2022

TRAJECTORIES OF MEDICATION NON-ADHERENCE WITH TIME-VARYING PREDICTORS AND ASSOCIATION WITH HEALTH OUTCOMES: A COMPARISON OF CLASSICAL STATISTICAL METHODS WITH MACHINE LEARNING ALGORITHMS

Vasco Pontinha
Virginia Commonwealth University

TRAJECTORIES OF MEDICATION NON-ADHERENCE WITH TIME-VARYING PREDICTORS AND ASSOCIATION WITH HEALTH OUTCOMES: A COMPARISON OF CLASSICAL STATISTICAL METHODS WITH MACHINE LEARNING ALGORITHMS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

by

Vasco Miguel Pontinha
PhD Candidate 2022
VCU School of Pharmacy
Department of Pharmacotherapy and Outcomes Science

Main Advisor:
David A. Holdford, R.Ph., M.S., Ph.D., FAPhA
Professor
Department of Pharmacotherapy and Outcomes Science

Virginia Commonwealth University
Richmond, VA
June 2022

Acknowledgments

In (optimistically, many) years from now, I can only hope to have lived a life worthy of the faith others have put on me. That I have inspired as much as those who have inspired me until today. That I have cherished, celebrated, and loved as much as I have been loved until this very same day today. As I play the last few years in my head, I feel ever so fortunate to be surrounded by people that have embraced me beyond any familial bond.

No measure is adequate enough to quantify the debt of gratitude I have for Dr. David Holdford. Since taking me in as a graduate student, Dr. Holdford taught me how to think like a scientist, the beauty of the scientific process, and how to be meticulous with my words so I can reach the audience. His relentless advocacy and mentorship provided me access to a wide variety of research, teaching, and service opportunities. Because of him, I have found my voice as a scientist and a hopeful academic. I truly enjoyed the countless meetings, all the advice, edits, laughs, and learning about the world's smallest violin. I can only hope to keep him in my corner, as a mentor and a great friend.

Ever since I moved to the United States, Dr. Karen Farris has been a constant in my life. From providing life advice, how to navigate the intricacies of American academia, and co-mentoring this research project, Dr. Farris has indelibly marked my life. Without her support, this research project and the next steps that hopefully are in the horizon would simply not be possible. This dissertation is a testament to her mentorship, friendship, and guidance.

I am truly honored to have been taught by Dr. Norman Carroll and be a witness of his wit, raised eyebrow whenever someone uttered something questionable, and a sheer joy of students successes. His Pharmacy Benefits Management course provided such a candid perspective into

the American health care industry. This dissertation stems from the multiple discussions held in the classroom.

Dr. Julie Patterson, who introduced me to the Health and Retirement Study, has been unwavering in her support during the last four years. Her love for research, teaching, and ability to tackle complicated problems are examples I will do my best to follow in the next steps of my professional life. Without her support, I would not have become familiar with group-based trajectory modeling, or how to use this method in a reasonably skillful way.

I am truly thankful for the countless letters of support I have asked Dr. Dave Dixon. He is a true inspiration in leadership and the professional development of others. I am so fortunate to be the recipient and a close spectator of his mentoring style.

Dr. D'Arcy Mays provided invaluable insight for the study design and analysis plan. I am truly thankful for his contributions and guidance towards the completion of this project.

During the PhD program I made friends for life. Having Purva and Elena as peers for the most part of the program was one of the highlights of the last years. I am truly appreciative of their guidance and friendship. I am honored to be a part of the group that survived McGuire 218, with all its peculiarities. Stephen, Tyler, Rotana, Jessica, and Haya made troubleshooting code and modelling a motivating endeavor, and research group meetings much more enjoyable. My worldwide ISPOR friends made me feel connected despite a raging global pandemic. I am also thankful for the financial support that the PhRMA Foundation provided towards the completion of this project.

A part of what I am is located an ocean away. I am truly thankful for my parents Berta, José, and brother Manuel for providing me with a foundation that emphasized the importance of education and the pursuit of what is good, worthwhile, important, and useful for society. Thank

you for giving me the world as my horizon. During the last decades, a significant number of people having been playing a vital role so that I can trail my own path. My father in law, Aldino Salgado was unique in that aspect. He always expressed a boundless confidence in what I could achieve (to the extent which I sometimes wondered if it was only wishful thinking). I am honored to have met him, been loved, and celebrated by him. I can only hope our family of three and this milestone prove him right. My mother in law, Odete has been an unbending rock of support and I am truly blessed to have her in my life. Rafa, Diogo, Pedro, and João have been my weekly injection of assurance that the life I am given could be larger than any of my wildest dreams.

Lastly, my wife Dr. Teresa Salgado has been relentless and a staunch supporter of all my goals. We have trailed thousands and thousands of miles working to achieve our dreams: we have built a family, a home, and a foundation for our Sebastian. I am so fortunate to go through this experience in this world with you. Thank you for your support and for showing me the value in being persistent. I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

ABM	Andersen's Behavior Model of health services use
ACA	Affordable Care Act
ACEI	Angiotensin-converting-enzyme inhibitors
ADL	Activities of Daily Living
aOR	Adjusted Odds Ratio
ARB	Angiotensin receptor blockers
AUC	Area Under the Curve
BIC	Bayesian Information Criterion
BMI	Body Mass Index
CART	Classification and Regression Trees
CHAMPUS/VA	Civilian Health and Medical Program of the Department of Veterans Affairs
CKD	Chronic Kidney Disease
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CMR	Comprehensive Medication Review
CMS	Centers for Medicare & Medicaid Services
DPP-IV	Dipeptidyl-peptidase 4
DRI	Direct renin inhibitors
DUA	Data Use Agreement
ESRD	End-stage Renal Disease
GBTM	Group-based trajectory model
GFR	Glomerular Filtration Rate
GLP-1	Glucagon-like peptide 1
HIPAA	Health Insurance Portability and Accountability Act
HRS	Health and Retirement Study
IADL	Instrumental Activities of Daily Living
ICD	International Classification of Diseases
IRB	Institutional Review Board
MA	Medicare Advantage
MedRIC	Medicare & Medicaid Resource Information Center
MeSH	Medical Subject Headings
MI	Myocardial infarction
ML	Machine learning
MPR	Medication Possession Ratio
MTM	Medication Therapy Management
NIA	National Institute of Aging
PQA	Pharmacy Quality Alliance
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QBP	Quality Bonus Payments

RDA	Restricted Data Agreement
RF	Random forest
ROC	Receiver Operating Characteristic
SA	Specific Aim
SLGT2	Sodium-glucose Cotransporter-2
TEFRA	Tax Equity and Fiscal Responsibility Act

ABM	Andersen's Behavior Model of health services use
ACA	Affordable Care Act
ACEI	Angiotensin-converting-enzyme inhibitors
ADL	Activities of Daily Living
aOR	Adjusted Odds Ratio
ARB	Angiotensin receptor blockers
AUC	Area Under the Curve
BIC	Bayesian Information Criterion
BMI	Body Mass Index
CART	Classification and Regression Trees
CHAMPUS/VA	Civilian Health and Medical Program of the Department of Veterans Affairs
CKD	Chronic Kidney Disease
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CMR	Comprehensive Medication Review
CMS	Centers for Medicare & Medicaid Services
DPP-IV	Dipeptidyl-peptidase 4
DRI	Direct renin inhibitors
DUA	Data Use Agreement
ESRD	End-stage Renal Disease
GBTM	Group-based trajectory model
GFR	Glomerular Filtration Rate
GLP-1	Glucagon-like peptide 1
HIPAA	Health Insurance Portability and Accountability Act
HRS	Health and Retirement Study
IADL	Instrumental Activities of Daily Living
ICD	International Classification of Diseases
IRB	Institutional Review Board
MA	Medicare Advantage
MedRIC	Medicare & Medicaid Resource Information Center
MeSH	Medical Subject Headings
MI	Myocardial infarction

ML	Machine learning
MPR	Medication Possession Ratio
MTM	Medication Therapy Management
NIA	National Institute of Aging
PQA	Pharmacy Quality Alliance
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QBP	Quality Bonus Payments
RDA	Restricted Data Agreement
RF	Random forest
ROC	Receiver Operating Characteristic
SA	Specific Aim
SLGT2	Sodium-glucose Cotransporter-2
TEFRA	Tax Equity and Fiscal Responsibility Act
TTY	Teletype
VIF	Variance Inflation Factor
VIM	Variable Importance Measure
WHO	World Health Organization

ABSTRACT

TRAJECTORIES OF MEDICATION NON-ADHERENCE WITH TIME-VARYING PREDICTORS AND ASSOCIATION WITH HEALTH OUTCOMES: A COMPARISON OF CLASSICAL STATISTICAL METHODS WITH MACHINE LEARNING ALGORITHMS

By Vasco Miguel Pontinha, MPharm, MA, PhD Candidate

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2022

Advisor:
David A. Holdford, R.Ph., M.S., Ph.D., FAPhA
Professor
Department of Pharmacotherapy and Outcomes Science

Background: Medication adherence is a major obstacle to improving health care outcomes in long-term therapies for chronic diseases. According to the World Health Organization, interventions for improving medication adherence can have a higher impact on the health of the population than any other advance in medical treatments. Approximately 125,000 individuals die every year in the U.S. because of non-adherence to medication, representing societal costs of $100-289 billion. Previous research has successfully used group-based trajectories methods to identify similar longitudinal medication adherence trajectories. However, medication adherence is not an isolated behavior and is influenced by many factors that current interventions fail to confront. This study aims to (1) identify longitudinal trajectories of medication adherence of chronic diseases treated with oral medications, and (2) distinguish the predisposing, enabling, and need characteristics, which have been identified following an Andersen's Behavior Model of health services use theoretical framework. Additionally, this study investigates the association between adherence trajectories membership and *a posteriori* consequences, that was examined

by deploying two alternative predictive methods, one based on classic logistic regression and the other based on machine learning algorithms.

Methods: Participants of the Health and Retirement Study were linked to respective Medicare administrative health care claims between 2008-2016. Group-based trajectory models were used to elicit the number and shape of medication adherence trajectories, among a sample of 11,068 individuals taking hypertension medications, statins, or diabetes medications. Time-fixed and time-varying risk factors were examined using logistic regression and multi group-based trajectory modeling, respectively. The association between medication adherence trajectories and outcomes, including myocardial infarction, stroke, and diabetes-specific outcomes (ophthalmic complications, nephropathy, neuropathy, and peripheral angiopathy) was investigated by comparing logistic regression models with machine learning algorithms based on random forests. The outcomes were identified by respective ICD-9 and ICD-10 codes. Predictive ability of the logistic regression compared to machine learning algorithms was examined using the *c-statistic*.

Results: Group-based trajectory models were estimated for the sample population taking hypertension medications (n=7,272), statins (n=8,221), and diabetes medications (n=3,214). In the hypertension model, three trajectories were identified, following *near-perfect adherence*, *slow*, and *rapid decline* trajectories, accounting for 47.5%, 33%, and 19.5% of individuals in that group respectively. Five trajectories were identified in individuals taking statins, including *near-perfect adherence* (35.5%), *slow decline* (17.1%), *low then increase* adherence (23.6%), *moderate decline* (12.6%), and *rapid decline* (11.2%). The diabetes medications yielded the model with the greatest number of trajectories, including *near-perfect adherence* (24.2%), *slow*

decline (16.9%), *high then increase* adherence (25.1%), *low then increase* (13.8%), *moderate decline* (10.7%), and *rapid decline* (9.3%).

Several socioeconomic factors were identified as predictors of non-adherence trajectories, which typically were indicative of lower socioeconomic status. While this study pioneered the use of multi group-based trajectories to identify time-varying predictors of medication adherence trajectories, no coherent trends were observed in the analysis. Nonetheless, loss of spouse was generally found to occur in parallel with decreases in adherence, or the opposite, in which regaining a spouse was met with increases or maintenance of high adherence.

Overall, based on the *c*-statistic, the logistic regression models exhibited better predictive ability than random forest machine learning algorithms in examining the relationship between medication adherence trajectories and outcomes. All non-adherence trajectories in all three models were found to be more likely to experience myocardial infarction compared to each respective *near-perfect adherence* trajectory. However, the same was not observed for stroke and diabetes-specific outcomes. All declining trajectories of patients taking hypertension medications were more likely to experience stroke. Additionally, only those in the *rapid decline* trajectory of statins model and those in the *slow decline* trajectory of the diabetes medications model were more likely to experience stroke compared to each respective *near-perfect adherence* trajectory. In the diabetes medications model, only patients following declining adherence trajectories (*slow*, *moderate*, and *rapid*) were more likely to experience nephropathy and peripheral angiopathy than those following *near-perfect adherence*. No statistically significant differences were found for ophthalmic complications and neuropathy between the *near-perfect adherence* trajectory and all other non-adherent trajectories.

Conclusions: The GBTM models displayed a nuanced perspective of how participants in the Health and Retirement Study are adherent to their medication for hypertension, statins, and diabetes and how time-varying factors can be investigated to identify patients at risk of falling into non-adherent trajectories. However, non-adherent trajectories are not equally and statistically significantly found be more at risk of health outcomes than near-perfect adherent trajectories. Quality and health policy implications are discussed in light of the results of this research study.

1. INTRODUCTION

1.1 MEDICATION NON-ADHERENCE.

Non-adherence to pharmacotherapy is a major obstacle to improving health care outcomes, especially in long-term therapies for chronic diseases. [1,2] The World Health Organization posits that improving medication adherence would improve health outcomes at a higher rate than any other innovation in health care. [1] It is estimated that approximately 50% of the patients with chronic diseases are non-adherent to their therapeutic plan. 125,000 deaths per year and added annual costs between $100-289 billion are estimated to result from non-adherence to medications in the United States (U.S.). [3-6]

With the passage of the Social Security Amendments in 1965 by the United States Congress, Medicare was established as a federal medical insurance program for people over 65 years old, younger disabled people, and patients in dialysis. [7] Currently, this program provides coverage for benefits such as Hospital Insurance (Part A), Medical Insurance (Part B), and prescription drugs (Part D). [8] Alternatively, Medicare beneficiaries can opt to enroll instead in a Health Maintenance Program or Preferred Provider Organization (PPO) that includes coverage for Parts A, B, and D and are commonly known as Medicare Advantage Plans (Part C). [8,9] Since the establishment of the Medicare Part D Prescription Drug Benefit in 2006, the U.S. Centers for Medicare and Medicaid Services has established a public-private partnership with the Pharmacy Quality Alliance (PQA) to conduct research to improve medication safety, adherence, and appropriate use. The PQA endorsed Proportion of Days Covered (PDC) as the preferred measure to quantify medication adherence. [10] In 2007, the Centers for Medicare & Medicaid Services (CMS) developed a 5-star quality rating system aimed at increasing the quality of care and provide accountability for quality and offer beneficiaries information to compare plans on quality

provided to Medicare Advantage plan beneficiaries. [11] After the passage of the Affordable Care Act (ACA) in 2010, the 5-star rating system allowed plans to be financially rewarded by the Quality Bonus Payments based on the quality performance in health plan specific metrics for Part C and prescription drug measures for Part D. Of all the metrics included in the rating system for part D plans, medication adherence is the one weighing the heaviest for the 5-star ratings. Currently, this value-based payment model rewards sponsors of Part D plans for quality metrics pertaining to medication adherence during the last year or last 91 days of the enrollment period for the following medication classes:

a) Diabetes: biguanides, a sulfonylureas, thiazolidinediones, DPP-IV inhibitors, incretin mimetics, meglitinides, or SGLT2 inhibitors.
b) Hypertension: Angiotensin converting enzyme inhibitors (ACEI), angiotensin receptor blocker (ARBs), or direct renin inhibitors (DRIs).
c) High blood cholesterol: statins.

