9 (1.6), $p = 0.89$ and 8 months: 6.8 (1.7) vs. 7.1 (1.3), $p = 0.75$.

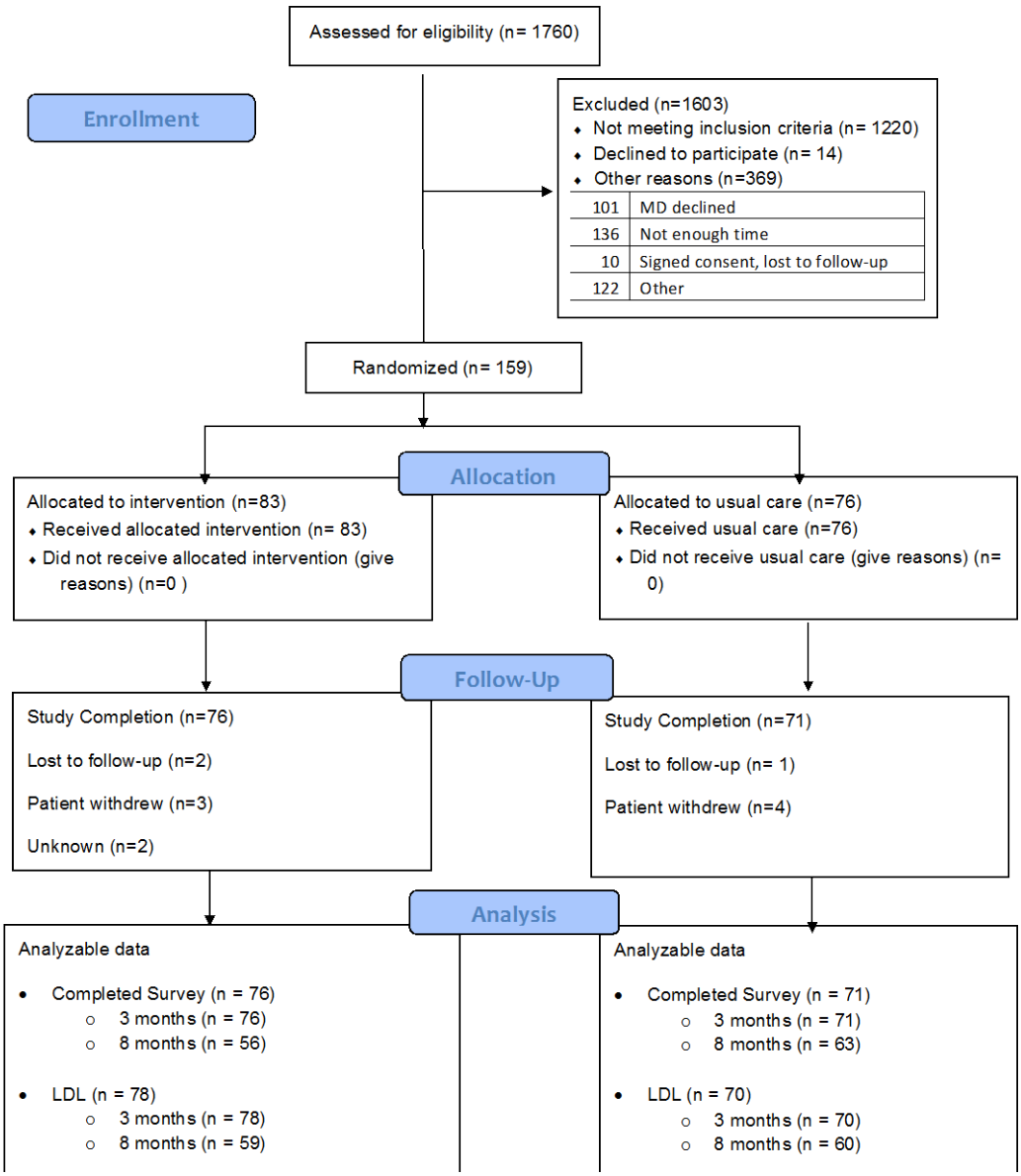

Figure 2. CONSORT diagram.

Table 1. Subject Demographics

Characteristic	Genotyping + Usual Care (N=83)[a]	Usual Care Only (N=76)[a]	Total (N=159)[a]
Age, mean (SD), yr	62.7 (10.2)	62.5 (11.5)	62.6 (10.8)
Gender – female	41 (49.4%)	50 (65.8%)	91 (57.2%)
Race			
White	66 (79.5%)	61 (80.3%)	127 (79.9%)
Black/African American	14 (16.9%)	11 (14.5%)	25 (15.7%)
Other	3 (3.6%)	4 (5.3%)	7 (4.4%)
Height, mean (SD), cm	169.9 (10.5)	167.3 (9.4)	168.7 (10.0)
Weight, mean (SD), kg	87.4 (16.8)	86.4 (23.3)	86.9 (20.1)
Body mass index, mean (SD), kg/m²	30.3 (5.9)	30.6 (6.8)	30.5 (6.3)
Pravastatin, rosuvastatin, or fluvastatin intolerence	27 (32.5%)	26 (34.2%)	53 (33.3%)
Number of statin intolerances			
1	51 (61.4%)	43 (56.6%)	94 (59.1%)
2	21 (25.3%)	23 (30.3%)	44 (27.7%)
>2-4	11 (13.3%)	10 (13.2%)	21 (13.2%)
Statin symptoms			
Muscle aches	61 (73.5%)	55 (72.4%)	116 (73.0%)
Joint aches	7 (8.4%)	11 (14.5%)	18 (11.3%)
Fatigue	7 (8.4%)	4 (5.3%)	11 (6.9%)
Other	30 (36.1%)	27 (35.5%)	57 (35.8%)
Comorbidities			
Cardiovascular disease or congestive heart failure	7 (8.4%)	5 (6.6%)	12 (7.5%)
Type 1 or Type 2 diabetes	20 (24.1%)	15 (19.7%)	35 (22.0%)
Hyperlipidemia	76 (91.6%)	71 (93.4%)	147 (92.5%)
Hypertension	46 (55.4%)	46 (60.5%)	92 (57.9%)
Hyperthyroidism	6 (7.2%)	10 (13.2%)	16 (10.1%)
Osteoporosis	6 (7.2%)	5 (6.6%)	11 (6.9%)
Concomitant medications			
<3	5 (6.0%)	7 (9.2%)	12 (7.5%)
3-8	43 (51.8%)	32 (42.1%)	75 (47.2%)
>8	35 (42.2%)	37 (48.7%)	72 (45.3%)
Study site			
DGMC	6 (7.2%)	7 (9.2%)	13 (8.2%)
Duke	77 (92.8%)	69 (90.8%)	146 (91.8%)
Cholesterol, mg/dL			
LDL	152.7 (41.1)	157.6 (41.9)	155.0 (41.4)
High-density lipoprotein (HDL)	49.8 (14.9)	55.5 (14.9)	52.5 (15.1)
Non-HDL	186.2 (42.8)	187.1 (41.7)	186.6 (42.1)
Triglycerides	167.7 (100.3)	150.3 (87.6)	159.4 (94.5)
Total cholesterol	235.9 (45.2)	242.6 (47.1)	239.1 (46.1)
Education level			
Less than high school/high school/GED	10 (12.0%)	10 (13.2%)	20 (12.6%)
Some college/associate degree	27 (32.5%)	28 (36.8%)	55 (34.6%)
Bachelor's/master's/doctorate degree	46 (55.4%)	38 (50.0%)	84 (52.8%)
Alcohol use	43 (54.4%)	47 (61.8%)	90 (58.1%)

[a]Data expressed as number (%) unless otherwise specified.

2.5.4 Secondary Outcomes

2.5.4.1 Cholesterol (LDL, HDL, Total, Non-HDL, and Triglycerides). At 3 and 8 months there was a significant improvement in LDL, total, and non-HDL cholesterol in the GGST arm vs. usual care (Figures 3-5). There was no difference in triglyceride levels (Figure 6), and although HDL cholesterol was different at 3 and 8 months (Figure 7), because of the difference in baseline HDL (Table 1), when adjusted for baseline HDL the differences at 3 and 8 months were no longer significant (p = 0.34 for both).