Traditionally, researchers and practitioners dichotomously classify patients as adherers (PDC≥80%) or non-adherers (PDC<80%) over a given period of time, assuming PDC≥80% as the optimal threshold for generating positive health outcomes. [12] However, this approach appears to be insensitive to temporal changes in patient behavior and blind to the rest of the year of coverage. Moreover, medication adherence has been shown to be influenced by a multitude of factors that extend way beyond the mere ability to follow doctors' orders. These include socioeconomic-, health-system-, therapy-, condition-, and patient-related factors. [13] As a function of the combination of several factors that inexorably can change with time, this research study proposes examining medication adherence as a dynamic behavior.

The combination of the different factors influencing medication adherence inherently results in a diverse patient experience, and likely, patient health outcomes, especially at the initiation phase of pharmacotherapeutic treatment plan. As Seixas and colleagues put it, the factors that influence medication adherence can be categorized in two groups, those that are non-modifiable and those that are modifiable. [14] Cognizant of the several factors that can influence medication adherence behavior, several social-behavioral frameworks have been proposed to date. [15-26] One way or another, all these models recognize that there are individual features, contextual factors, and environmental characteristics that influence medication adherence. From a developmental behavioral perspective, the Andersen's Behavioral Model of health services use (ABM) is the one that seems to best encapsulate predisposing individual characteristics, needs of the patients, as well as factors pertaining to the environmental, social, health care provider and health-system characteristics. [25] The ABM theory appears to resonate well with the socioeconomic-, health-system-, therapy-, condition-, and patient-related factors that influence adherence. [1] Moreover, it is possible that some of these factors may be modifiable over time, or by means of interventions, which suggests the fluidity of medication adherence behavior. Therefore, medication adherence should be investigated using a method that allows the characterization of a behavior longitudinally, allowing practitioners to evaluate optimal evolutions of improvement in medication taking behavior.

Human behavior research[27] such as disease modeling is based on the notion that behavior or a disease is the result of continued exposure to risk factors and once onset, the disease follows a natural progression. [28,29] For example, hypertension is not established after a single meal with excessive content in salt inexorably conducive to a deadly outcome. Instead, it is the continued

use and exposure to risk factors that results in the establishment hypertension, that with time can result in myocardial infarction (MI), or stroke. [30]

Seemingly, behaviors like medication adherence are not established overnight, which makes them ideal candidates for an analysis conducted from a developmental perspective. Investigating medication adherence behaviors from a developmental perspective will potentially allow pharmacy practice researchers to understand how patients engage with their medication and pharmacotherapeutic treatment plan and distinguish *lifetime developmental trajectories* of medication use. While value-payment schemes such as the Quality Bonus Payments indirectly tie medication adherence performance with health outcomes, it is possible that patients with fluid medication adherence behavior (sometimes perfect, sometimes less than) have comparable outcomes like the ones perfectly adherent over time.

This research study adds to existing literature on medication adherence by eliciting longitudinal developmental trajectories of medication adherence to pharmacotherapy for diabetes, hypertension, and high blood cholesterol, investigating the modifiable and non-modifiable predictors of those trajectories, and examining the relationship between medication adherence trajectories and clinical outcomes.

Group-based trajectory modeling was used in this research study to analyze administrative health care claims of Medicare beneficiaries and elicit medication adherence trajectories. The Medicare beneficiaries were identified via participation in the Health and Retirement Study (HRS), a longitudinal panel study with a representative sample of approximately 20,000 people in America sponsored by the National Institute of Aging (grant number NIA U01AG009740) and conducted by the University of Michigan. The survey data was used to inform the modifiable and non-modifiable factors that can influence medication adherence trajectories. Finally, the

relationship between medication adherence trajectories and health outcomes was examined by deploying classification methods. More specifically, we compared the predictive value of logistic regression and machine learning algorithms such as random forest and boosted random forest algorithms.

The study rationale, specific aims and significance are provided in the next sections of Chapter 1. The background and a systematic literature review on medication adherence and group-based trajectory modeling are provided in Chapter 2. The methods followed in this research study are explained in chapter 3, while the results from the data analysis are presented in chapter 4. In chapter 5, we discuss the results of the research study, influence of limitations in the results, overall conclusions, and directions for future research.

1.2 STUDY RATIONALE AND SPECIFIC AIMS

To overcome the high financial and societal burden of medication non-adherence, this research study enhances the current methods used to describe longitudinal trends in medication adherence and discusses whether current value-based payment schemes that reward improvements in medication adherence truly reflect improvements in patient outcomes. Additionally, this study compares existing optimal thresholds of medication adherence with longitudinal non-adherent trajectories to discuss if optimal levels of adherence are warranted and whether these thresholds are medication or disease specific. The fluid nature of medication adherence makes it hard to identify which patient- and context-specific characteristics are predictors of changes in medication adherence and what is the optimal medication adherence behavior. This research project addresses these two challenges by deploying an innovative methodological approach to identify both longitudinal trends of medication adherence behaviors

and longitudinal covariates responsible for changes in medication adherence, while also investigating the association of medication adherence trajectories and outcomes to establish optimal non-adherent trajectories.

The specific aims (SA) of this research study are summarized as follows:

Specific Aim 1: Trajectories group model in chronic patients initiating oral-dosage forms

- Specific Aim 1a: Estimate a group-based trajectory model to identify the number of medication adherence trajectories in patients initiating pharmacotherapy for different diseases:
    - Diabetes (biguanides, a sulfonylureas, thiazolidinediones, DPP-IV inhibitors, incretin mimetics, meglitinides, or SGLT2 inhibitors)
    - Hypertension (Angiotensin converting enzyme inhibitors (ACEI), angiotensin receptor blocker (ARBs), or direct renin inhibitors (DRIs))
    - High blood cholesterol (statins).
- Specific Aim 1b: Compare the proportion of patients traditionally considered as adherers and non-adherers with the medication adherence trajectories identified in SA1a.

Specific Aim 2: Identification of time-fixed and time-varying predictors of medication adherence trajectories

- Specific Aim 2a: Identify the modifiable and non-modifiable determinants of medication adherence that are associated with medication adherence trajectory membership.
- Specific Aim 2b: To identify the concurrent trends of the time-varying risk factors that are associated with medication adherence trajectory membership.

- Specific Aim 2c: Build a data visualization tool displaying the evolution of the time-dependent predictors and the medication adherence trajectories.

Specific Aim 3: Predictive model linking medication adherence trajectories to health outcomes:

- Specific Aim 3a: Examine the relationship between medication adherence trajectories and health outcomes, by comparing two classification methods: logistic regression and random forest algorithms.
- Specific Aim 3b: Determine predictive ability by comparison of the c-statistic to identify the best predictive model and examine the strength of association between medication adherence trajectories and outcomes, including myocardial infarction and stroke for all models, and ophthalmic complications, nephropathy, neuropathy, and diabetic peripheral angiopathy as diabetes-specific outcomes.

2. BACKGROUND AND LITERATURE REVIEW

2.1 VALUE-BASED PROGRAMS IN MEDICARE

The movement towards enacting policies that transform how health care is paid for is inextricably related with the rise in health care costs in the United States. Today, more than 80 million Americans rely on federal programs that finance their health care benefits. [31] These programs include Medicare, Medicaid, and the Children's Health Insurance Program, administered under the U.S. Department of Health and Human Services. Medicare and Medicaid alone account for more than 1.5 trillion dollars in 2020, representing a 98% increase since 2008. [32] Figure 1 displays the growth and annual percent change in Government-sponsored expenditures pertaining to the Medicare and Medicaid programs.

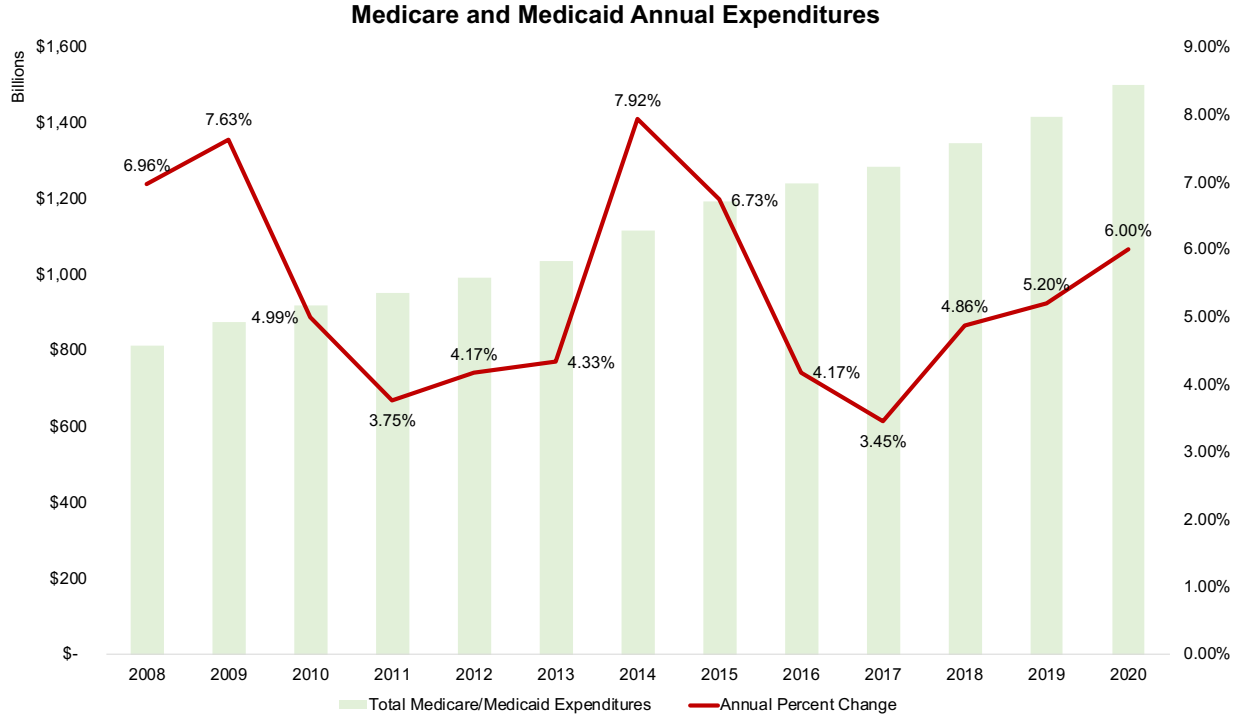

Figure 1 - Historical data of the Federal Government-sponsored Expenditures towards Medicare and Medicaid programs

Despite the increase in the population over 65 years old bound to become eligible for Medicare, the rate of increase in spending in Medicare funded by the Federal Government has

been trending downwards in the past decade, despite a slight increase since 2018 (Figure 2). [31,33] It appears that this downward trend occurred right after significant reforms in health care.

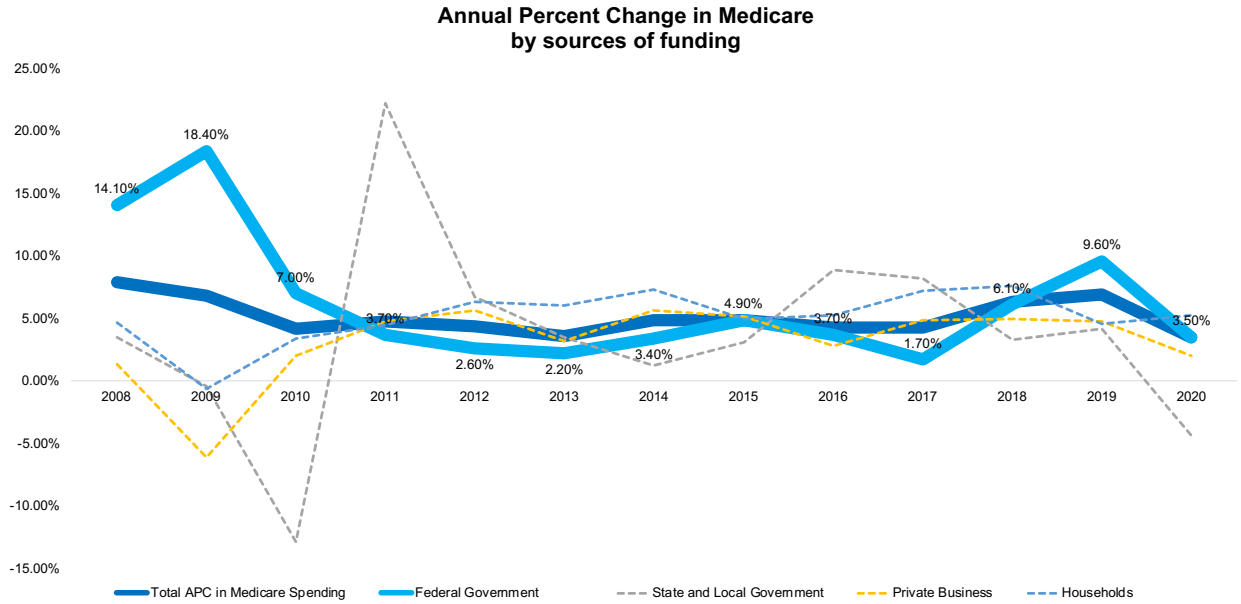

Figure 2 - Annual percent change in the Medicare spending and respective sources of funding

Since the passage of the Affordable Care Act and the impetus provided by the Obama administration, CMS introduced a series of policies aimed at transforming a traditional Fee-For-Service (FFS) into a value-based purchasing of health care services, while also increasing the quality of service provided and expanding the number of insured individuals in the United States. [34-36] A subset of those policies and strategies focused specifically on prescription drug payment reform and establishment of well-defined quality standards. The key strategic imperatives referred to "*monitor*[ing] *and assess*[ing] *the quality of health care services in MAOs* [Medicare Advantage Organizations] *and Part D sponsors*" and "*provid*[ing] *incentives for MAOs and for Part D sponsors for improving and/or excelling on quality assessments*". [35]

Traditionally, the Medicare program allowed beneficiaries to enroll in health care insurance plans that include Parts A and B, which loosely refer to hospital and medical insurance respectively. [37,38] However, the Tax Equity and Fiscal Responsibility Act (TEFRA), allowed

Medicare to create what is now known as Medicare Part C or Medicare Advantage (MA) plans. [39] The policy goal was to bundle Parts A and B and, if willing, Part D (i.e., drug benefits) so that enrollees benefit from reduced cost-sharing expenses. In turn, MA plan sponsors would contract a risk-sharing agreement with Medicare, who would pay them a predetermined risk-adjusted amount for each covered beneficiary. [40] While the creation of MA plans resulted in expanded choice and enrollment, Part C also ended up costing significantly much more than traditional Medicare plans. [41] Evidence suggested that the payment benchmarks in MA plans were flawed and did not promote gains in efficiency when compared to traditional Medicare plans. [42] Thus, when the ACA passed in 2010, the payment rates for MA plans saw a significant decrease in the reimbursement rates in order to reduce overpayment. [43] However, the ACA also included the opportunity for MA plan sponsors (including those offering Part D coverage) from benefiting from Quality Bonus Payments (QBPs).