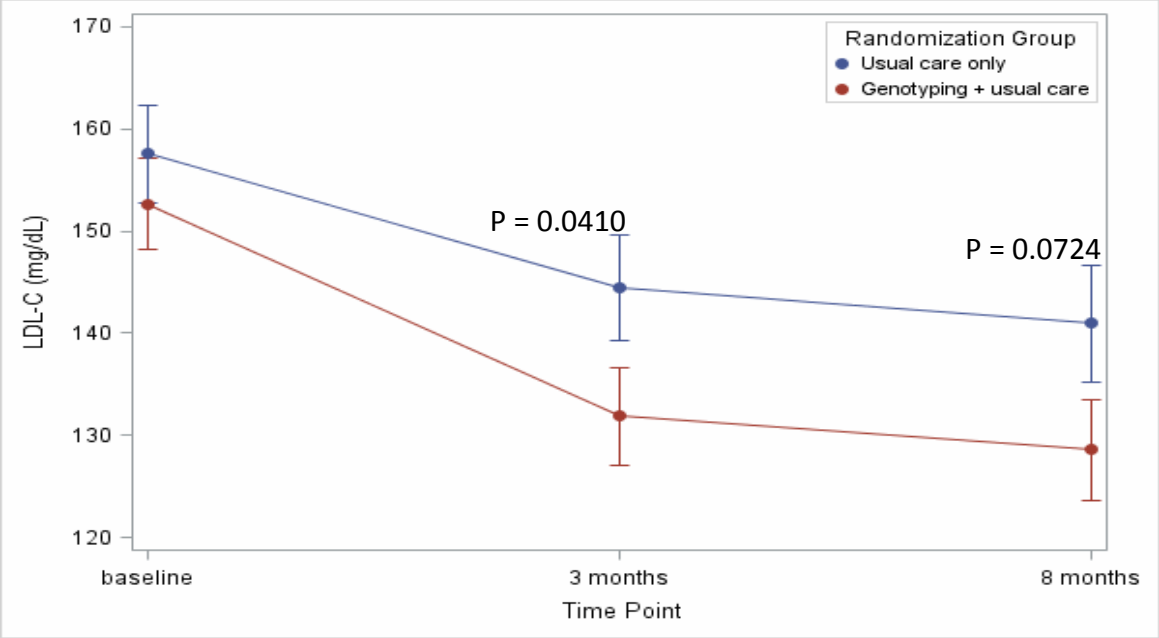

Figure 3. LDL cholesterol.

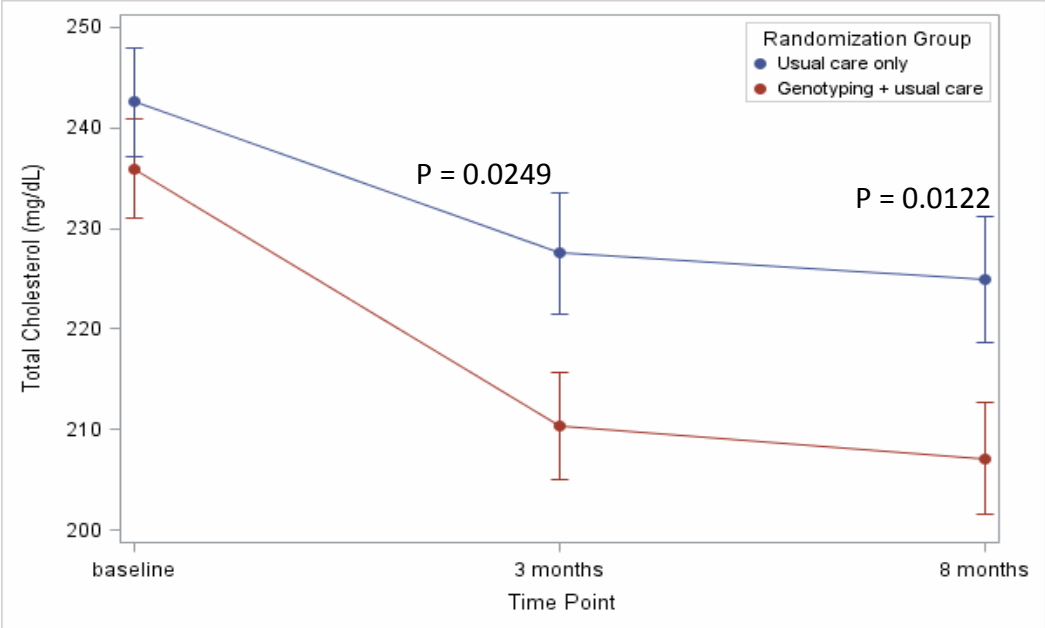

Figure 4. Total cholesterol.

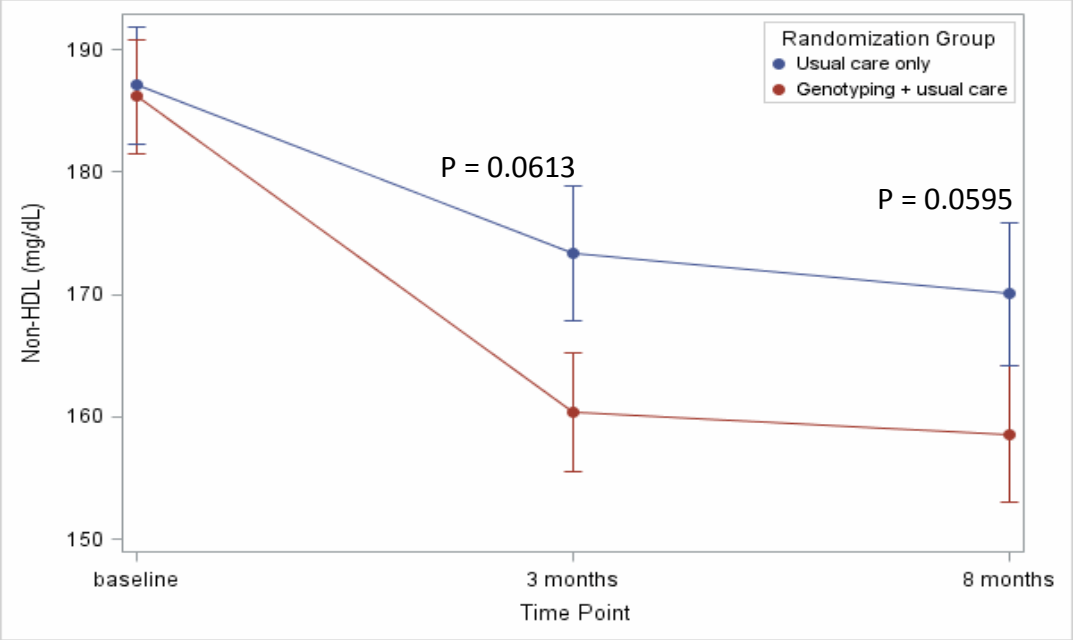

Figure 5. Non-HDL cholesterol.

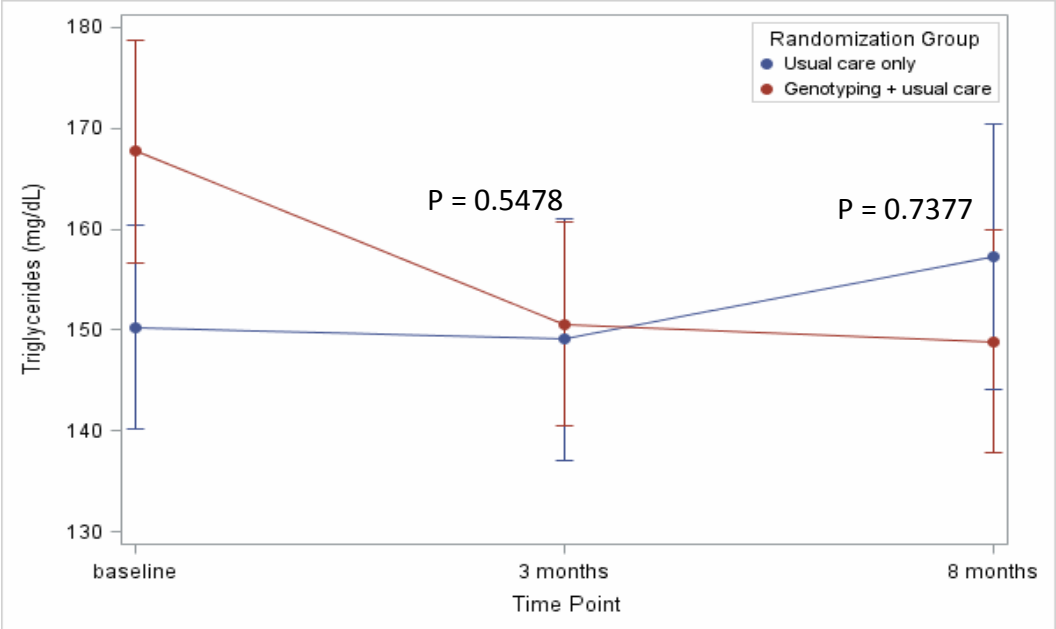

Figure 6. Triglycerides.

2.5.4.2 Statin Prescriptions. Based on patient reports at 3 and 8 months, statin prescriptions were higher in the GGST vs. usual care arms (Figure 8).