2.2 MEDICARE PART C AND D STAR RATINGS

The QBPs were established as an additional revenue source for MA and Part D plan sponsors based on predefined quality metrics. Simultaneously, plan sponsors would be rated on a 1–5-stars scale, in what is known as the Medicare Part C and D Star Ratings (Star Ratings). The ACA was crucial to factor the Star Ratings into payments to plans, by mandating QBPs to all contracts earning 4 or higher in the Medicare Star Ratings Program. [44] In addition to requiring plan sponsors to report quality metrics, CMS also publishes the Star Ratings in a consumer-friendly fashion to help beneficiaries evaluate the relationship between quality and cost of MA and Part D plans as active health consumers (Figure 3). [45]

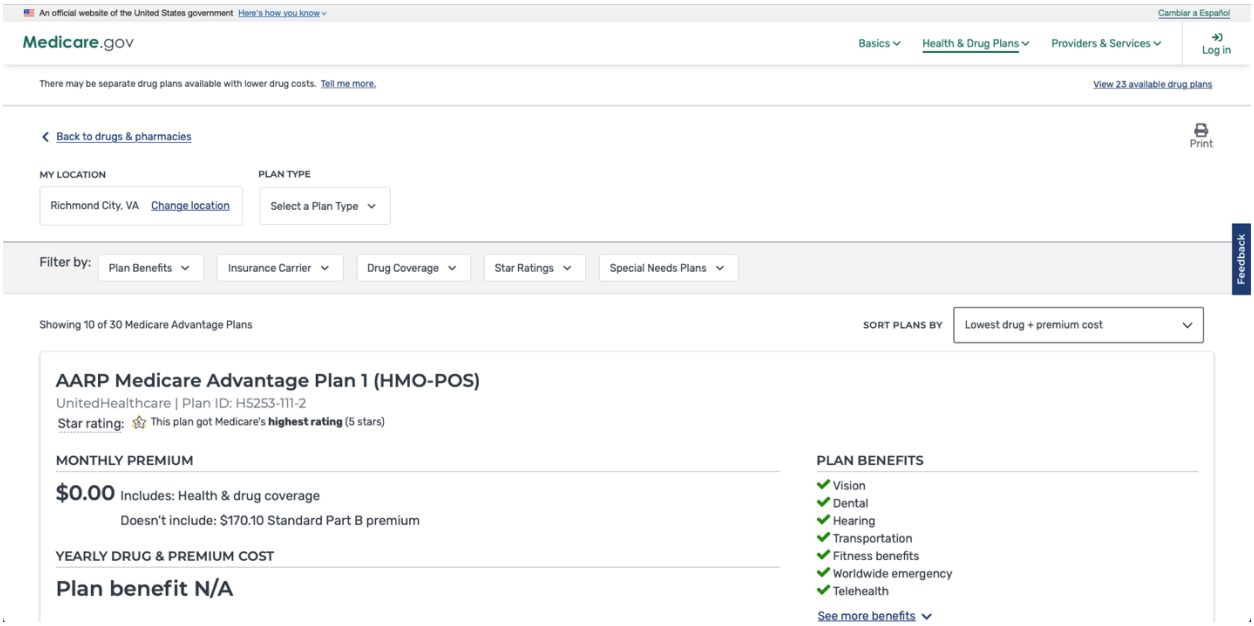

Figure 3 – Example of the Medicare Advantage Plan Selection on Medicare's Website

The quality metrics associated with the Star Ratings program include medical services measures for MA plans and medication-related measures for MA prescription drug and Medicare Part D plans. Table 1 lists the quality measures utilized in the calculation of the Star Ratings as of 2022 and national average ratings. [46] In general, the aspects evaluated in these measures pertain to provision of medication therapy management within comprehensive medication review services, adherence measures for diabetes, hypertension, and high blood cholesterol medication, level of satisfaction with the prescription drug plan, and beneficiary-stated willingness to change prescription drug plan. Overall, the measures associated with medication adherence account for approximately 33% of the aggregate rating of the plan. Simply put, the adherence scores are calculated based on the proportion of patients considered adherent to the medication. [47,48]

Table 1 - Quality measures included in the Medicare Prescription Drug Star Ratings

Measure	2022 National Average Score	
	MA-PD Plans	PDP Plans
MTM Program Completion Rate for CMR†	83.35	53.74
Medication Adherence for Cholesterol (Statins) †	86.24	87.08
Medication Adherence for Diabetes Medications†	86.03	86.68
Medication Adherence for Hypertension†	87.04	88.45
Statin Use in Persons with Diabetes†	82.86	80.34
Rating of Drug Plan†	86.43	84.15
Getting Needed Prescription Drugs†	91.05	90.56
Complaints about the Plan*	0.21	0.06
Call Center – Foreign Language Interpreter and TTY availability†	91.02	88.71
Members Choosing to Leave the Plan**	14.68	10.65

MA-PD – Medicare Advantage Prescription Drug; PDP – Medicare Part D Prescription Drug plan; MTM – Medication Therapy Management; CMR – Comprehensive Medication Review; TTY – teletype

* The lower the score the better for Complaints about the plan and Members choosing to leave the plan

† Measure scored from 0-100, the higher the better

Since the establishment of the Medicare Part D Prescription Drug Benefit in 2006, the U.S. Centers for Medicare and Medicaid Services has established a public-private partnership with the Pharmacy Quality Alliance (PQA), a non-profit organization aiming to conduct research to improve medication safety, adherence, and appropriate use. [49] As a national quality organization, the PQA endorsed Proportion of Days Covered (PDC) as the preferred measure to quantify medication adherence. [10] PDC is calculated by dividing the number of covered days by the number of days in the measurement period and discounting any days that might have overlap between prescription fills. CMS considers drug coverage as having at least one drug of the target drug class during the enrollment period. [48] Beneficiaries are only considered in the Star Ratings analysis if the first fill of the target drugs occurred at least 91 days before the end of the enrollment period. [48] Therefore, the number of days in the measurement period is at least 91 days

or, if the patient has already initiated treatment before the enrollment period, 365 days. In conclusion, CMS dichotomizes beneficiaries' medication use behavior into adherent or not adherent irrespective of the time-period using the 80% threshold for medications for hypertension (renin-angiotensin system agonists), diabetes (except insulin), and for high blood cholesterol (statins). From a practical standpoint, this approach is problematic if the goal of the quality metric and respective QBP is to reward health care professionals for addressing the underlying causes of medication non-adherence. Let us consider two scenarios of patients filling prescription within a 90-day period (Figure 4). Both patients' PDC would be 66% (both had two fills of 30-day supply within a 90-day period). However, it is possible that Patient 1 interrupted his medication in the second month because they could not afford the out-of-pocket cost of the drug, while Patient 2 simply discontinued the medication because of failure to achieve treatment goals. Armed with the context and pattern of adherence, the pharmacist could work with the patient and prescribing physician to transition Patient 1 to a drug within the same drug class that the patient can afford. However, the current measurement system does not reward potential health care interventions because the PDC calculation does not provide the overall pattern and evolution of medication adherence.

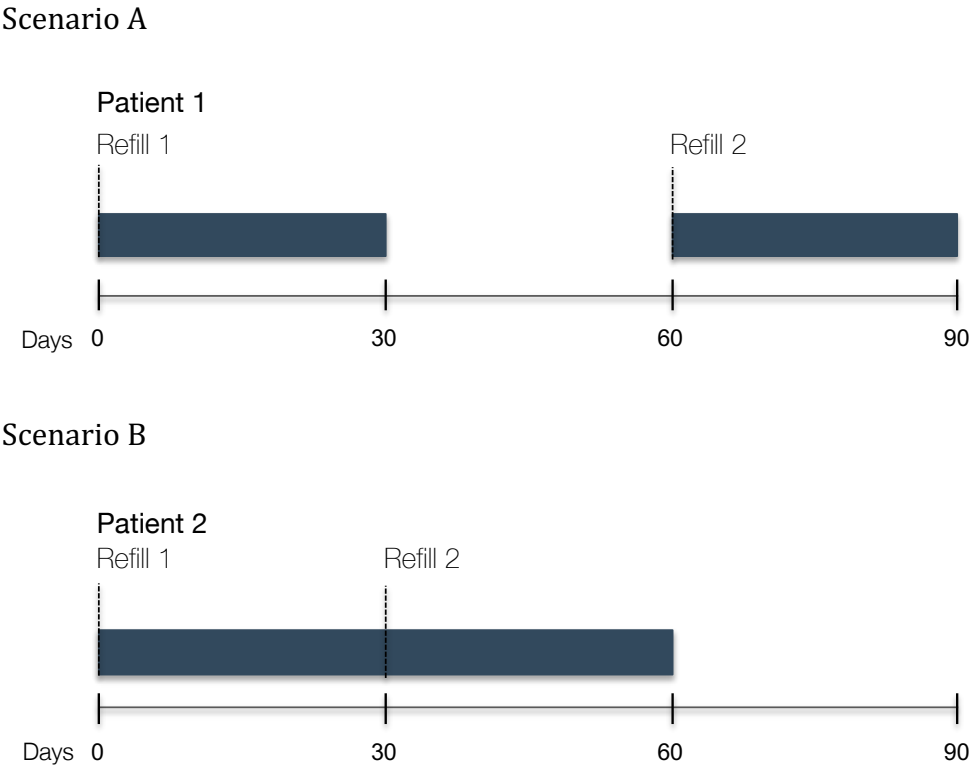

Figure 4 - Simulation of prescription fills of two patients within a 90-day period

PDC calculations are a validated proxy-measurement for the behavior of medication adherence. [50] However, several studies and systematic reviews point out lingering questions about establishing an optimal threshold medication adherence. [51-53] Moreover, the current time-fixed approach to calculating adherence ignores temporal patterns and does not consider the time it takes for interventions to generate their results. [54,55]

Group-based trajectory modeling (GBTM) is a statistical approach that examines a specific phenomenon, an outcome that is quantifiable, regardless of being a behavior, biologic, or physical outcome over age or time. [56,57] As Nagin put it, "understanding the developmental trajectories is among the most fundamental and empirically important research topics in social and behavioral sciences and medicine". [56] This research work posits using GBTM to examine the

medication adherence behavior as a better approach to establishing optimal medication trajectories as potential indicators of different health outcomes.

2.3 GROUP-BASED TRAJECTORY MODELING (GBTM)

GBTM is relatively new methodological approach increasingly used in the social sciences. This method elicits common longitudinal trajectories of a particular outcome within in a population over a period of time. This method is based on finite mixture models, which assumes that there is a finite number of unobserved groups in any given sample. The primary assumption of GBTM is that any population of *n* individuals can be grouped in a maximum of *n* groups. In other words, a population cannot be grouped by any feature in more groups than the total number of individuals in that population. In essence, the GBTM assumes that in a sample, there is a determined number of trajectories (groups) that take up the shape of polynomial functions. The number of those trajectories are estimated initially by examining the goodness of fit measures – such as the Bayesian Information Criterion (BIC). GBTM is a specialized application of the maximum likelihood function. The number of trajectory groups are estimated by maximizing the probability distributions (maximum likelihood) for each trajectory group based on the best possible goodness of fit combining single-group models within a common multiple-group model structure. [56,57] In simpler terms, GBTM clusters individuals with seemingly similar behavior over time in probabilistic trajectories. Each trajectory (*j*) suggests a latent variable, which can be interpreted as the probability of individual *i* to engage in the behavior of interest (P(y_{it}^*) at a specific Age *t*. This trajectory definition is generally referenced to as the "tobit regression" (Equation 1). [58]

$$y_{it}^* = \beta_0^j + \beta_1^j Age_{it} + \beta_2^j Age_{it}^2 + \beta_3^j Age_{it}^3 + \varepsilon_{it}$$

Equation 1 - Tobit regression

This method assumes that the shape of a trajectory *j* is described by a polynomial function of time and that individual differences in trajectories are summarized for a finite set of different polynomials functions of time. As such, the probability of an individual to behave a certain way can be described by a simple equation (Equation 2). The unconditional probability of individual *i*'s observation of longitudinal behavioral measurements ($P(Y_i)$ is defined by the sum of the product of all conditional probabilities of Y_i belonging to a given trajectory *j*, $P^j(Y_i)$, and probability of a randomly chosen population member belonging to group *j*, π_j. [56]

$$P(Y_i) = \sum_{j}^{J} \pi_j \, P^j(Y_i)$$

Equation 2 – Unconditional probability of *i* in any group-based trajectory model

Therefore, the likelihood function of the entire sample is defined by the product of the individual likelihood functions of the N individuals that make up the sample. This method is an extension of finite mixtures because a GBTM model sums across a finite number of discrete trajectory groups that are latent in the sample population (Equation 3).

$$L = \prod^{N} P(Y_i).$$

Equation 3 – Maximum Likelihood formula

The behavior (outcome) that can be studied using GBTM can assume several formats. GBTM allows the analysis of outcomes in the form of scales, counts, or binary data. Medication adherence behavior can be measured in a continuous scale (PDC ranges from 0-100%), which results in an ideal behavior to be analyzed with GBTM. [27,59-61] Moreover, this method appears to overcome the limitations of static measures like annual or quarterly PDC calculations. [62,63] The fluid nature of medication-taking behavior and the underlying factors driving that behavior are

not adequately captured with PDC and MPR ratios calculated for large periods alone. Patient adherence may vary over time depending on an individual's employment status, family situation, new disease diagnosis, change in therapy, or any host of other factors. By quantifying adherence with a single ratio, both PDC and MPR collapse a broad spectrum of adherence behaviors into a single number, thereby masking complex, dynamic and longitudinal patterns of behavior.[10,14] For example, a patient with a PDC of 0.6 may have been (1) highly adherent during the early stages of therapy but less adherent as time went on, (2) poorly adherent during the early stages of therapy but more adherent over time, or (3) intermittently adherent throughout the follow-up period.[10] Each of these three patterns of medication taking would be classified as nonadherent using conventional measures and be treated as a homogenous adherence group. Ignoring these underlying differences in the refill patterns and patients' adherence behavior may lead to missed opportunities to address barriers to appropriate medication use.

2.4 MEDICATION ADHERENCE IN THE CONTEXT OF THE ANDERSEN'S BEHAVIORAL MODEL OF HEALTH SERVICES USE

Several previous studies examining the causes of non-adherence were conducted within the theoretical framework of the Andersen's Behavioral Model of Health Services Use (ABM) [64-68]. While it was originally developed to study the family unit and its contextual circumstance to explain the rate of health services use, this model proved to provide both an explanatory and predictive ability of the interaction with health services, including those associated with the use of medicines. [25,64] This model has gone through an iterative process, which precisely added the dimensions of the health care system and external environments and customer satisfaction. In the latest iteration of the ABM framework, Andersen and Davidson recognize the relevance of

individual characteristics as well as the importance of the contextual situation of the patient, both comprising each dimension as follows[25] (Figure 5):

a) **Predisposing factors,** which represent the biological imperatives that suggest the need for health services, such as the demographic characteristics, education, mental factors, including health beliefs and attitudes, broader community support, cultural norms, and political perspectives.

b) **Enabling factors** as the financing and organizational characteristics of source of [health] care, means of transportation, travel and waiting times, affluence, health insurance coverage, cost of goods and [health] services, degree of communication with physician and involvement with treatment plan, health literacy and available information.

c) Need factors as those related with how people perceived their own need for health services, general health, and functional state. It also encompasses the same aspects as evaluated by a health care professional, such as those epidemiological indicators such as comorbidities, disabilities, and mortality.