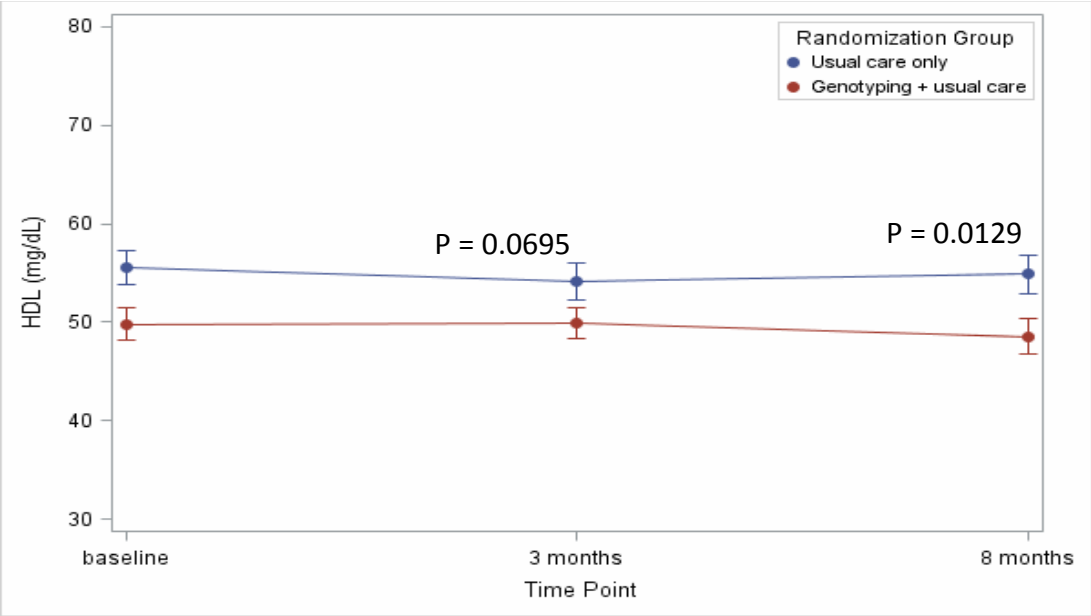

Figure 7. HDL cholesterol.

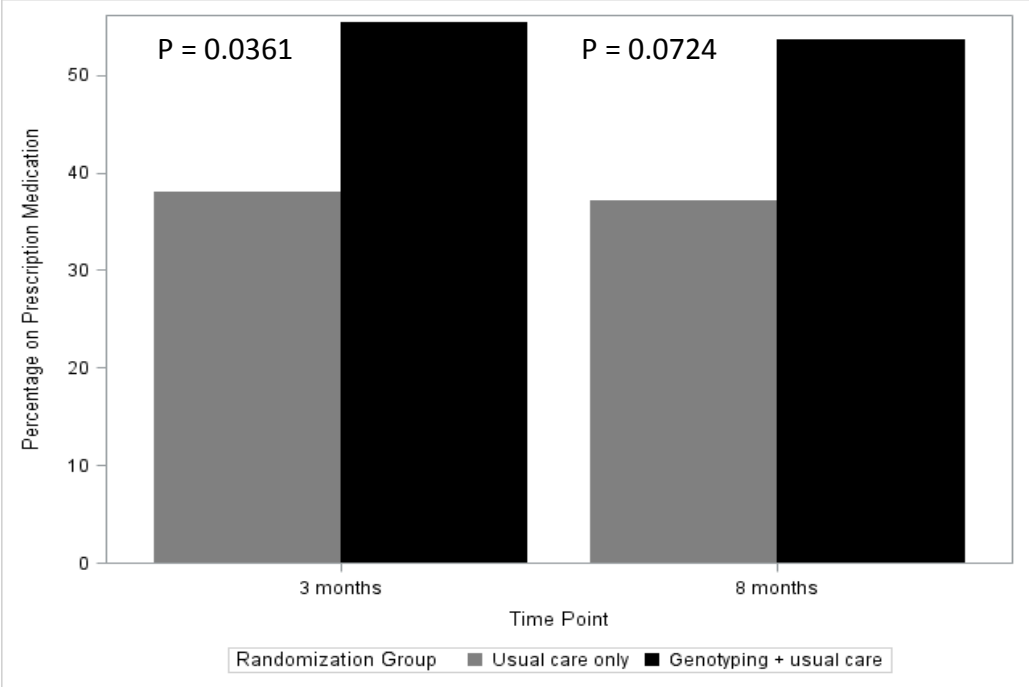

Figure 8. Statin prescriptions.

2.5.4.3 BMQ. Compared to patients in the usual care arm, those in the GGST arm had an improvement in the BMQ Necessity score at 3 months but not at 8 months (Figure 9). There were trends toward lower BMQ Concerns in the GGST arm at both 3 and 8 months (Figure 10). The combined effects on the BMQ Necessity and Concerns resulted in trends in improvements in the BMQ Necessity-Concerns differential at both 3 and 8 months in the GGST arm (Figure 11).

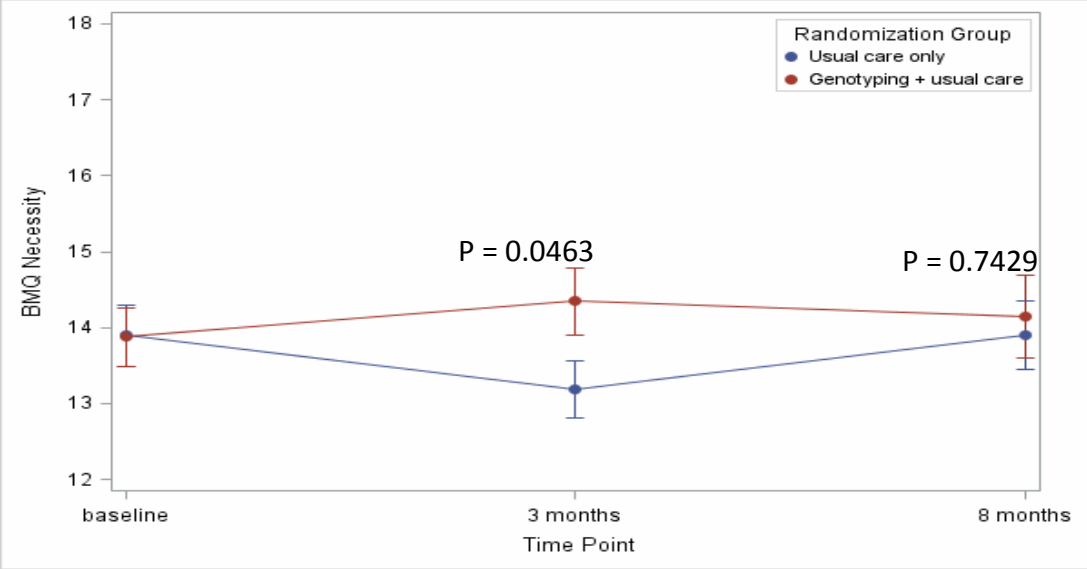

Figure 9. BMQ Necessity.

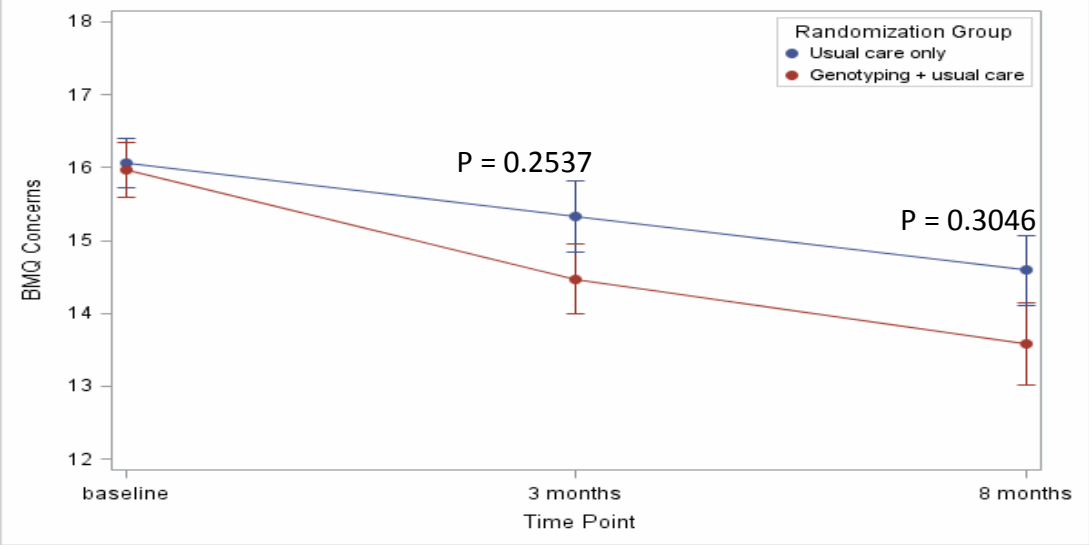

Figure 10. BMQ Concerns.

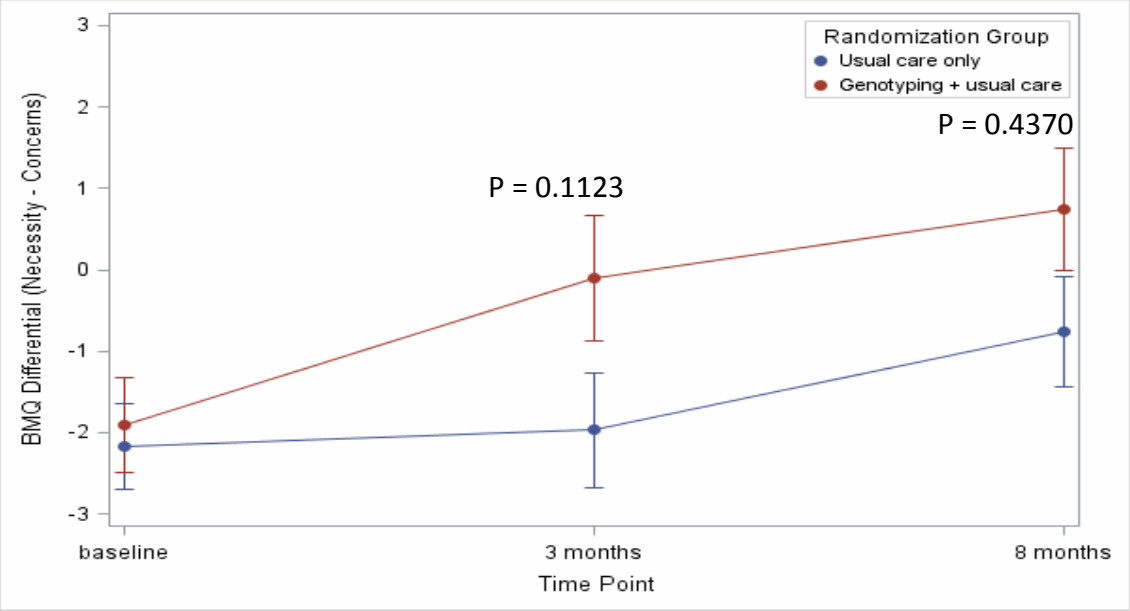

Figure 11. BMQ Necessity-Concerns differential.

2.5.4.4 BPI, SF-12, and MPR. There were no significant differences in the BPI, SF-12, or MPR between GGST and usual care arms.

2.5.5 Exploratory Outcomes: Post-Study Cross-Over Outcomes. As part of the study protocol, *SLCO1B1*5* genetic test results were returned to patients and providers randomized to usual care after all study visits were completed. Genotype results were returned in an identical manner as the intervention arm. As a further test of our hypothesis that a genotype-guided strategy of statin prescription would improve LDL cholesterol, we compared LDL cholesterol values after study completion in the usual care arm (that had received genotype results after study completion) vs. the GGST arm that received no new information after study completion. We found that after receiving *5 genotype results, patients in the usual care arm had significant decrease in their LDL cholesterol post-study compared to patients in the GGST arm (Figure 12). Therefore, by the end of the study, when all patients in both the GGST and usual care arms had both received their *5 genotype results, LDL cholesterol values had declined by similar amounts (Figure 13) compared to baseline. The GGST arm patients had a decline in their LDL within study, whereas the usual care patients had a decline in their LDL post-study during the usual care and GGST arm patients had similar LDL cholesterol values at the end of the study (Figure 14).

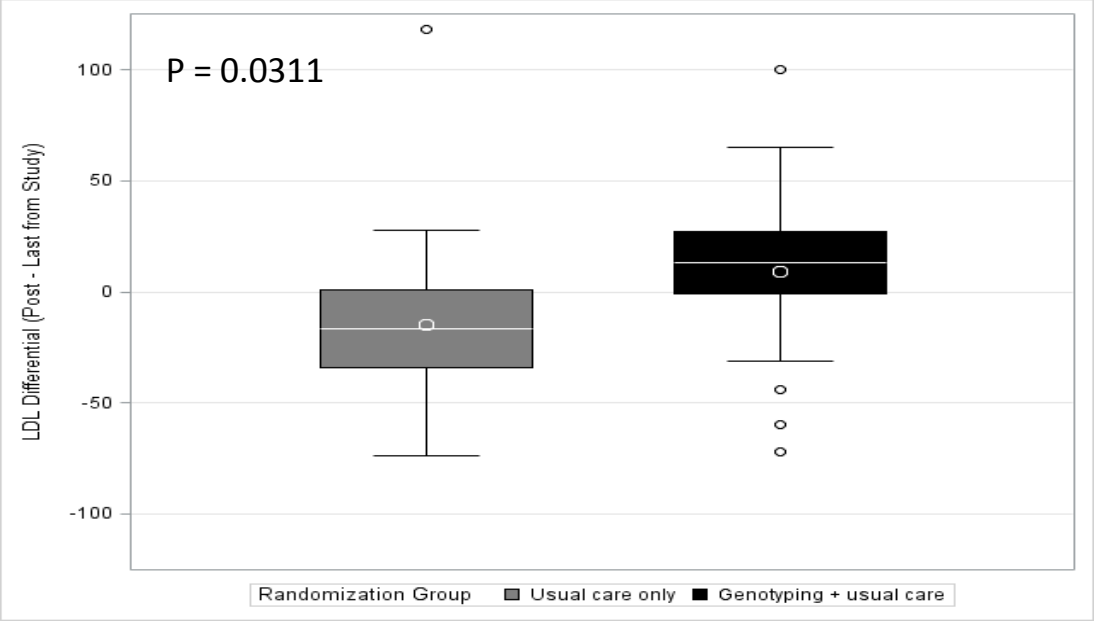

Figure 12. Change in LDL post-study.

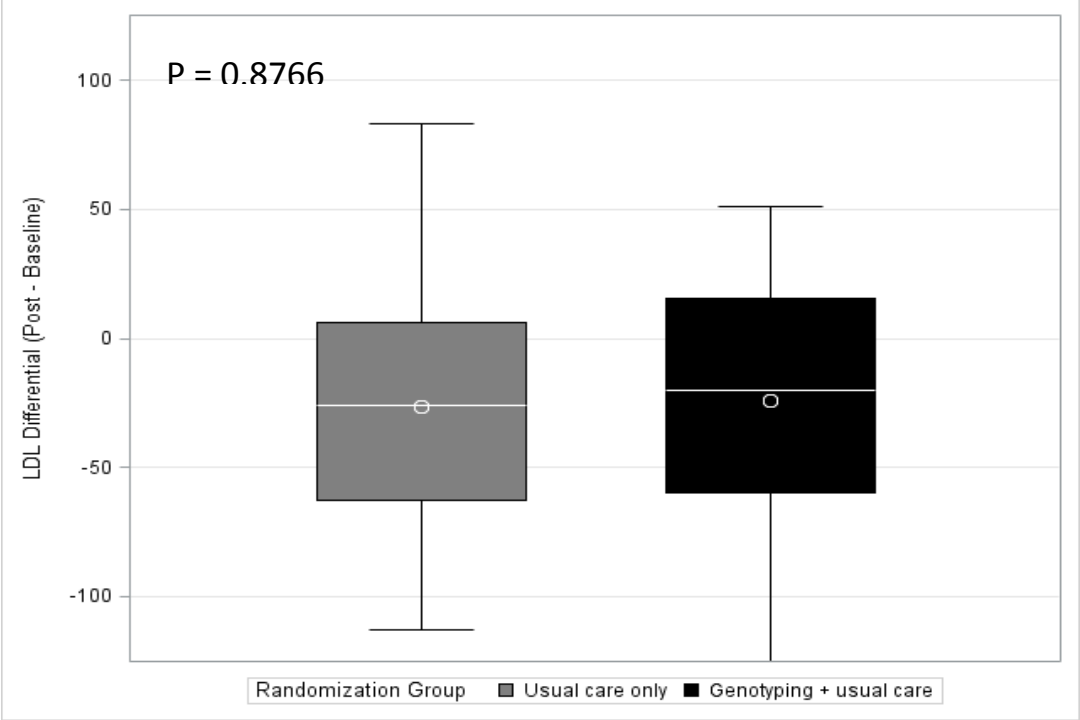

Figure 13. Comparison of changes in LDL cholesterol from baseline to post-study.