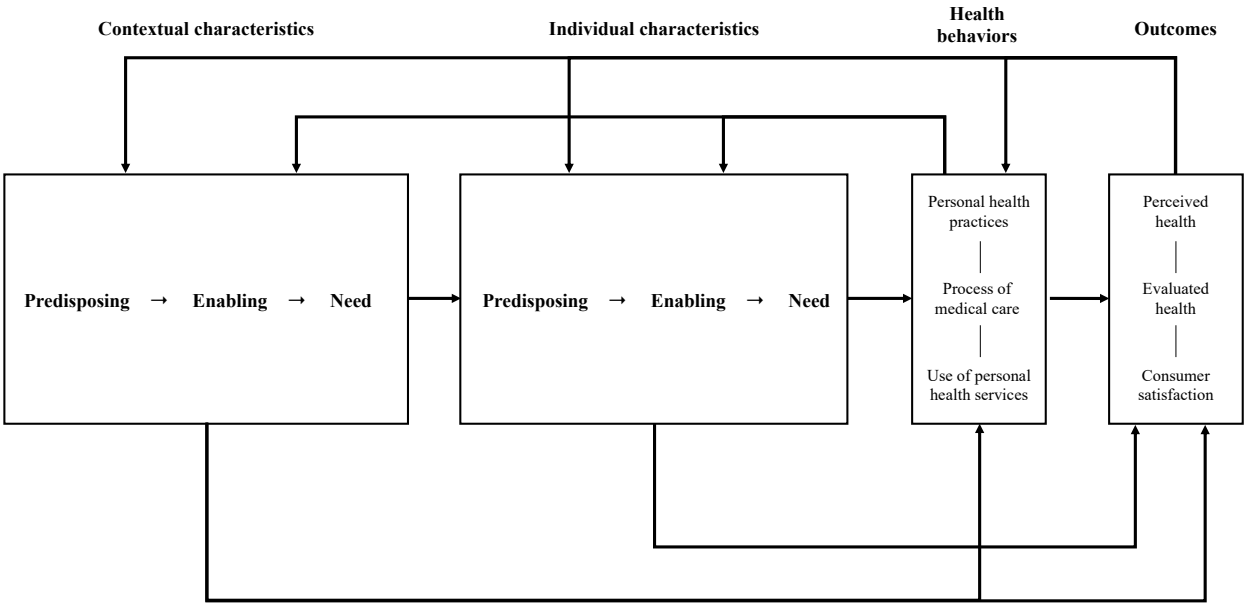

Figure 5 - The Andersen's Behavioral Model of health services use (adapted from Andersen and Davidson[64])

Furthermore, the WHO issued a report in 2003 urging health care professionals to commit to improving medication adherence, as it would generate "far greater impact on the health of population than any improvement in specific medical treatments". [1] Included in this report were the multifactorial causes of non-adherence: socioeconomic, health care team and health system, disease-related, therapy-related, and patient-related factors. [1,13] At face-value, these factors seem to be consistent with the predisposing characteristics (socio-demographic, social structure, and health beliefs), enabling resources (personal, family, and community), and need aspects that are included in the original ABM conceptualization. Table 2 compares the two theoretical frameworks by demonstrating how compatible the operationalization of each dimension can be.

Table 2 - Operationalization of the dimensions of two conceptual frameworks: the Andersen's Behavioral Model of Health Services Use and the Causes of Non-Adherence summarized by the World Health Organization

		WHO Report: Causes of Non-Adherence				
		Socioeconomic	Health care team / Health care system	Disease-related factors	Therapy-related factors	Patient-related factors
Andersen's Behavioral Model of Health Services Use	Predisposing characteristics	Education, race, ethnicity, income, occupation, marital status	Trust in medical organizations/health care team	Health-beliefs		Transportation, distance to health services, substance abuse
	Enabling factors	Urbanicity, Medicaid eligibility	Access to health care services, wait times, difficulty filling prescriptions, cost, health information, integration of health care team, physician-patient communication, facetime with health care providers			Health insurance, social/family support, health literacy
	Need characteristics			Evaluated health-status, comorbidities (MI, stroke, cancer), severity, symptoms	Treatment complexity, route of administration, side effects, duration, degree of behavioral change required	Activities of daily living, limitations in activities/profession, risk-factors (obesity, smoking, alcohol use)

Despite the seemingly broad agreement to using this ABM as theoretical framework for studying health services use, the operationalization of specific items for each of dimensions, predisposing characteristics, enabling, need resources has been inconsistent. [69] Thus, for the purpose of this research work and to make sure our implementation of the ABM theory is consistent with previous studies, we followed the operationalization of variables for each of the dimensions of the ABM theoretical framework according to the systematic literature review by Babitsch and colleagues. [69]

2.5 USE OF GBTM IN THE STUDY OF MEDICATION ADHERENCE: A SYSTEMATIC LITERATURE REVIEW*

Group-based trajectory modeling (GBTM) is method commonly used in social sciences that elicits common trajectories, or progressions of how a group of people behaves over a period of time. Lately, it has been used to evaluate adherence, overcoming the limitations posed by measuring this behavior with static measures such as annual PDC, or MPR. Instead of providing a single adherence ratio, it describes different medication adherence patterns as trajectories over time. With GBTM, patient administrative claims data are used to calculate monthly measures of adherence, which are then analyzed over time, clustering patients into similar patterns of longitudinal adherence behavior (e.g., always adherent or decreasingly adherent). [63] The number of trajectory groups may be identified based on several criteria, including Bayesian information criteria (BIC), where a lower value indicates better fit, [27] and Nagin's criteria for model adequacy, [56] with consideration given to group size such that the proportion of patients in each group is not less than 5% of total sample. [27,57] The estimation of the proportion of patients assigned to each trajectory is based on each individual's highest probability to belong to a certain group. Once the final number of groups is selected, trajectories are plotted and presented in graphical form. Regression models may then be used to estimate and characterize the predictors of each trajectory. [57,63]

Given the growing literature on GBTM medication adherence studies, there is a need to examine the commonalities and differences in models and the trajectories of adherence identified by them. The premise of using GBTM methods to investigate patterns of medication adherence

* Published in Alhazami M, Pontinha VM, Patterson JA, Holdford DA. Medication Adherence Trajectories: A Systematic Literature Review. J Manag Care Spec Pharm. 2020 Sep;26(9):1138-1152. doi: 10.18553/jmcp.2020.26.9.1138. PMID: 32857646.

is based on the limitations of measures currently used. Medication adherence is a complex phenomenon that warrants an individualized approach. If pharmacists, especially those within managed care organizations, can identify different adherence trajectories and underlying causes, they will be able to develop targeted individualized interventions. The identification of patients most likely to benefit from tailored interventions is not as straightforward when using annual or quarterly static measures like PDC or MPR. Consequently, several questions arise with studying medication adherence patterns using GBTMs. For example, what are the general characteristics of the studies in terms of the selected treatments, populations examined, and study features? What is the overall shape of medication adherence trajectories found in the literature? Are there consistently similar trajectories across studies? Finally, are there any specific trajectory patterns associated with individual patient populations or types of treatment? This section describes a systematic literature review of studies that utilized group-based trajectory models to identify patterns of medication adherence. The overarching goal of this review is to provide a summary of the patterns of medication adherence trajectories across studies and characteristics that were found to be associated with trajectory membership.

2.5.1 SEARCH STRATEGY

The literature review was conducted on April 2020 in PubMed and CINAHL databases using both MeSH terms and key words in appropriate combinations and in accordance with PRISMA guidelines. Search terms were combinations of: (group based trajectory modeling OR group based trajectory models OR group based trajectory OR trajectory) AND (medication adherence OR "Medication Adherence"[Mesh]), (group based trajectory modeling OR group based trajectory models OR group based trajectory OR trajectory) AND (fill prescription), (group based trajectory modeling OR group based trajectory models OR group based trajectory

OR trajectory) AND (medication compliance), (group based trajectory modeling OR group based trajectory models OR group based trajectory OR trajectory) AND (medication persistence). To ensure complete capture of relevant articles, the authors conducted backward snowballing, by reviewing the references of the included articles.

2.5.2 REVIEW PROCEDURES AND STUDY SELECTION

Titles and abstracts from search result articles were screened by two investigators based on inclusion and exclusion criteria. Disagreements pertaining to the inclusion and exclusion criteria were resolved by discussing with a third researcher. Studies that evaluated patient adherence to oral medications over time using group-based trajectory models and were published in English were included. Studies focusing on select special populations (i.e., children and pregnant women) were excluded due to the unique nature of medication taking behaviors in these groups. Papers examining adherence to provider-administered biologics were excluded because adherence behavior may systematically differ from self-administered medications obtained from a pharmacy. Review articles, case-reports, and prospective studies without results were also excluded.

2.5.3 DATA EXTRACTION

Two researchers extracted data concerning authorship, year of publication, country, population, type of treatment, sample size, number and shape of trajectory groups identified, clinical outcomes that were being assessed, and general conclusions that results from GBTM analysis.

2.5.4 SUMMARY OF RESULTS

The search on PubMed and CINAHL yielded a total of 248 articles. After screening for duplicates and applying inclusion and exclusion criteria, 20 research articles remained for analysis (Figure 6). After the full-text review, included articles' references lists were searched, yielding 8 additional studies for inclusion.

All 28 papers were published in the last ten years (2010-2020). Most studies were conducted in the US (n=17, 60.7%),[59,60,62,63,70-83] while the remaining were conducted in Europe (n=7, 25%),[61,84-89] Taiwan (n=1, 3.6%),[90] and Australia (n=2, 7.14%).[91,92] Although a variety of medical conditions were explored in GBTM populations, the majority of studies focused on populations diagnosed with cardiovascular diseases (n=19, 67.9%).[60-63,70,72,74,75,78,79,81-86,89,91,93] Study duration ranged from nine months[61] to six years, [90] with most lasting two years or less (n=22, 78.6%). Table 3 summarizes the studies included in this review by the population characteristics, type of treatment, study duration, adherence measure, country, sample size, number of trajectory groups identified, and overview of the conclusions.

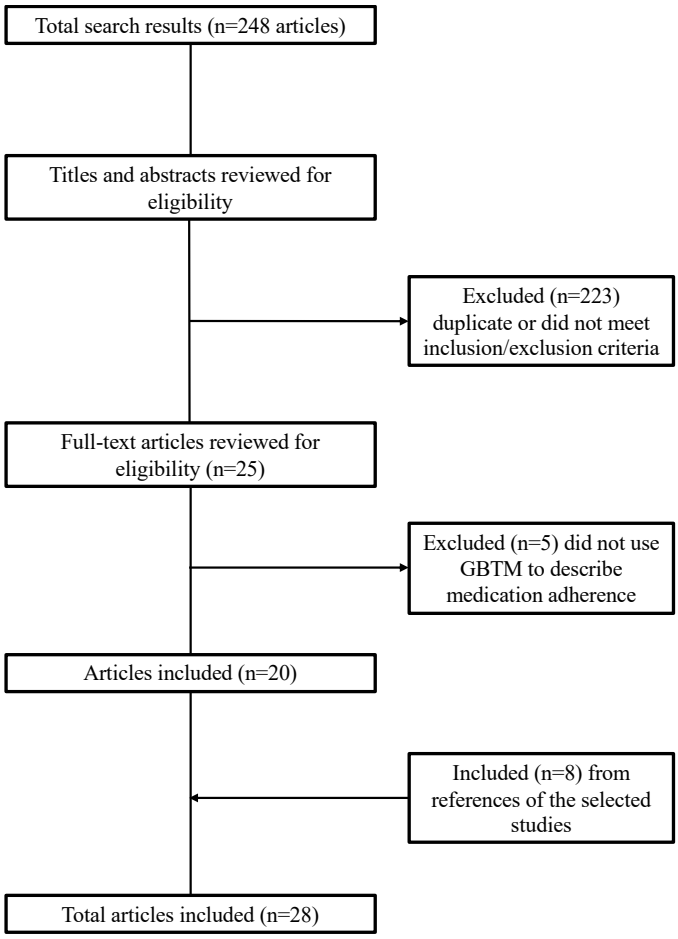

Figure 6 - PRISMA Diagram

Table 3 - Summary of the included studies

Authors	Objective	Population	Type of treatment	Duration of study	Country	Sample size	Adherence measure	Trajectory group	Clinical outcomes	Conclusion
1.Dillon et al., 2019[30]	1.Classify adherence to antihypertensives using PDC and GBTM 2. Evaluate the longitudinal association between community pharmacy antihypertensive refill-adherence metrics (GBTM and PDC) and the number of hospital visits and GP visits	Patients aged ≥ 65 years old	Multiple classes of antihypertensives	1 year	Republic of Ireland	905	PDC	1. Perfect adherence 2. High adherence 3. Low adherence	Hospital visit, general practitioners (GP) visit	PDC could be used to rank patients by adherence level and be combined with other adherence measurements, such as individual adherence trajectory graphs, to provide a richer picture of patient adherence behavior
2.Marcum et al., 2019[16]	1.Examine antihypertensive and statin adherence 2.Compare people who went on to develop dementia to those who did not	Patients aged ≥ 65 years old	Multiple classes of antihypertensives and statins	3 years	US	4,368	PDC	1. Perfect adherence 2. Moderate non-adherence 3. *Slow decline* 4. *Rapid decline*	-	Patterns of medication adherence may be useful to identify a subset of people at higher likelihood of developing dementia
3.Winn et al., 2019[17]	1. Examine how adherence to ET for women with breast cancer was impacted by reducing copayments for ETs by the introduction of generic ETs among women who do not receive a subsidy compared with those that do receive a subsidy and are not exposed to any changes in copayments by using GBTM	Women with breast cancer age>66 years' old	Endocrine therapy (ET)	1 year	US	3,344	PDC	1. Perfect adherence 2. *Slow decline* then increase 3. *Rapid decline*, then increase 4. *Slow decline* 5. *Rapid decline* 6. Low adherence	-	This study describes a new approach to identify het-erogenous effects when using an interrupted time series research design

4.MacEwan et al., 201818	1. Identify patients' antipsychotic adherence patterns 2. Assess whether antipsychotic adherence patterns can predict other medications adherence patterns	Patients with serious mental illness who initiated an atypical antipsychotic and were also taking an SSRI, biguanide, or an ACE inhibitor	Atypical antipsychotic, SSRI, biguanides, and ACE inhibitors	1 year	US	431,591	PDC	1. Perfect adherence 2. *Rapid decline*, then increase 3. *Slow decline* 4. *Rapid decline*	Adherence patterns for SSRIs, biguanides, and ACE inhibitors	Among patients with multiple chronic mental and physical illnesses, patterns of atypical antipsychotic adherence were useful predictors of adherence patterns to a patient's adherence to ACE inhibitors, biguanides, and SSRIs
5.Feldman et al., 201836	1. Identify adherence trajectories for patients taking Hydroxychloroquine (HCQ) 2. Identify predictors of non- adherence to HCQ	Adult Medicaid patients with systemic lupus erythematosus (SLE)	Hydroxychloroquine (HCQ)	1 year	US	10,406	PDC	1. Perfect adherence 2. Moderate non-adherence 3. Low decline 4. Very *rapid decline*	-	HCQ adherence is a dynamic behavior that declines over the first year of use
6.Franklin et al., 201820	1. Identify adherence trajectories for statins, antihypertensives, and oral antidiabetics 2. Identify predictors of non-adherence trajectories 3.Evaluate the accuracy of predictions of medication adherence based on EHR data versus claims	Tufts Health Plan Medicare Advantage beneficiaries aged 65 and older, receiving care at Harvard Vanguard Medical Associates (HVMA)	Statins, multiple classes of antihypertensives, and oral antidiabetics	1 year	US	11,479	PDC	1. Perfect adherence 2. Decline after 9 months 3. Moderate non-adherence 4. Decline after 6 months 5. *Rapid decline*	-	EHR data can provide good predictions of adherence trajectory