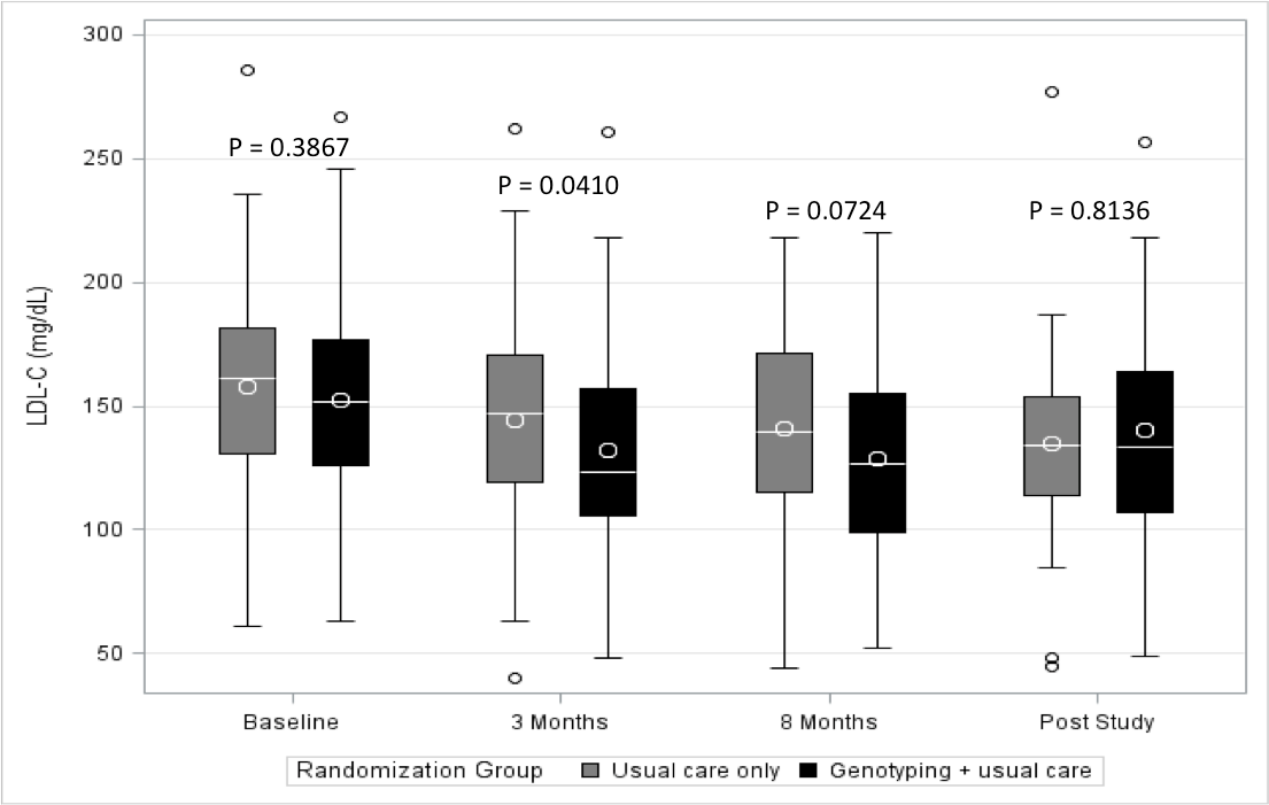

Figure 14. Comparison of LDL values during study (baseline, 3 months, and 8 months) when only GGST patients received SLCO1B1*5 genotype results post-study when both usual care and GGST arms received *5 genotype information.

2.6 Conclusions

This prospective, randomized clinical trial was the first to test the effect of *SLCO1B1*5* guided statin therapy and clinically relevant outcomes like lipid lowering and medication adherence. In primary care patients with statin intolerance, *SLCO1B1* GGST is effective in lowering LDLc compared to usual care. GGST, however, did not improve statin adherence in those who reinitiated statin therapy. As genomics provides us with the tools, more and more pharmacogenomics biomarkers are being discovered for certain drugs. At the same time, there is a lack of evidence that use of such biomarkers can improve health outcomes that are important to patients and providers. Statins are no exception, and with evidence to support that statin-induced side effects affect adherence and cardiovascular outcomes, understanding whether using genomics could improve adherence and minimize side effects for better outcomes becomes important. Observational studies, retrospective analyses, and case-control studies have identified the association between *SLCO1B1*5*, but investigating its impact in real-world clinical practice has not yet been accomplished. The cost of genomic profiling is rapidly declining with improvement in technology, and the financial and practical feasibility of applying it in routine clinical practice is no longer considered an insurmountable barrier. With all pharmacogenomics studies, clinical implementation for personalized medicine is the next step to prove both utility and cost effectiveness, as well as establish the validity of the described association in prospective randomized controlled studies.

3.0 OBJECTIVE 2: PREDICTIVE MODEL OF STATIN NONADHERENCE

3.1 Introduction

Interventions to improve medication adherence must be preemptive. A key, first-step to deploying interventions that aim to prevent nonadherence is the accurate identification of individuals at high risk for statin nonadherence. Therefore, we designed Objective 2 to allow the military to identify those individuals at highest risk for nonadherence. Using the large (~400,000 beneficiaries) datasets available through the Military Health System (MHS) and our statistical expertise, we aimed to develop a robust model to predict nonadherence. Such a tool could then be deployed within the military to preemptively target future interventions to improve statin adherence and ultimately cardiovascular health.

3.2 Data Generation

3.2.1 Data Sources. Data for this study were collected from the MHS Management Analysis and Reporting Tool, which contains subject Medical Treatment Facility enrollment data as well as inpatient and outpatient healthcare service elements provided by the Medical Treatment Facility, civilian providers, hospitals, and managed care support contractors. The Pharmacy Data Transaction Service as well as the MHS Composite Health Care System provided pharmaceutical and laboratory data, respectively.

3.2.2 Inclusion Criteria. We defined the study cohort to be statin-naïve MHS adults between 18 and 65 years of age and continuously enrolled in TRICARE PRIME for at least 5 years on 1 January 2005. These subjects were prescribed their first statin medication between 1 January 2005 and 31 December 2006. National Drug Codes were used to identify prescriptions for statin medications. Subjects were counted only once during the 2-year period and were included in the cohort irrespective of medication refills, discontinuation, or switches in statin drug class.

3.2.3 Baseline Data. Demographic information from the above-mentioned sources, in addition to baseline laboratory elements of LDL cholesterol, creatinine, creatinine kinase, HDL cholesterol, total cholesterol, and triglycerides, were collected for each study cohort subject prior to the first statin prescription in the 2-year period. The most recent laboratory data were collected for each patient soonest to the start of the stain prescription within a 2-year look-back period. Any “other” medications prescribed and filled after statin initiation through 6 months were also captured and recorded.

3.2.4 Outcome Data. Selected outcome data were collected for three calendar years after the first filled statin prescription for each subject. These outcome elements included laboratory data and selected medication refills. The “other” prescribed medications as described above were assessed; those that were the two most commonly prescribed and filled medications in each subject were included with the statin for aggregate medication outcome data. In addition, both direct care and purchased care inpatient records were assessed for cardiovascular events including acute coronary syndrome, angina pectoris, and stroke.

3.3 Data Preparation

The initial dataset consisted of medication fill amounts, laboratory names and values, and Healthcare Common Procedure Coding System (HCPCS) codes for procedures and diagnoses as described above. Dates for all were truncated by the sponsor at the month to partially address patient privacy concerns. Demographic variables available on each patient include sex, ethnicity, race, marital status, and drug delivery program. The sample included data from 449,895 unique patients. We dropped from our analysis the following collections of patients:

1. Any patients who did not fill a statin (defined by National Drug Code) prescription on the start date identified in the dataset
2. Patients who had filled a prescription for a statin in the 12 months before the statin start date to limit the analysis to new starts
3. Patients whose initial prescription duration was not equal to 30, 60, 90, or 180 days
4. Patients with less than 1 year of follow-up

3.4 Outcomes

To evaluate the predictive models developed in this study, we analyzed three related outcomes: (1) statin adherence, (2) cholesterol, and (3) hospital admissions for atherosclerosis-related events in the 3 years following statin initiation.

To assess adherence to statin therapy, we used the percent days covered (PDC), which was calculated beginning on the statin start date and continuing until the last observation of any kind made by the health system. PDC was computed as the ratio of total days of purchased statin medication (of any type) to the number of days of follow-up. In addition, binary indicators identifying (1) patients who purchased their first refill and (2) patients with PDC > 80% were computed. For models that incorporate an indicator of whether the first statin refill was purchased, we also calculated a follow-up PDC beginning at the end of the duration of the first prescription and not counting the pills in the first prescription.

To examine the effects of statin therapy on cholesterol-related laboratory outcomes, we chose very low-density lipoprotein (VLDL), HDL, LDL, and total cholesterol. Indicators of increasing (i.e., nondecreasing) or decreasing values were computed by comparing the last value measured before or on the statin start date to the first value measured 2 months subsequent to the statin start date.

To examine the effects on clinical outcomes known to be prevented by statin therapy, we examined hospitalizations for three diagnoses: acute myocardial infarction (MI), stroke, and coronary artery disease (CAD). Binary outcomes were created by identifying patients with a code from one of these disease-based sets within 1 year of the statin start date.

3.4.1 Independent Variables. Potentially predictive variables were collected from four sources: laboratory, prescription fill, HCPCS codes, and demographic data collected prior to statin start date. Count matrices were computed for each observation type (laboratory, prescription, and codes) where the i^{th} element of the j^{th} row indicates the number of times observation type i was recorded in the record of patient j in the 6 months up to and including the statin start date. This process generated 26,291 potential independent variables for use in predicting future adherence

behavior. We eliminated from consideration any of these potential predictors that did not appear in the records of at least 100 different patients.

In addition, we obtained for each patient (1) the last laboratory value (for every laboratory test available in the dataset) measured before or at the statin start date and (2) the prior PDC for each chronic medication available in the record. Laboratory values for which we did not have data on at least 100 patients were dropped from the analysis.

Prior PDC was computed for each medication and each patient using the number of days from the first prescription of the medication in the dataset to the statin start date. Note that because we are interested in adherence after the *first* prescription of a statin, there are no statins in the list of medications for which we have prior PDC. Prior PDC was only computed for a particular patient/medication pairing if there were at least 2 months of data to use for computation. Medications for which we did not have prior PDC values in at least 100 patients were dropped from analysis.

After filtering variables observed in less than 100 patients, 850 count variables, 106 lab value variables, and 874 prior PDC values on 140k patients remained.

3.4.2 Other Variables. In addition to the clinical variables described thus far, we considered the inclusion of age; gender; race; smoking status (based on HCPCS code); total number of health system interactions in the past 6 months split by labs, drugs, and codes; statin strength (mg); and initial supply duration for inclusion in models. These variables along with their association with first refill behavior (Kruskal-Wallis) are listed in Table 2.