7.Schaffer et al., 201735	1. Compare statin adherence in individuals initiating combined amlodipine/atorvastatin therapy as a fixed-dose (FDC) or free combination 2.Identified subgroups benefiting most from FDCs 3. Identified predictors of adherence trajectories	All patients initiating amlodipine in combination with atorvastatin, either as an FDC or in free combination	amlodipine and atorvastatin as an FDC and free combination	2 years	Australia	3996 (FDC), 5434 (free combination) 905	PDC	1. Perfect adherent 2. Moderate adherence 3. *Slow decline* 4. Very *rapid decline*	-	The amlodipine/ atorvastatin FDC was associated with greater statin adherence among prevalent statin users
8. Dillon et al., 201829	1.Characterize adherence to antihypertensive medication 2. Test predictive validity of GBTM in BP measurement	Patients aged ≥ 65 years old	Multiple classes of antihypertensives and statins	1 year	Republic of Ireland	905	PDC	1.Perfect adherence 2.High adherence 3.Moderate non-adherence	-	GBTM identified 3 trajectories. However, did not show predictive validity with BP measurement
9. Hargrove et al., 201721	1.Use GBTM to identify antihypertensive adherence trajectories 2.Comapre adherence trajectories to traditional adherence measures 3. Identify patients characteristics associated with adherence trajectories	Medicare patients aged ≥ 65 years old	Multiple classes of antihypertensives and statins	1 year	US	282,520	PDC Used to compare identifying adherent and non-adherent months	1. Perfect adherence 2. *Rapid decline*, then increase 3. *Moderate decline*, then increase 4. *Moderate decline* 5. *Rapid decline* 6. Very *rapid decline*	-	GBTM is an effective method to identify patterns of medication adherence compared to PDC

10.MacEwan et al., 201622	1. Identify adherence trajectory for oral atypical antipsychotics (OAA) 2. Identify associated factors	Adult schizophrenia patients	Oral atypical antipsychotic	1 year	US	29,607	PDC	1. Perfect adherence 2. Decline after 3 months 3. Decline after 6 months 4. Decline after 9 months 5. *Rapid decline*, then increase after 6 months 6. Very *rapid decline*	Psychiatric inpatient admission and ED visit	Adherence patterns identified by GBTM are more varied than research based on PDC -Lower adherence trajectories associated with higher ED visits
11.Aarnio et al.,201631	1. Identify adherence trajectories of statin 2. Examine association between SEP and adherence trajectories	Patients aged 45 to 75 years	Statin	18 months	Finland	116,846	PDC Used to compare predictors ability to identify non-adherent participants.	1. Perfect adherence 2. High adherence 3. *Rapid decline*, then increase 4. *Moderate decline* 5. *Rapid decline* 6. Very *rapid decline*	-	SEP is associated with low adherence groups. Overall, GBTM provide insight to dynamics of adherence behavior
12.Librero et al., 201632	1. Identify adherence trajectories for ACEI, statin BB, and antiplatelet 2. Identify associated factors	Patients discharged with coronary heart disease (CHD)	ACEI/ARB, beta-blockers, statins, and antiplatelet	9 months	Spain	7,462	PDC	1. Perfect adherence 2. Low adherence, then increase 3. *Moderate decline*, then increase 4. *Moderate decline* 5. *Rapid decline*	-	GBTM identified distinct adherence trajectories for difference preventive medication for CHD. It showed advantage over traditional measure

13.Mårdby et al., 201633	1. Identify adherence trajectories of antidepressants 2. Identify associated factors	Patients aged 18–85 years	Citalopram	24 months	Sweden	54,248	MPR	1. Perfect adherence 2. *Moderate decline* 3. *Rapid decline*, then increase 4. *Rapid decline* 5. Very *rapid decline*	-	-GBTM identified 5 distinct patterns -Low adherence trajectories associated with lower SEP
14. Lo-Ciganic et al., 201623	1. Identify adherence trajectories of oral hypoglycemic and associated factors 2. Examine association with clinical events	Patients aged 18–64 years with diabetes (DM)	Oral hypoglycemic medication	1 year	US	16,256	PDC Used to compare GBTM in the ability to predict outcome	1. Perfect adherence 2. High adherence 3. *Moderate decline* 4. Moderate non-adherence 5. *Rapid decline*, then increase 6. *Rapid decline* 7. Very *rapid decline*	1. DM related ED visit/hospitalization 2. All cause ED visit/hospitalization	Lower adherence trajectories associated with higher ED/hospitalization events
15. Winn et al., 201624	1. Identify adherence trajectories of ET 2. Identify associated factors 3. Examine association with mortality	Women with breast cancer	Endocrine therapy (ET)	1 year	US	9,492	PDC Used to compare GBTM in the ability to predict outcome	1. Perfect adherence 2. *Slow decline* 3. *Rapid decline*, then increase 4. *Moderate decline* 5. *Rapid decline*	All-cause mortality	Low adherence groups associated with higher mortality rate

16. Chen et al., 201634	1. Identify adherence trajectories of hypoglycemic 2. Identify associated factors	Patients aged ≥ 18 years type 2 diabetes	Oral hypoglycemic medication	6 years	Taiwan	12,123	MPR	1. Perfect adherence 2. Moderate adherence, then increase 3. *Moderate decline* 4. Low adherence, then increase	-	GBTM help in identifying heterogeneity of medication adherence
17. Franklin et al., 201525	Identify adherence trajectories of statin	Patients aged ≥ 65 years old	Statins	1 year	US	77,703	PDC	1. Perfect adherence 2. Moderate adherence 3. *Rapid decline*, then increase 4. *Moderate decline*, then increase 5. *Rapid decline* 6. Very *rapid decline*	-	Initial adherence behavior associated with better future adherence
18.Newman-casey et al., 201526	1. Identify adherence trajectories of glaucoma medication 2. Identify associated factors	Patients ≥40 years old treated for glaucoma	Glaucoma medication	4 years	US	1,234	MPR	1. Perfect adherence 2. Moderate non-adherence 3. *Moderate decline* 4. Low adherence, then increase 5. *Rapid decline*	-	Adherence patterns for first year, had great impact on future adherence behavior

19.Juarez et al., 201527	1. Identify adherence trajectories of ACEI 2. Identify associated factors	Patients with Congestive Heart failure (CHF)	ACEI	6 years	US	10,986	MPR	1. Perfect adherence 2. *Moderate decline* 3. Low adherence, then increase 4. *Rapid decline*	-	Patients factors associated with low adherence trajectories can be used to target interventions
20. Franklin et al., 201528	1. Identify adherence trajectories of statin 2. Examine association with clinical events	Patients in the UnitedHealth Optum Research Datamart, aged 35-64 years	Statins	1 year	US	519,842	PDC Used cross-validation methods to ascertain if initial PDC could predict trajectory membership	1. Perfect adherence 2. *Rapid decline* then increase 3. *Moderate decline* 4. *Moderate decline*, then increase 5. *Rapid decline* 6. Very *rapid decline*	Hospitalization for an acute coronary event, revascularization, cerebrovascular event, or heart failure	Adherence trajectories predicts future clinical outcomes better than PDC
21.Franklin et al., 201310	1. Identify adherence trajectories of statin 2. Identify associated factors	Patients initiating a statin in CVS Caremark	Statins	15 months	US	264,789	PDC	1. Perfect adherence 2. *Rapid decline* then increase 3. *Moderate decline* 4. Moderate non-adherence 5. *Rapid decline* 6. Very *rapid decline*	-	GBTM summarized adherence patterns better than traditional measures

22. Vadhariya et al., 2019[82]	1. Identify distinct trajectories of adherence to statin therapy 2. Identify sociodemographic and clinical predictors of trajectory membership	Patients enrolled in a Medicare Advantage Plan	Statins	12 months	US	7,850	PDC	1. High/Nearly perfect adherence 2. Declining adherence, then increase 3. Gradual decline 4. *Rapid decline*	-	Adherence trajectories are consistent with previous findings in the literature. Patient characteristics are found to be predictors of trajectory membership
23. Lambert-Côté., 2020[88]	1. Identify adherence trajectories and characterize trajectory shapes 2. Identify factors associated with group membership	Women with breast cancer living in metropolitan areas covered by National Health Insurance plan	Adjuvant endocrine therapy	5-years	France	674	PDC	1. Very high adherence 2. High adherence 3. *Slow decline* 4. *Moderate decline* 5. Quick decline	-	Prior exposure to chemotherapy and personalized care plan were found to be predictors of trajectory membership Dynamic behavior of adherence not fully captured byb traditional annualization of PDC
24. Ajrouche et al., 2020[89]	1. Identify adherence trajectories of low-dose aspirin (LDA) 2. describe trajectories based on primary or secondary indication for LDA 3. identify predictors of LDA adherence trajectory membership	Patients with at least 3 months of follow-up after first LDA delivery	Low-dose aspirin	3-years	France	11,793	PDC	1. Perfect adherence 2. Declining adherence 3. Declining adherence then increase 4. Low adherence	Hemorrhagic events	GBTM provides a better understanding of longitudinal trends by yielding visual developmental trajectories Non-adherence trajectories seem to be associated with uncertain efficacy of using LDA in primary prevention

2.5.4.1 Number and types of trajectory groups

In Table 3, the column "Trajectory group" encompasses the number and general description of the adherence trajectories found in each of the included studies. The number of trajectory groups identified ranged from three to seven, although most papers (n=18, 85.7%) identified four to six trajectory groups. Regardless of the total number of trajectories identified, 19 studies identified the following four similar adherence trajectory groups: (1) consistent high adherence, (2) declining adherence, (3) early and consistent non-adherence, and (4) initial non-adherence followed by a slight increase in adherence.

Consistent high adherence. All of the published papers identified a "perfect adherence" or a "nearly perfect adherence" trajectory composed of patients who consistently registered a monthly probability of PDC or MPR over 80% for the duration of the study. In most studies, (n=22, 78.6%), the consistently high adherence trajectory group comprised the largest proportion of patients, ranging from 21% to 80% of patients in the study. More specifically, fourteen studies observed a consistently adherent group which represented more than 40% of the sample. [59-61,70-72,74-76,78,81,82,84-86,90-93]

Declining adherence. Studies that identified more than one declining adherence trajectory varied primarily in the shape and speed of decline. The most identified declining adherence trajectory (n = 26, 92.9%) was one characterized by a gradual decline in adherence, with "falling" adherence rates starting anywhere between the third and ninth month. [60-63,71,72,74-77,79,86,87,90,91,94] Studies also often (n = 22, 78.6%) identified a rapid trajectory, in which adherence immediately declines within the first two months of therapy initiation. [60,62,63,71,72,74-77,79,82,86-88,91-94]

Early and consistent non-adherence. All articles reported early non-adherent behaviors. In these trajectory groups, adherence steeply declined well below 80% after the first months. In some papers (n=4), this group was designated as *rapid decline* in adherence, or persistent non-adherers. [60,72,91,94] In all studies, the steep decline in adherence was usually followed by a consistently very low adherence throughout the study duration. However, adherence analyses based on administrative claims data require at least one fill of the prescribed medicine. Notably, adherence measures using administrative claims data do not capture patients with primary nonadherence who never pick up their initial fill. GBTM models using claims data can only identify patients with secondary nonadherence and may therefore underestimate the true proportion of consistently nonadherent patients, which would more accurately include both patients with primary and secondary nonadherence.

Initial non-adherence followed by an increase. The fourth commonly identified trajectory group described patients whose adherence was poor at the beginning of the study period but would gradually improve over time (n=18, 64.3%).[59-63,71,72,75-79,82,86,87,89,90,93] This rebound in adherence was more frequent following a period of rapid declining adherence (n=12, 42.9%), even though four studies identified trajectories with improving medication adherence following both rapid and gradual declines. [60,62,71,79,83] This trajectory was different from the others which appeared to display an initial pattern of non-adherence, which evolved to a pattern of increasing adherence. Despite not being referenced in the included studies, since most use data from administrative claims, this transition may be indicative of some event or intervention that explains the change in the adherence patterns of patients. The studies with a follow-up period between 9 and 18 months, there was a clear pattern in the time point at which an increase would

be observable in the declining adherence trajectory (4-6 months). [59-63,72,75,76,79,82,83,86,93] The studies with longer follow-up periods (3-6 years) described increases at different time points, primary between the first and second year. [78,89,90,92] The type of drug being analyzed also varied, ranging from statins, antipsychotics, antidepressants, ACEIs, which does not indicate a common underlying reason for the increase in adherence after a period of declining adherence. [77,89,90,92] In one study with a follow-up period of 5 years, a declining adherence trajectory increased to a stable level of adherence, suggesting that a trend of reaching a plateau in adherence may be observed if follow-up periods were extended beyond 18 months. [92] It is possible, that the trend of reaching a *plateau* in adherence would be observed in the shorter follow-up period studies.

2.5.4.2 Drug classes studied

Except for one study, [61] all studies focused on adherence to a single drug class over time. Single drug class studies examined trajectories of a wide selection of medications: antihypertensives, statins, oral antihyperglycemic drugs, antipsychotics, antidepressants, treatment for systematic lupus erythematosus, glaucoma, and endocrine therapy (Table 3).

Four of the six studies that focused exclusively on statin adherence identified six adherence trajectories, while the other two found only 4 trajectories. Studies that examined both statins and antihypertensives, however, varied in the number of identified trajectories (3-6 trajectories), as did studies that included oral antidiabetic agents (4-7 trajectories).