Table 2. Other Variables

Variable[a]	Mean/Probability, Adherent (14328)	Mean/Probability, Nonadherent (124403)	p-value
Age (yr)	54.111	49.2978	<1e-20
Female	0.42525 (6093)	0.42712 (53135)	0.677
Black	0.079634 (1141)	0.079154 (9847)	0.84
Tobacco	0.051368 (736)	0.043552 (5418)	1.69E-05
Follow-Up (days)	1325.7711	1260.4477	<1e-20
Product Strength (mg)	45.9049	43.6006	0.000923
Days Supply	68.7734	66.6115	2.23E-17
# Labs	3.7628	3.4569	0.00018
# Drugs	5.0409	3.808	<1e-20
# Codes	1.6607	0.8532	<1e-20

[a]Additional variables included in the model and association with PDC>80%.

Figure 15 (a-d) shows the association between PDC and changes in cholesterol levels for patients in our study. Note that there is a strong association between failing to purchase medication (low PDC) and increased LDL (c) and total cholesterol levels (d). There is a weaker but also statistically significant negative association between PDC and VLDL – mean PDC 0.44 vs. 0.39 among patients with decreasing and increasing VLDL, respectively, Kruskal-Wallis p-value $<10^{-20}$. There is also a statistically significant positive association between PDC and HDL – mean PDC 0.42 vs. 0.43 among patients with decreasing and increasing HDL, respectively, Kruskal-Wallis p-value $<10^{-20}$. These patterns are consistent with the known effects of statin therapy on lipid profiles.

Paradoxical outcomes. One of the challenges in working with observational data is the sometimes strong relationship between comorbidities and behavior. Statin therapy directly lowers lipids – particularly LDL cholesterol. However, the ultimate goal of statin therapy is to decrease incidence of CAD, acute myocardial infarction and perhaps stroke. However, there is a surprising paradoxical relationship between PDC for statins and the incidence of MI and CAD – Figure 15 (e & f). We hypothesize that this may be because patients with a higher likelihood of these comorbidities are likely aware of the severity of their disease and therefore more adherent.

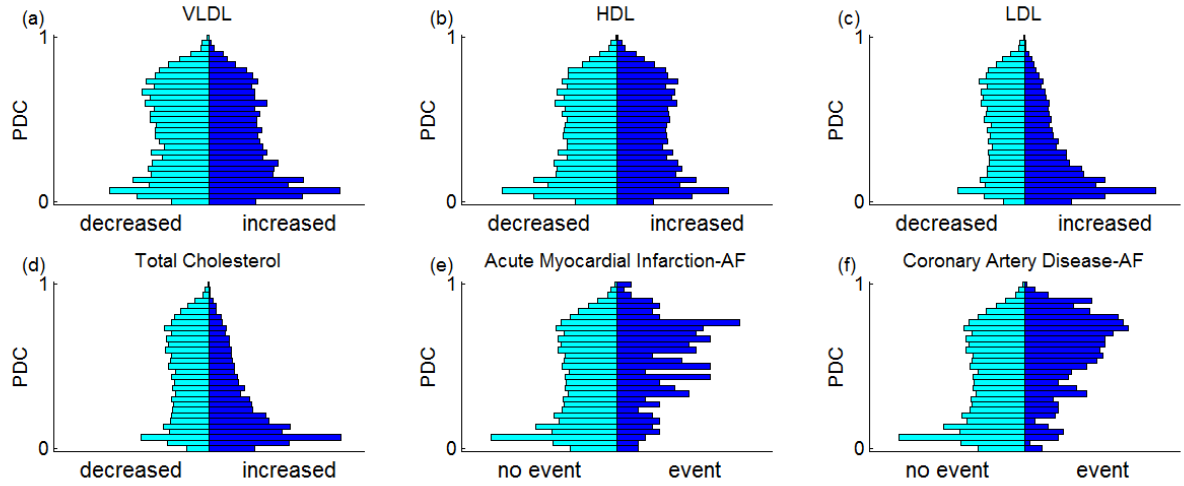

Figure 15. Comparison of PDC with various lipid-related outcomes. *Each figure shows a pair of histograms. The y-axis shows PDC and the x-axis shows frequency. Panel (c) shows that patients with lower PDC are much more likely to have increased LDL. Similarly, nonadherent patients are more likely to have increased total cholesterol (d). Paradoxically, patients with high PDC are also at higher risk for MI (e) and coronary artery disease (f). This is likely due to increased adherence in response to events relating to those diseases.*

3.4.3 Analysis Approach. Based on our study design, there is a clear threshold date – the date of first statin prescription. The character of the data (particularly which codes were collected) changes at this date. Time 0 was set to this threshold date for each patient independently; all predictive models and analyses of statistical association use data from before the threshold to predict occurrences after.

3.4.4 Hypothesis Testing. Independent tests of association between predictors and either PDC or an indicator of first refill produced significant but clearly unexpected results (Figure 16). Specifically, all strongly significant results are in the direction of promoting adherence. This suggests that some bias exists in the dataset such that patients who are seen by the health system more often are also more likely to be adherent.

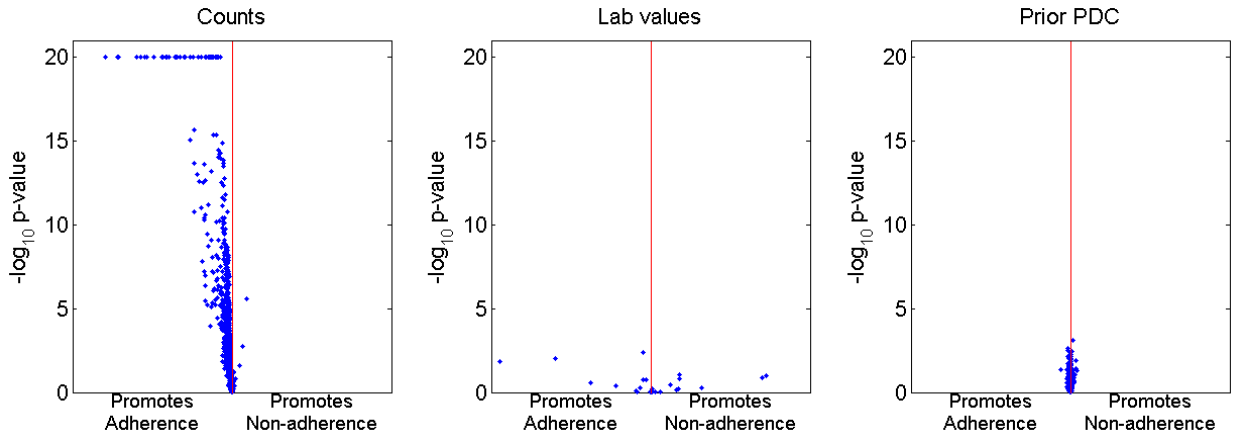

Figure 16. Statistical association between the potential explanatory variables. *y-axis shows -log10 p-value and x-axis shows the effect size (t-test, change in mean). Effects are distorted because we didn't control for important confounding variables.*

To correct for this potential bias, we reanalyzed all independent predictors while simultaneously controlling for the variables in Table 2. This was performed as an independent logistic regression analysis for each predictor variable that always included the control variable set. For these results, the outcome was an indicator of first statin refill, but using PDC as a continuous metric or an indicator of PDC > 80% produced similar results. Results obtained when controlling for demographic variables and comorbidity are shown in Figure 17.

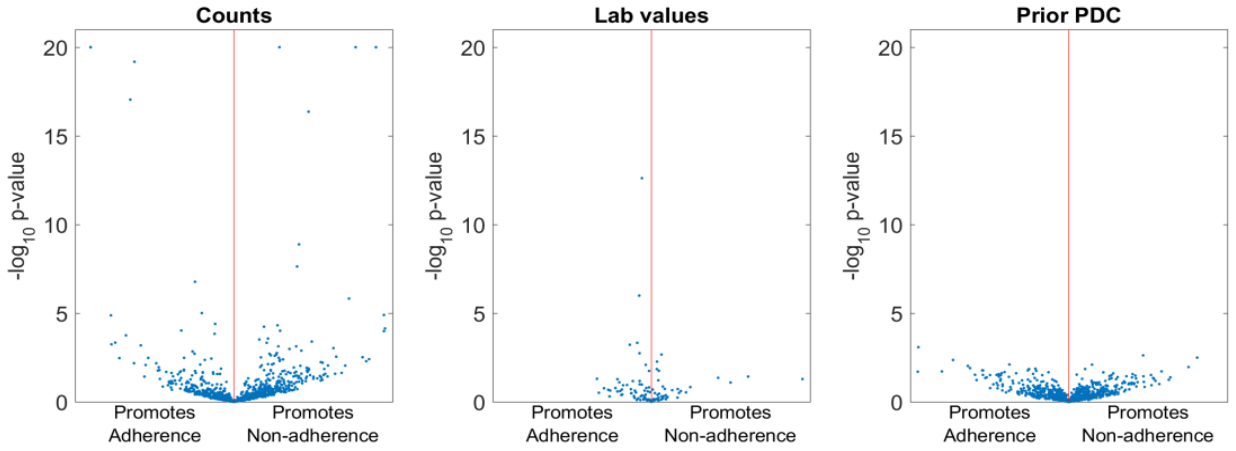

Figure 17. Association between explanatory variables and outcomes are much more reasonable after controlling for design variables listed in Table 2.