2.5.4.3 Patient Characteristics Associated with Trajectories

Patient socio-demographic and clinical characteristics were commonly associated with the likelihood of belonging to specific adherence trajectories in the included studies. Juarez et al.,

found that race, such as being Black, Asian, and Pacific Islanders was associated with belonging to non-adherence trajectories when compared to white patients. [78] The same types of associations were found in other studies in which individuals who were younger, non-white, males, with more severe disease and comorbidities, and having lower educational attainment and socioeconomic status being more likely to follow low adherence trajectories. [59,63,76-78,82,87,90,94] Conflicting results were found in two studies, in which women were more likely to belong to non-adherence trajectories. [82,89] Low adherence trajectories were found for patients from a lower socioeconomic status[1, 30, 32] and for individuals with more severe diseases and comorbidities. [75,94] Treatment complexity, race, prior history of the disease, frailty, concomitant use of opioids, and comorbidities were also associated with lower or non-adherence trajectories. [59-61,74-77,84,86,87,94] Adherence trajectories were influenced by patients' experiences taking other medications. MacEwan and colleagues found that adherence to atypical antipsychotic drugs can be a predictor of adherence to other medications. [72] Prior medication use was also found to be a predictor of higher adherence in other studies. [60,84,91] Patients in groups defined by declining adherence followed by an increase seem to have distinct features. Although most studies found typical sociodemographic characteristics in patients in non-adherence trajectories, drug abuse, alcoholism[75], hypertension[60], a higher number of comorbid diagnoses[61], and higher levels of out of pocket expenditures[61] were each reported to be overrepresented in declining adherence or non-adherence trajectories. Contrastingly, other characteristics seem to make patients less likely to belong to declining adherence trajectories or declining adherence followed by an increase. These included higher adherence to prior medication[72], and being a women outside of the labor market[86].Finally, the only study with a pre- and post- approach found that patients who were previously in a declining followed by an increase in adherence trajectory were more likely to

belong to nearly perfect adherence trajectories after a critical medical event (i.e. acute coronary syndrome). [89]

2.5.4.4 Trajectories and Health care Resource Utilization (HCRU) and Health Outcomes

Five articles examined and found a relationship between adherence trajectories and health care events (e.g., hospitalization events, emergency department (ED) visit, adverse events, or death), [59,75,76,79,85] consistently reporting that consistently adherent were associated with lower HCRU and improved health outcomes. Specifically, fewer hospitalizations were seen among patients in consistently adherent trajectories for antihypertensive, [85] statins, [79] and oral antidiabetic drugs, [59] with similar trends for reductions in ED visits in studies on antipsychotics[75] and oral antidiabetics. [59] Higher utilization of health care resources was not always associated with belonging to poor adherence trajectories. Even though not specific to increased health care resource utilization, a more individualized approach to treatment plan seems to improve the outlook of medication adherence. In Lambert-Côté and colleagues, women with breast cancer who had received a personalized treatment plan, that comprised receiving a written document with information and resources available throughout the treatment duration provided at diagnosis were more likely to belong to consistently adherent trajectories. [88] Conversely, patients in consistent non-adherent or declining adherence trajectories had a significantly increased risk of adverse outcomes. Patients on endocrine therapy in the rapidly and slowly declining adherence trajectories had a significantly higher risk of death compared to those in perfectly adherent trajectories. [71] Consistent non-adherence trajectory group membership was associated with a higher risk of cardiovascular events (hospitalization for acute coronary event, resvascularization, cerebrovascular event, or heart failure) in patients taking statins. [79]

2.5.4.5 Predictive Validity of GBTM

Five studies compared GBTM to conventional dichotomous measures (e.g., PDC), and all five concluded that GBTM is advantageous, namely because of the ability to encapsulate the trends of medication adherence over a period of time. [59,60,76,79,86] Of these, only three studies used multinomial logistic regression and computed the c-statistic to compare predictive validity of using GBTM *versus* dichotomized periodic PDC. [59,60,76] However, the comparison focused on different aspects: in Hargrove et al., reported that a six group trajectory model better distinguished between adherent and non-adherent months than either the PDC or even Proportion of Months Covered (i.e., another possession-based adherence measure). [60] Lo Ciganic et al. reported the c-statistic to compare dichotomized PDC and GBTM model in predicting diabetes-related hospitalization and ED visits than PDC measures. [59] In Franklin et al., the c-statistic was used to compare the GBTM model and dichotomized PDC in predicting cardiovascular events. In all situations, the c-statistic of the GBTM fared better than the traditional dichotomized PDC ≥ 80% threshold approach. While the other two studies did not report specific statistics, Franklin and colleagues refer the more refined approach that GBTM provides in identifying longitudinal trends. [79] Aarnio et al., found GBTM to be comparable to PDC to identify relationships between non-adherence and lower socioeconomic status, but that GBTM provided a more differentiating approach to analyzing non-adherent groups, [86] and Winn et al. also determined that GBTM better differentiated adherence groups. [76]

2.5.5 MAIN FINDINGS FROM THE SYSTEMATIC LITERATURE REVIEW

Our review found that four to six trajectory groups could be used in most cases to describe medication taking behavior over time. The most frequent trajectories can be described using the

following labels: (1) consistently adherent, (2) declining adherence, (3) consistent non-adherence, and (4) initial non-adherence followed by an increase. Studies that identified fewer or more trajectories than four to six may have been affected by high adherence of patients studied, bias in selecting study subjects, the nature of the drugs studied, unique characteristics of patient populations, and differences in the rate of decline from adherence to non-adherence. [59,85] Figure 7 depicts the general trend identified in this literature review. Generally, adherence trajectories can be summarized by four trajectory groups, with declining adherence trajectory possibly occurring at different rates (fast and slow dashed lines).

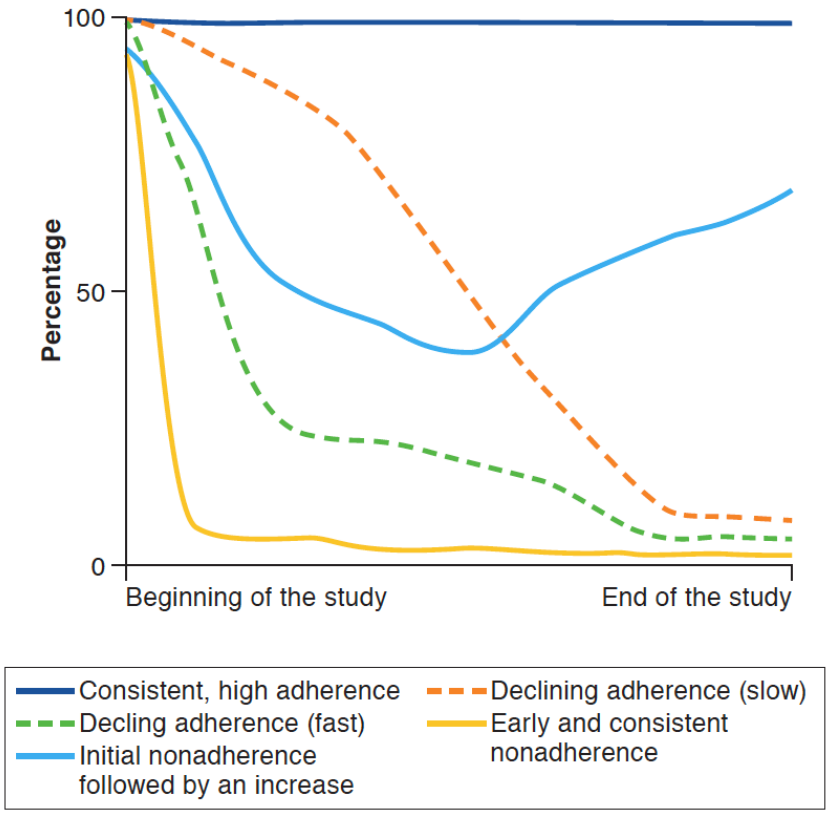

Figure 7 - Illustration of Similar Medication Adherence Trajectories[95]

The value provided by GBTM over current measures such as PDC and MPR is that it describes behavior over time instead of at discrete time periods. Understanding the temporal

changes in behavior is important in evaluating and predicting influences on patient adherence. Hypothetically, a patient who interrupts their medication taking due to complications might have a different trajectory than one who stops due to an issue of affordability or other issue, assuming affordability and complications happen in different time points. Rather than classifying patients into dichotomous adherent and non-adherent groups, GBTM provides the longitudinal perspective to understand and describe patterns of non-adherence. The distillation of adherent trajectories has the power to differentiate between underlying causes of non-adherence in a way that periodic dichotomization of PDC cannot. Early non-adherence can be a result of adverse events that otherwise would not be identified if the patient did not follow-up with a health care professional. [60,72,91,94] Similarly, declining adherence followed by an increase, may be related to specific medical events that resulted in increased motivation to resume the treatment plan, suggesting that targeted interventions following medical events may be effective in improving adherence in these patients. [89] Those patterns should be explored through further research into the causes and potential solutions to non-adherence trajectory groups.

When contrasted with dichotomous PDC and MPR measures, GBTM offers a richer description of patient behavior. In Märdby and colleagues the simple dichotomized PDC analysis showed that 50% of the patient population was non-adherent. However, when a GBTM analysis was conducted, they found that a significant portion of the patients initially classified as non-adherent belonged to group with declining adherence that later increased. The finding of a subgroup with GBTM suggests heterogeneity among non-adherents that require a more sophisticated array of interventions to improve medication adherence.

Other implications result from this review. Payers like the CMS and other private health insurers that provide discounts and rebates for providers who meet certain quality measures, such

as medication adherence. Specifically, Medicare Part D plans incorporate medication adherence as a significant component of their quality-based payment program (Medicare Star Ratings). [10,96] In one of the studies, two trajectories (declining followed by an increase and *slow decline*) were found to have no difference in risk of mortality when compared to the consistent adherent trajectory. [76] Therefore, it is likely that further research may elicit other *optimal* thresholds of adherence. Thus, systems that either reward or penalize providers based on medication adherence measures should utilize methods that appropriately demonstrate variations in medication adherence. Moreover, GBTM has already served to develop a tailored motivational interviewing system based on adherence trajectories. The individualization of interventions not only allows for a better use of available resources but is able to produce significant improvements in adherence. [97] Likewise, using GBTM may also allow to identify patients who are keen on following a specific adherence trajectory, and consequently develop an action plan to invert the trend. Future research should clarify significant thresholds of medication adherence improvements, that result in improvement of patient outcomes, as opposed to computing an average medication adherence percentage in a given period.

Finally, we also found several inconsistencies in the model selection and identification of predictors of trajectory memberships. These include different implementations medication adherence, in which some studies considered the outcome as the probability of being adherent or non-adherent, assuming the 80% PDC threshold, while other studies considered the outcome to be the nominal monthly PDC measurement. Additionally, some of the factors that are known to influence medication adherence are not likely to remain constant over time. Therefore, a method that analyzes how time-varying predictors influence medication adherence over time might be helpful. GBTM also allows the visualization of each trajectory of a desired outcome in

combination with the time-dependent covariate, and it have never been used before in the context of describing medication adherence and its predictors in trajectories. Traditionally, the identification of covariates that suggest trajectory membership have used multinomial logistic regressions. However, like medication adherence, the predictors of medication adherence trajectories can themselves change with time. Thus, using GBTM would allow investigating the extent to which the variation of a covariate over time influences the behavior of interest. This approach has been applied successfully in clinical research to investigate biomarkers as predictors of End-Stage Renal Disease (ESRD) in patients with chronic kidney disease (CKD). [98] In this study, the outcome of interest, glomerular filtration rate (GFR) was time dependent, as well as all the other predictors (anemia, metabolic acidosis, secondary hyperparathyroidism, and hyperphosphatemia). The resulting model was able to identify which time-dependent predictors are associated with worsening of the renal function, in addition to predicting which trajectories are associated with end-stage renal disease. The model showed a powerful predictive capability, despite the acute nature of ESRD: misclassification rate of approximately 35% and 20%, in consecutive time points. [98]

2.6 MACHINE LEARNING ALGORITHMS IN HEALTH SERVICES RESEARCH

Conventional statistical methods like logistic regressions have been used for purposes of classification and prediction. However, these methods sometimes are marred by methodological limitations. For example, logistic regressions fit the data in models to calculate the log odds of the dependent variables to a linear combination of the explanatory variables. From social to medical sciences, researchers have been using logistic regressions to identify risk factors in diseases (sick or healthy), behavioral trajectories because of its clarity and succinctness. [99-103]

Despite its simplicity, logistic regressions assume a linear correlation between the log odds of the dependent variables and the covariates in the model. [104] With the increase of computational power, researchers have now the ability to implement more complex methods like machine learning algorithms for making sense of the data: classification and prediction, while overcoming the limitations of traditional statistical methods, especially when those limitations are detrimental to the predictive performance. [105,106] Random forest (RF) algorithms are a part of the family of the machine algorithms and these methods assume a non-linear correlation between the features and outcomes – they are non-parametric methods. [107,108] RF algorithms are a predictive modeling tool that has gained traction as a classification technique using high-dimensional data. For example, in one study investigating almost 20,000 datasets, investigators concluded that random forest algorithm consistently performed better than multinomial logistic regression for predictive ability and identification of relevant outcome predictors. [109]

In its original form, as proposed by Breiman, RF refers to the ensemble of decision trees, each using a particular combination of features (predictors). [110] Unlike logistic regression, the random forest algorithm focuses on the predictive component rather than trying to fit the data in a single model. [111] It does so by computing a large number of possible decision trees based on a bootstrap sample randomly selected. [110,112] The forest in RF refers to the multitude of trees that represent all possible combinations of features. A critical characteristic of RF is the randomness with which features are selected to build the tree. Each decision tree is built using random covariates. In the end of each decision tree, the algorithm evaluates whether that combination of covariates is a good predictor of the outcome or not (Figure 8). The evaluation refers to a count of the number of *votes* for each outcome being predicted.

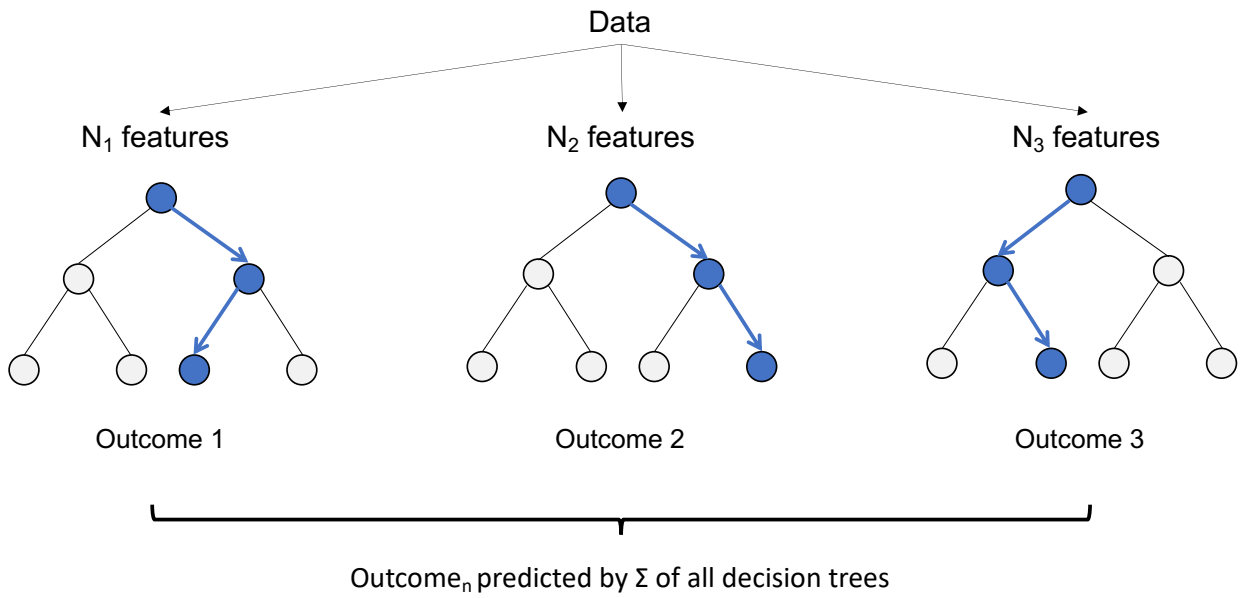

Figure 8 - Illustration of random forest decision trees to elicit predictors or certain outcomes

In a pure RF model, final predictions are based on the greatest number of correct votes. [110] Each tree vote counts the same in the final prediction. A significant limitation can be identified from this approach. While the voting method follows the original algorithm, it may seem illogical that a decision tree, who has failed to correctly characterize the sample, has a vote with equal weight as compared to another tree who characterized the sample significantly better. Because there are trees that classify the sample better than others, researchers can use boosting methods, which allow to weigh the "quality" of each tree in the final voting tally. Boosting methods, such as AdaBoost (adaptative boosting) or gradient boosting algorithms are still based on random forests and appear to optimize the random forest approach to yield better predictions.

Of these, the most popular boosting techniques are AdaBoost and Gradient Boosting learning algorithms:

2.6.1 ADABOOST

AdaBoost only builds the beginning of a tree ("*stump*") with only two 2 leaves. [113] The reason why RF builds a complete tree is because, by themselves, tree stumps are weak learners

(i.e. predictors). However, AdaBoost iteratively weighs how well each stump can classify the sample. After running several iterations, the final classification is a combination of the votes of the forest of stumps, which are weighed by the "amount of say" (i.e., the total error) of each stump. In short, AdaBoost learns by considering the forest of good learners (higher "amount of say") and weak learners (less "amount of say"). The implementation of AdaBoost is relatively simple but is limited to classification problems, that is, when the outcome is binary (e.g.: sick/not sick). For this study, the main driver to consider machine learning algorithms is to investigate whether these methods better explain the association between medication adherence trajectories and health outcomes (usually operationalized in categorical variables, such as "exhibit complications/does not exhibit complications") and economic outcomes (e.g.: total medical expenditures or number of hospitalizations). This two-fold aim would not be possible using solely AdaBoost. Moreover, while AdaBoost has been shown to be immune to overfitting, AdaBoost seems to be wildly subject to the effects of noise. [112] This means, that if the labelling of the outcome is changed randomly, the AdaBoost model loses predictive ability, when compared to the models following the pure RF algorithm. Thus, the training dataset needs to be of high quality with as little noise as possible and with outcomes labelled correctly. The ability to be resistant to noise is helpful to identify, for example, misdiagnosis: if a model is resistant to noise and with high predictive ability (e.g.: c-stat > 0.95), they would be able to identify observations who have been wrongfully labelled.

2.6.2 EXTREME GRADIENT BOOSTING

Like with AdaBoost, gradient boosting uses a Random Forest approach and can be used to model both categorical and continuous outcomes (called gradient boosting for regression). This is particularly important because, as stated, this study investigates the link between medication

adherence trajectories and outcomes which can be operationalized as both binary and continuous variables.

Gradient boosting builds trees that are *deeper* than the tree "stump", but still considered weak learners as compared to the full-fledged trees built with the pure RF algorithm. Generally, researchers use a depth (number of leaves) ranging from 8-32 leaves. Furthermore, gradient boosting also scales the trees according to level of error, but it does so for all the trees in each iteration (the algorithm specifies a learning rate, as opposed to the "amount of say" automatically calculated in each iteration with the AdaBoost algorithm). Subsequent trees built by the gradient boosting algorithm are still based on the error of the previous trees, until the number of trees specified in the model are achieved or the computation of additional trees fails to improve the predictive ability of the model. In simple terms, gradient boosting works sequentially to maximize the learning ability of the features in the model, so that the residuals (approximately understood as the different between predicted and observed) is as close to zero. The inclusion of a learning rate is so that the model is not subject to high variance (i.e., it would take more trees to reach the level of residuals of zero). Thus, it is no surprise that gradient boosting is usually the boosting technique that yields the best predictive models. [114,115] Moreover, the available packages of gradient boosting for R software allow the inclusion of cross-validation techniques, so that researchers are able to identify the best number of trees for prediction. [116,117] Additionally, the number of parameters that are required to fine tune a model with gradient boosting are of simple comprehension and limited number:

- Number of trees (usually more than 100)
- Depth of the tree (8-32)
- Learning rate (ranging from 0.01 – 0.1)

- Distribution type (binary, multinomial, or continuous)
- *n* folds for cross validation

Regardless of RF approach, once predictors are found, the algorithms (original or boosted version), by means of the CART method of random draw and Decrease Gini Impurity splitting criterion, ranks the covariates by order of relevance for prediction. This ranking is usually called Variable Importance Measures (VIM). [111]

The assumption that medication adherence trajectories, which are defined by polynomial functions of time, can have a linear correlation with, potentially, time-dependent covariates seems questionable. Therefore, it is pertinent to investigate the predictors of trajectories of medication adherence using widely accepted methods (multinomial logistic regression) and compare it to more innovative techniques like random forests. Furthermore, should specific medication adherence trajectories be associated with certain health outcomes, the better predictive model will be able to identify patients at-risk during the follow-up period, given that for each trajectory, relevant risk factor, i.e., predictors will be identified.

2.7 SPECIFIC AIMS AND HYPOTHESIS

Approximately half of the chronic patients do not take their medicines as directed. Developing effective medication adherence interventions is critical to the extent to which improvements in medication adherence can potentially yield better health outcomes that any other medical innovation. Moreover, the reasons for non-adherence have been shown to be diverse and changeable over time. As a result, the current measures used to assess adherence may not show the fluidity of the adherence behavior as well as the factors that influence it.

Group-based trajectory models have been gradually more used to investigate the medication adherence of oral and non-oral dosage forms. We conducted a systematic literature review that showed that, in general, 4-6 trajectories of medication adherence are seen longitudinally irrespective of the disease state. Further, several factors were found to be predictive of adherence trajectory membership. However, the studies that used GBTM to investigate trajectories of medication adherence showed displayed significant heterogeneity. Furthermore, the studies that investigated predictors of those trajectories used classic logistic regression models. This classic approach simply fits the data in a regression equation. Other advanced methods like RF algorithms have showed to provide better predictive capability and better identification of risk factors. We implemented an extension of GBTM, multi-trajectory group-based modeling, which investigates medication adherence trajectories as well as time-dependent trajectories of the predictors of medication adherence. Identifying how predictors of non-adherence change over time and influence adherence is critical to develop new interventions. Moreover, this research can potentially help identify non-adherent trajectories with equal disease-specific outcomes. For example, currently a simple 80% threshold of Proportion Days Covered (i.e., a rough estimate of the monthly supply of medicine) has been considered

ideal. Nonetheless, it is possible that different trajectories and types of medication yield similar outcomes based on different adherence rates over time.

To address these issues, we conducted a retrospective analysis of longitudinal of more than 10 years of data comprising medication adherence to oral dosage forms using multi-trajectory group-based modeling. This research focused on the drugs classes for conditions for which medication adherence is currently rewarded by the MA and Part D Plan Star Ratings program. These include:

a) Diabetes: biguanides, a sulfonylureas, thiazolidinediones, DPP-IV inhibitors, incretin mimetics, meglitinides, or SGLT2 inhibitors.
b) Hypertension: Angiotensin converting enzyme inhibitors (ACEI), angiotensin receptor blocker (ARBs), or direct renin inhibitors (DRIs).
c) Blood cholesterol: statins.

Additionally, we identified the patient- and contextual-related factors that influenced different trajectory membership. Figure 9 encompasses a conceptual framework and the focus of each specific aim of the present study.

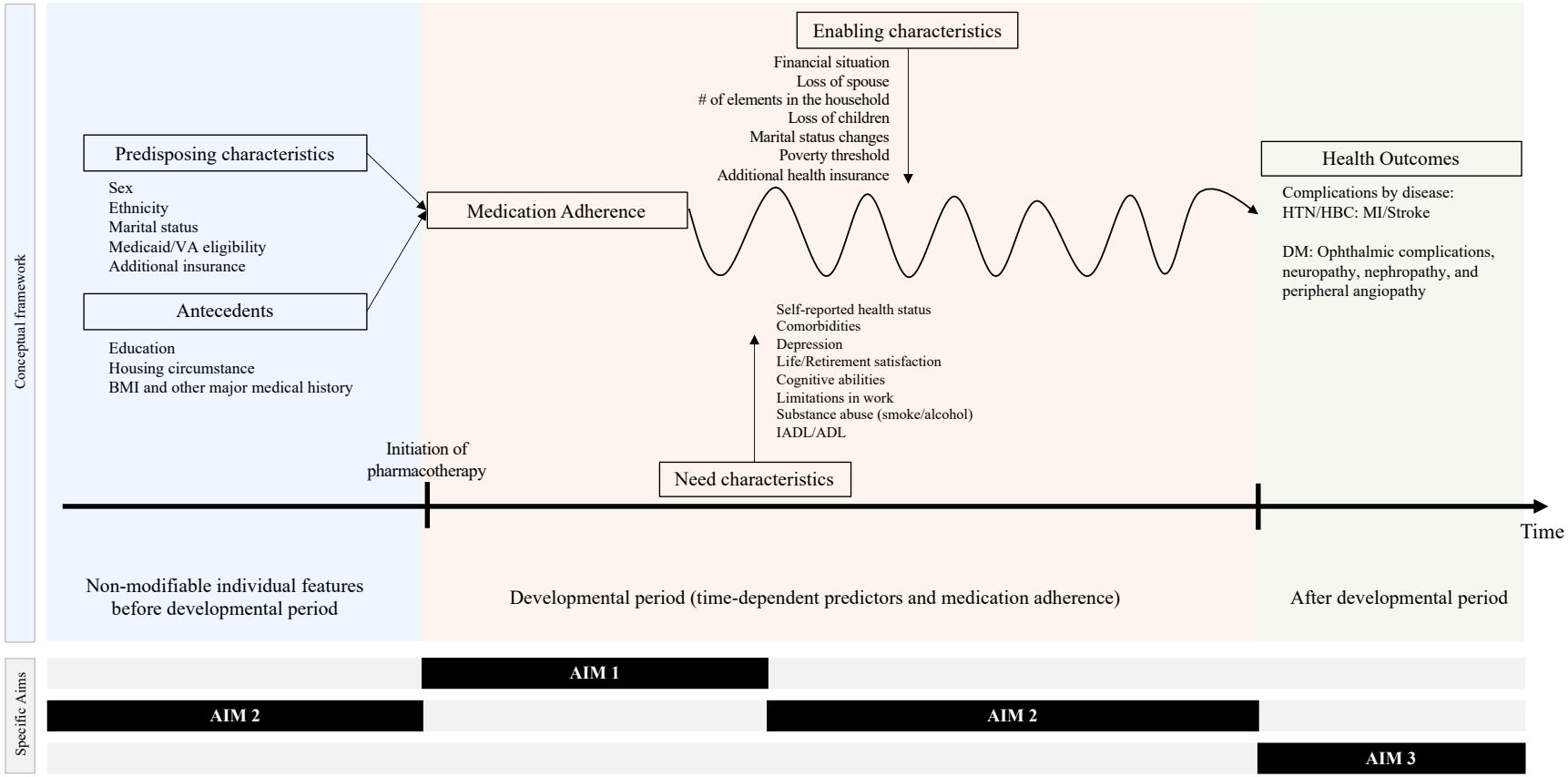

Figure 9 - Conceptual framework and specific aims (HTN – hypertension, DM: diabetes, HBC: high blood cholesterol, MI: myocardial infarction)

Furthermore, we investigated whether random forest algorithm provides a better predictive capability of identifying fixed and time-dependent factors that are associated with each medication adherence trajectories. Finally, we examined whether the identified medication trajectories have significantly different economic burden and health outcomes in the predictive model. This allowed a contrast between non-adherent trajectories with the currently perfect or close to perfect trajectories (with >80% medication adherence threshold). The specific aims (SA) of this study are summarized as follows:

Specific Aim 1: Trajectories group model in chronic patients taking oral-dosage forms

- Specific Aim 1a: Estimate a group-based trajectory model to identify the number of medication adherence trajectories in patients initiating pharmacotherapy for different diseases:
    - Diabetes (biguanides, a sulfonylureas, thiazolidinediones, DPP-IV inhibitors, incretin mimetics, meglitinides, or SGLT2 inhibitors).
    - Hypertension (Angiotensin converting enzyme inhibitors (ACEI), angiotensin receptor blocker (ARBs), or direct renin inhibitors (DRIs)).
    - High blood cholesterol (statins).
- Specific Aim 1b: Compare the proportion of patients traditionally considered as adherers and non-adherers with the medication adherence trajectories identified in SA1a.

Specific Aim 2: Identification of time-fixed and time-varying predictors of medication adherence trajectories

- Specific Aim 2a: Identify the modifiable and non-modifiable determinants of medication adherence are associated with medication adherence trajectory membership.
- Specific Aim 2b: To identify the concurrent trends of the time-varying risk factors that are associated with medication adherence trajectory membership.
- Specific Aim 2c: Build a data visualization tool displaying the evolution of the time-dependent predictors and the medication adherence trajectories.

Specific Aim 3: Predictive model linking medication adherence trajectories to health outcomes:

- Specific Aim 3a: Examine the relationship between medication adherence trajectories and health outcomes, by comparing two classification methods: logistic regression and random forest algorithms.
- Specific Aim 3b: Determine predictive ability by comparison of the c-statistic to identify the best predictive model and examine the strength of association between medication adherence trajectories and outcomes, including myocardial infarction and stroke for all models, and ophthalmic complications, nephropathy, neuropathy, and diabetic peripheral angiopathy as diabetes-specific outcomes.

This study was adequately powered to detect meaningful medication adherence differences between groups, from a nationally representative longitudinal panel study surveying a population

of aging individuals in the US. This dataset allows linkages to administrative claims from the CMS, which were used in this research study.

3. METHODS

3.1 STUDY DESIGN

This research work followed a retrospective longitudinal observational study design employing repeated measures in participants over a given period. The retrospective nature of study implies that study participants have already experienced events that are of relevance, namely initiation of medication of interest, utilization of health care resources, and significant events in each patient's life that possibly shaped chronic medication use patterns. The study design includes from the HRS, which is a longitudinal panel study with a representative sample of approximately 20,000 people in America sponsored by the National Institute of Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. Additionally, the HRS data was linked with administrative health care claims pertaining to Medicare Parts A, B, and D. Figure 10 displays a simplified description of the database relationship of the different data sources.

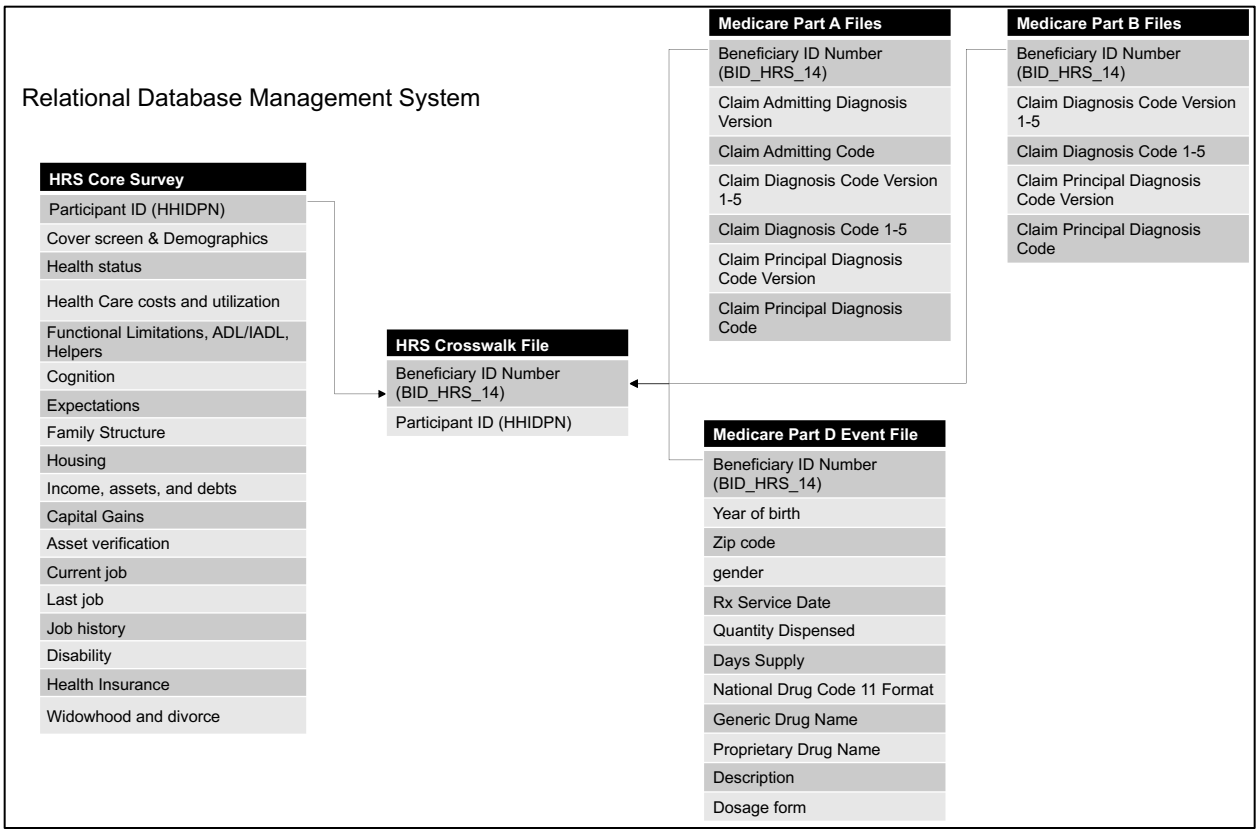

Figure 10 - Database Relationship Management Scheme

3.2 STUDY POPULATION AND DATASET

The HRS survey primarily designed to inform research on aging and provide information about the individual aging experience in the United States. Thus, the survey combines information about the general health, retirement, disability, resources, and family support of aging adults, collecting data every 2 years. The dataset provides general information on a variety of the related topics. The content areas of the survey used to retrieve specific variables for this research work are listed in Table 4 with their respective section identifier per year.