3.4.5 Dimension Reduction. Because many codes co-occur on patients' charts, there is a high degree of correlation in the data. By modeling this we have the opportunity to decrease the dimension of the data with which we will generate predictive models. We use non-negative matrix factorization on a matrix indicating the presence/absence of an independent variable for each patient. Let M be the $p \times n$-dimensional indicator matrix. Non-negative matrix factorization represents M as

$$M \approx \beta S$$

where β is a $p \times k$ dimensional matrix of loadings and S is a $k \times n$ dimensional matrix of factor scores. For this application, we tested $k = 15, 30, 40, 50$ for accuracy of prediction and interpretability. The results presented in this report are for $k = 30$ because we found no improvement in accuracy or interpretability for $k > 30$.

Both β and S are relevant to our results. The columns of β are weights across independent variables. By looking at each column, we can identify which independent variables most strongly affect that factor. For each factor we identified those variables that were weighted >10 times more heavily on the factor than they were in the overall counts. We discuss the use of S to predict adherence to statins in section 3.4.7 below.

3.4.6 Modeling Disease Risk. We modeled the likelihood of poor outcomes (defined as a composite of acute MI, stroke, kidney disease, CAD) in the follow-up period using data collected prior to the statin start date. Independent variables are the same as those used for modeling medication adherence (described in detail above). Because of the large number of potential independent variables, we used a two-step approach – dimension reduction by non-negative matrix factorization followed by classification using outcome tree regression on the reduced dimension variables. The dimension reduction approach is intended to collapse similar variables into meaningful groups and thereby minimize overfitting of the risk model. The regression approach is intended to borrow strength across similar outcomes; this leads to results that are more robust when the outcomes are related – particularly in the case of rare events.

We used 30-fold cross-validation to assess model accuracy. Classification results for an indicator of event occurrence within 3 years post-statin start are shown in Figure 18. Experimentation showed no improvement in accuracy for Cox proportional hazard models. Kaplan-Meyer survival curves for tertiles of risk based on the same model are shown in Figure 19.

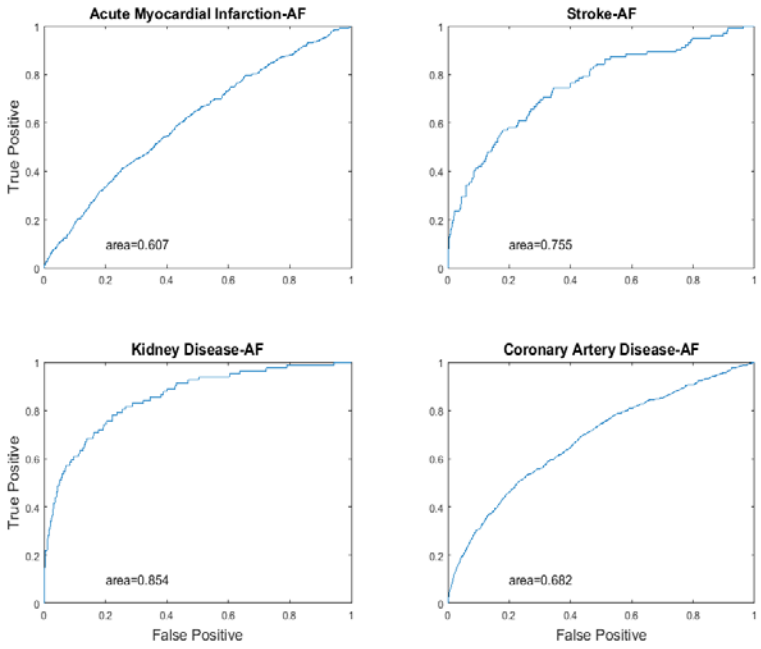

Figure 18. ROC curves showing predictive accuracy for the factor-regression model designed to predict occurrence within 3 years.

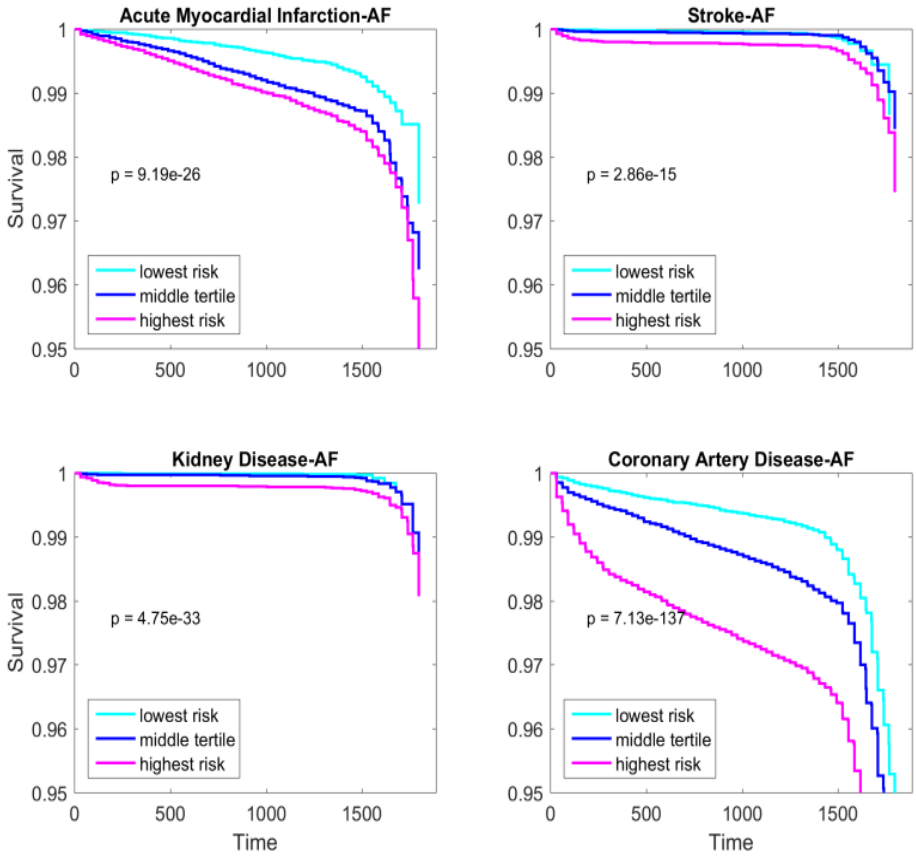

Figure 19. Kaplan-Meier curves for each of three risk tertiles.

3.4.7 Modeling Adherence. After dimension reduction (see section 3.4.5 above), we obtain 30 independent variables for use in predictive modeling. We tested for association between each of these independent variables and adherence using regression to control for all variables in Table 2 as well as the overall disease risk variable – obtained as described in section 3.4.6. The resulting significance of associations for each factor is shown in Figure 20.

We used random forests to build models to predict adherence to statin therapy (PDC > 0.8) (Figure 21). Because random forests are highly susceptible to overfitting, we used 30-fold cross-validation to generate estimates of accuracy for out-of-sample prediction. The c-statistic for this predictive model is 0.736.

3.4.8 Prediction with First Refill Behavior. An obvious approach to predicting medication adherence involves waiting to see whether the patient purchases the first medication refill. Indeed, as can be seen in Figure 22, PDC after the first prescription ends is strongly associated with first refill purchasing behavior. Using first refill behavior alone, we have a sensitivity, specificity, positive predictive value, and negative predictive value of 0.96, 0.44, 0.14, and 0.99, respectively. To assess whether our risk model adds to our ability to predict adherence above and beyond first refill behavior, we regressed PDC > 0.8 on our previously developed risk score and an indicator of whether a first refill was purchased. For this model, we recomputed PDC for a time window beginning after the first prescription had completed.

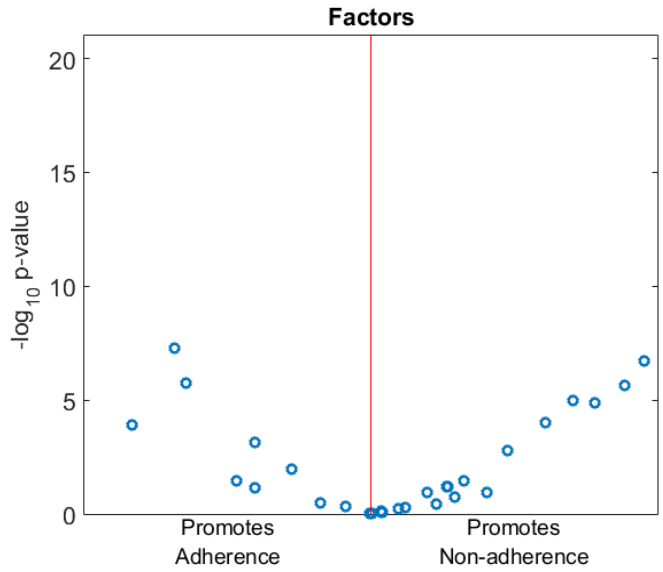

Figure 20. Association between factors and adherence. *A volcano plot showing the p-value (y-axis) and effect size (x-axis) for association between factors and statin adherence behavior.*

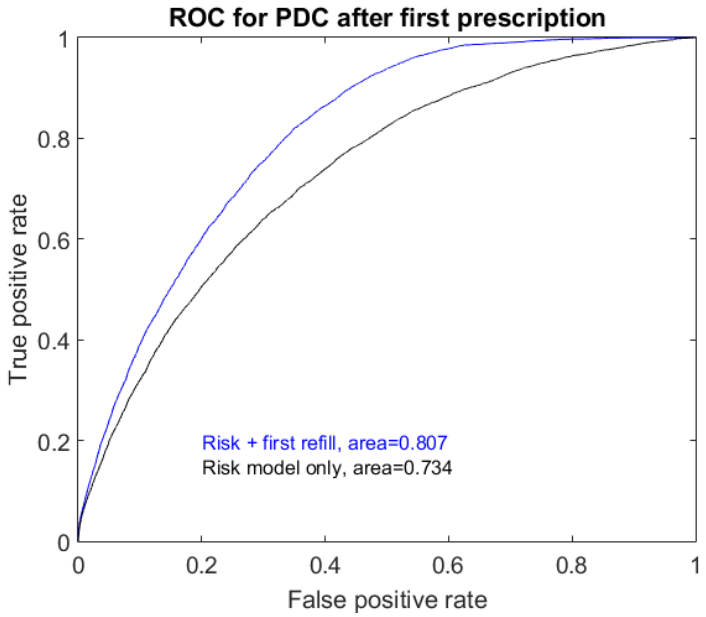

Figure 21. ROC for statin adherence. *Model is random forest using design matrix, disease risk predictor, and dimension-reduced predictors. Outcome is PDC > 0.8.*

The resulting risk score has a c-statistic of 0.81; the ROC curves for this model and for the risk model alone are shown in Figure 21. If we fix the sensitivity to be equal to that obtained from first refill behavior alone, we find an increase in specificity of 1.7% with the other statistics remaining the same. An additional advantage of this approach is the availability of a continuous risk marker. This allows the stratification of risk even in patients with consistent first refill behavior. If we are in a situation in which a health system is planning to intervene on at-risk patients but has limited resources, the availability of a more fine-grained risk stratification can be employed to appropriately target a smaller population.

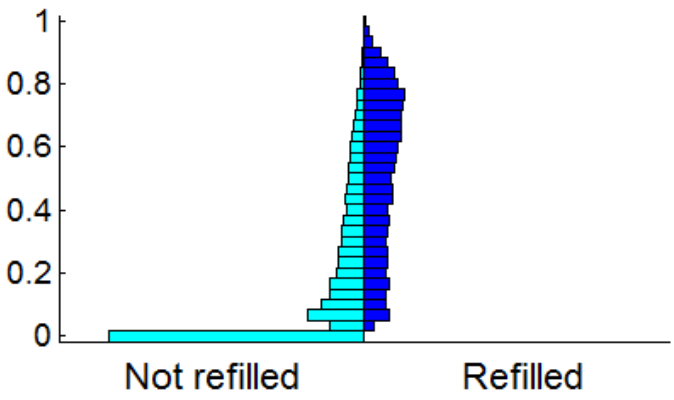

Figure 22. First refill versus PDC. *For this figure, PDC is computed during the time period after the first prescription runs.*

3.4.9 Statin Adherence and Disease Risk. Although statins directly affect lipid levels – particularly LDL levels – the clinical reason for prescribing statins is to prevent arterial disease and its associated clinical outcomes. We tested (Cox proportional hazards) for association between our predicted "risk" of statin adherence (risk is in quotes because we are using the term to predict good behavior rather than bad) and each of the four diseases in Figure 23. We controlled for all variables in Table 2 along with the disease risk variable derived in section 3.4.6. Resulting association with events is shown in Table 3.

As expected, the disease risk variable is strongly associated with the observed disease outcomes. Differences in p-values reported in Table 3 and those shown in Figure 19 are because we are controlling for variables in Table 2 in this analysis. The effect size for predicted statin adherence is in the correct direction, although it is only statistically significant in the case of kidney disease and CAD.

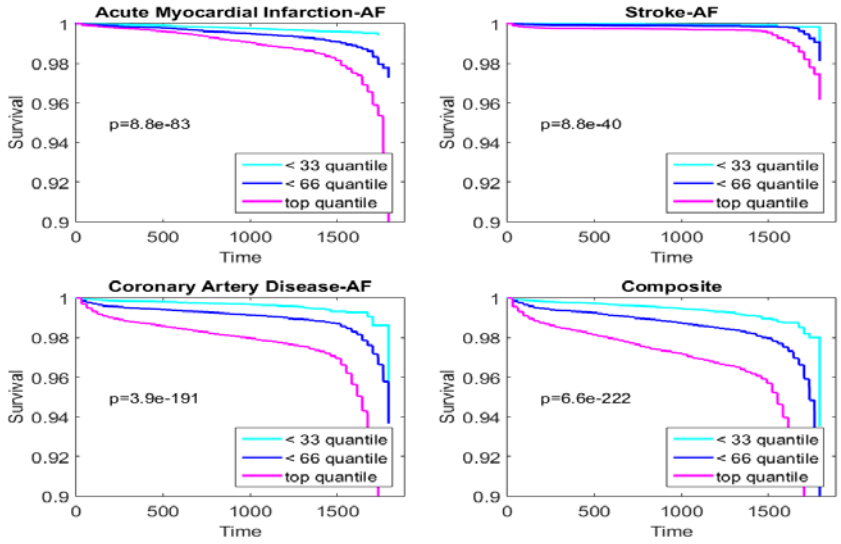

Figure 23. Full model of outcomes. *Results from a Cox survival model incorporating disease risk, design variables, and nonadherence risk.*

Table 3. Predictor of Statin Adherence is Associated with Better Outcomes

Outcome	Predicted Disease Risk Effect		Predicted Statin Adherence Effect	
	Size	p-value	Size	p-value
Acute MI	0.20523	0.00027274	-0.21651	0.00025274
Stroke	0.22133	0.048973	-0.26024	0.01874
CAD	0.3956	0.12E-17	-0.20765	1.13E-06
Composite	0.30474	3.40E-16	-0.1717	1.87E-06

3.4.10 Statin Adherence and Changes in Cholesterol. The direct effect of statins is to lower LDL cholesterol levels. We tested (Pearson correlation) for an association between the adherence risk score and (1) change in LDL cholesterol and (2) change in total cholesterol. The correlations are 0.14 and 0.15, respectively; both p-values are < 1e-20.

3.5 Conclusions

We demonstrate that a robust statistical model using readily available electronic medical record data can be used to predict meaningful outcomes related to statin adherence in the military healthcare population. In addition to this model outperforming using 1st refill as a proxy for future nonadherence, this model also identifies individuals who will not receive the cardioprotective effects of statins, presumably because they are not adhering to statin therapy. Future directions will involve improvement of the predictive model by (1) consideration of additional data sources (i.e., text, medical imaging, or geospatial data), (2) novel statistical method development tailored to these types of data and external validation of the model using contemporary MHS data or data from alternate healthcare systems (e.g., Duke), and (3) implementation of the predictive model within an electronic health system such that an "alert" can be provided at the point of statin prescription to the patient and his/her provider that statin adherence may be challenging in the future. Existing interventions that improve statin adherence can then be deployed in a targeted and preemptive fashion to improve adherence, lower cholesterol, and reduce cardiovascular risk.

4.0 REFERENCES

1. Kiani J, Imam SZ. Medicinal importance of grapefruit juice and its interaction with various drugs. Nutr J. 2007; 6:33.
2. Afonso NM, Nassif G, Aranha AN, DeLor B, Cardozo LJ. Low-density lipoprotein cholesterol goal attainment among high-risk patients: does a combined intervention targeting patients and providers work? Am J Manag Care. 2006; 12(10):589-594.
3. Horne R, Weinman J, Hankins M. The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. Psychol Health. 1999; 14(1):1-24.
4. Poquet N, Lin C. The Brief Pain Inventory (BPI). J Physiother. 2016; 62(1):52.
5. American Thoracic Society. SF-12 health survey. 2008. [Accessed 1 Dec 2016]. Available from http://qol.thoracic.org/sections/instruments/pt/pages/sf12.html.

LIST OF ABBREVIATIONS AND ACRONYMS

BMQ	Beliefs about Medications Questionnaire
BPI	Brief Pain Inventory
CAD	coronary artery disease
DGMC	David Grant USAF Medical Center
EMR	electronic medical record
GGST	genotype-guided statin therapy
HCPCS	Healthcare Common Procedure Coding System
HDL	high-density lipoprotein
LDLc	low-density lipoprotein cholesterol
MHS	Military Health System
MI	myocardial infarction
MMAS	Morisky Medication Adherence Scale
MPR	Medication Possession Ratio
PDC	percent days covered
PI	principal investigator
SD	standard deviation
VLDL	very low-density lipoprotein