Table 4 - Health and Retirement Study Core Survey: List of the content areas by years

Content area	2008	2010	2012	2014	2016
Cover screen & Demographics	A & B	A & B	A & B	A & B	A & B
Health status	C	C	C	C	C
Health Care costs and utilization	N	N	N	N	N
Functional Limitations, ADL/IADL, Helpers	G	G	G	G	G
Cognition	D	D	D	D	D
Expectations	P	P	P	P	P
Family Structure	E	E	E	E	E
Housing	H	H	H	H	H
Income, assets, and debts	Q	Q	Q	Q	Q
Capital Gains	R	R	R	R	R
Asset verification	U	U	U	U	U
Current job	J	J	J	J	J
Last job	J	K	K	K	K
Job history	J	L	L	L	L
Disability	M	M	M	M	M
Health Insurance	N	N	N	N	N
Widowhood and divorce	S	S	S	S	S

Each letter represents the section naming used in each collection year

While it would be tempting to include many variables collected in every two years, the HRS is marred by missing data. Effectively, several variables have over 70% of missing data. Considering that with study drop-outs and incoming participants, there is an average of 20,000 respondents/biennium, we would only be able to include data from potentially 6,000 patients. In addition, the prevalence of hypertension does not exceed 63.1% in the population >60 years old, 11.4 for hypercholesterolemia, and 22% for diabetes type II. [118-120] Therefore, being overinclusive in the variables from each section of the HRS would potentially result in sample sizes per disease smaller than those required for an adequately powered study.

As described previously, these are fundamental aspects that the literature describes as potential factors that influence health behaviors, such as medication adherence, within the context of the Andersen's behavior model of health services. [1,14,15] This study focused on participants who are initiating or have already initiated pharmacotherapy for diabetes, hypertension, and high blood cholesterol. The inclusion criteria are described below:

Inclusion criteria

- Age above 65 years in January 2008
- Linked HRS survey response to Medicare claims.
- At least one claim 6 months prior to January 1, 2008, for the following drug classes: biguanides, sulfonylureas, thiazolidinediones, DPP-IV inhibitors, incretin mimetics, meglitinides, SGLT2 inhibitors, ACEIs, ARBs, DRIs, and statins.

3.3 DATA ANALYSIS

Because the HRS includes several cohorts of participants, we considered the first fill for any of the target drugs as the index date (T_0) for analytical purposes. Participants who had not entered the study until the first fill were marked as missing values until the first record, or otherwise removed from the analysis. Thus, the monthly PDC trajectory calculation was considered as time elapsed after the index date, instead of a particular date interval. Medication adherence was estimated using SAS 9.4[121], while the group-based trajectory analyses were computed using STATA 17.0 MP and the STATA plugin traj, assuming a significant level $\alpha = 0.05$. [122,123] Finally, the caret package and dplyr, ranger, magrittr, tidyverse, xgboost, and caTools dependencies for R were used to investigate the association between trajectory assignment probability and health outcomes. [116,124]

3.3.1 SPECIFIC AIM 1: TRAJECTORIES GROUP MODEL IN CHRONIC PATIENTS USING CHRONIC MEDICATIONS

Linked Medicare administrative claims data were obtained from the participants in the HRS. Pharmacy claims were obtained from the Part D Event file, which allowed identifying claims of the medications of interest in this study: oral medications for hypertension, high blood cholesterol, and diabetes. The drugs of interest were identified based on the CMS Star Ratings Part C and D Performance Data Technical Notes pertaining to the CMS Star Ratings Program and summarized list with drug name and drug class is available in Appendix 1.[48] Table 5 summarizes the drug classes considered for each specific condition.

Table 5 - Drug classes included in the study by condition

Indication	Drug Class
Diabetes Medications	Biguanides
	DPP-4 inhibitors
	GLP-1 receptor agonists
	Meglitinides
	Sodium glucose co-transporter2 (SGLT2) inhibitors
	Sulfonylureas
	Thiazolidinediones
Hypertension Medications	ACE inhibitor medications
	ARB medications
	Direct renin inhibitor medications
High Blood Cholesterol Medications	Statins

DPP-4: Dipeptidyl peptidase 4; GLP-1: Glucagon-like peptide 1; SLGT2: Sodium-glucose cotransporter-2

Using the drug benefit administrative claims, medication adherence was estimated using the Proportion of Days Covered (PDC), as recommended by the Pharmacy Quality Alliance

(Equation 4). Medication adherence was then calculated in intervals of 30 days according to the calendar after the first fill. The PDC method of estimating adherence allowed to adjust cases of oversupply when the refill of the medication occurred before the end previous fill supply.

$$PDC = \frac{Total\ Days\ Drug\ Available}{Days\ of\ Follow\ Up}$$

Equation 4 - Proportion of Days Covered

The ISPOR's checklist for studies informed the calculation method of medication adherence in this study. [125] However, given the absence of explicit guidelines on how to conduct research on multi-medication adherence, [126] we calculated monthly PDCs for each drug grouped by pharmacotherapeutic class. This approach prevents ignoring medication switching that can occur within drug classes, assuming that drug switch is more likely to happen in patients taking drugs from the same pharmacotherapeutic class. Moreover, previous research suggested that estimating all [drug class average] PDC provides a more accurate measurement of multiple medication adherence. [126] The group-based trajectory model for medication adherence of each pharmacotherapeutic class was obtained through the following stages:

1) *Estimation of (within class average) monthly medication adherence:*

   Choudry and colleagues' approach to capturing concurrent adherence of multiple medications was implemented by adapting the SAS code suggested by Chu and colleagues[127,128]:

$$\text{PDC (by drug class)} = \sum_{i=1}^{n} \frac{(\text{PDC Drug 1}) + (\text{PDC Drug 2}) + (\text{PDC Drug n})}{n}$$

Equation 5 - Proportion of Days Covered (by drug class)

2) *Estimation of the number of trajectories:*

Models with 2 through 7 trajectories were estimated for each drug class. Even though the identification of the exact number of trajectories is in itself an unrealistic endeavor (there is no "true" number of trajectories), this method clustered participants with similar distinctive features in their monthly medication adherence behavior. [98] The parsimoniousness of the number of trajectories was assessed by calculating the Bayesian Information Criterion (BIC). [56,57] This measure estimates the logarithm of the maximum likelihood function (L) and it is penalized by the product of the logarithm of the sample size (n) and the number of parameters in each model (k) (Equation 6).

$$BIC = -2\log(L) \mp \log(n) * k$$

Equation 6 - Bayesian Information Criterion (BIC)

Maximum likelihood refers to the process to best identify how the data are distributed. In GBTM, the model fit refers to the optimal number of trajectories that the model is trying to fit according to the distribution of continuous data. When *forcing* the number of trajectories, the likelihood function will determine how each observation is either close or far away from each "average" trajectory. If the datapoints of the observations are far away from the average of each proposed distribution in each assumed trajectory, that means that the *likelihood* of observing those values is low. This would yield a low maximum likelihood estimate. On the other hand, if the number of trajectories were able to encompass the greatest number of observations (i.e., the distance of the observations of the "average" of trajectories were low), then the maximum likelihood would be higher. The higher the maximum likelihood, the better the model

explains the data. The maximum likelihood estimate is a good indicator for model selection – ideally, the best model is the one with the highest maximum likelihood estimate. Complex models that simply encompass a high number of parameters can artificially appear to fit well. [129] This is also true in Group-based trajectory model: if we just increase the number of trajectories, the number of observations "covered" in the model would increase – theoretically, it would improve when we reached *n* numbers of parameters. Because the goal of modeling is to develop the simplest model that best explains the data, BIC is an ideal goodness of fit measure. Model selection is based on the smallest value of BIC. The model with the least negative BIC parameter indicates the model with the most adequate number of probabilistic medication adherence trajectories (model adequacy). [56] Additionally, a Bayesian approach was implemented to determine the appropriateness of the model with the seemingly optimal number of trajectories that can also be considered as absolute model fit statistics. [130] Consequently, the following Bayesian parameters were estimated: average posterior probabilities of each trajectory, odds of correct classification and observed classification proportion versus the expected classification proportion expected. According to Nagin, if each trajectory in the model yields >70% of average posterior probability and 5 and higher odds of correct classification, the model is deemed as having good fit. [56]

3) *Estimation of the polynomial function of the trajectories:*

In the previous step, all models were estimated assuming that each trajectory would follow a quadratic function. In this step, the shape of the trajectory was determined by specifying the better fitting polynomial function. The highest order that indicates a *p*-

value over 0.05 will be considered the better fitting polynomial function, The STATA® plugin to calculate GBTMs, calculates for each polynomial function of each trajectory an estimate, standard error, and respective *p*-value. [56,122] The selection of the polynomial function for each of the trajectories was based on a combination of inspection of the standard errors and comparison of computed BICs. A trajectory displaying large standard errors in the linear and quadratic functions is indicative that the parameters in addition to the zero order function are not required to describe this particular trajectory. [56] Similarly, a deterioration of the BIC of the model following higher order polynomial functions for the same number of trajectories is suggestive that increasing the number of parameters *k* overpowers any possible improvement in model fit. [56]

4) *Probabilistic group assignment:*

As previously mentioned, the elicitation of clusters of individuals based on seemingly similar behavior is a probabilistic approach. This means that individuals are not deterministically assigned to a particular trajectory. Consequently, we determined trajectory group assignment by estimating the *posterior* likelihood of each participant to belong to a particular medication adherence trajectory. The trajectory group assignment is based on the Bayes' Theorem while also accounting for apparent group size. [56] Accordingly, a larger the posterior probability of trajectory membership for a small trajectory requires that the behavior of interest be so consistent throughout the period of analysis. Bayesian-based posterior probabilities consistently adjust for group size, so that the sheer group size in itself is not deemed as indicator of probability of group membership. [56]

3.3.2 SPECIFIC AIM 2: MULTI-TRAJECTORY GROUP-BASED MODEL FOR MEDICATION ADHERENCE AND PREDICTOR TRAJECTORIES

We investigated which covariates are associated with trajectory group membership. According to the conceptual framework present in the last chapter, some patient attributes are non-modifiable characteristics (i.e., predisposing characteristics and antecedents), while others such as enabling, need, and provider/care characteristics can change with time. We implemented two different methods to investigate how modifiable and non-modifiable characteristics can influence and predict medication adherence trajectories.

Non-modifiable individual features

To investigate how non-modifiable features influence medication adherence trajectory membership for each of the 3 disease states, a risk factor variation was implemented in the group-based trajectory modeling. In essence, this implementation performs a generalized logistic regression to each of the group-based trajectory models, in which time-stable covariates are tested for their ability to chance group-membership probability. [122] A generalized logistic regression is an ideal approach because the parameters for each trajectory θ_j of the multinomial logistic regression are able to denote the probability (P) of an individual i's membership in group j ($\pi_j(x_i)$), given the vector of variables that determine trajectory group membership (x_i): [56,131,132]

$$P(Y_i) = \sum_{j}^{i} \left[\frac{e^{x_i \theta_j}}{\sum_j e^{x_i \theta_j}} \right] P^j\,(Y_i).$$

Equation 7

Such that, the effect each vector of non-modifiable risk factor over time is modeled without loss of generality $\theta_1 = 0$: [122]

$$\pi_j(x_i) = \frac{e^{x_i\theta_j}}{\sum_j e^{x_i\theta_j}}$$

Equation 8

Each risk factor was investigated individually, followed by an adjusted model including all covariates found to be statistically significant in predicting membership to least one medication adherence trajectory. For each risk factor, regression estimates, odds ratios, standard errors, and *p*-values were estimated to demonstrate the strength of association between each risk factor and trajectory membership.

To investigate the statistical significance of risk factors and their influence in trajectory membership, the traj plugin for STATA 17 MP returns an output that is interpretable for binary variables only. For that reason, dichotomized *dummy* variables of the non-modifiable characteristics were generated.

To ensure that the influence of some characteristics that could eventually change with time were still able to be analyzed in combination with the non-modifiable characteristics, a *dummy* variable comprising the latest observation of such characteristics was generated. The complete list of non-modifiable characteristics is presented in Table 6.

Table 6 - Non-modifiable Characteristics Data Dichotomization

Characteristics	***Dummy* variable**	**Recoding**
Sex	Being a female	0 Male 1 Female
Born in the US	Foreign born	
Race	Non-white	0 White 1 Non-white
Hispanicity	Hispanic	0 Non-Hispanic 1 Hispanic
Marital status	Not married	0 Married 1 Never married
Medicaid Eligibility	Medicaid beneficiary	0 Not Medicaid beneficiary 1 Medicaid beneficiary

Characteristics	*Dummy* variable	Recoding
Poverty index	Lives below poverty threshold	0 Household above poverty threshold 1 Household below poverty threshold
CHAMPUS/VA Eligibility	CHAMPUS/VA beneficiary	0 Not CAMPUS/VA beneficiary 1 CAMPUS/VA beneficiary
College education	No college education	0 College/Graduate school education 1 Less than college education
Housing status	Not homeowner	0 Homeowner 1 Not homeowner
	Lives in mobile home	0 House/Apartment 1 Mobile home
	Lives in nursing home	0 Not living in nursing home 1 Self or partner living in nursing home
BMI	Overweight	0 Not excessive weight 1 Excessive weight
Cognitive Ability	Cognitive impairment	0 Without cognitive impairment 1 With cognitive impairment
Cancer survivorship status	Cancer ever	0 Never had cancer 1 Cancer survivor
Stroke survivorship status	Stroke survivor	0 Never had a stroke 1 Stroke survivor
Heart problem survivorship status	Heart problems survivor	0 Never had heart problems 1 Heart problems survivor

Given the possibility of multicollinearity, the variance inflation factor (VIF) was computed to determine by how much each risk factor estimate ($\pi_j(x_i)$) is increased because of high correlation with other risk factors. When VIF is equal to 1, the coefficient of determination (R^2) = 0, which means that the risk factor is not linearly related to other variables. [133] As a rule of thumb, a VIF greater than 5 is indicative of multicollinearity. [134] Nevertheless, we first explored the statistical association of each risk with each group-based trajectory individually. Once statistical significance was demonstrated, we proceeded to compute the VIF and R^2 of each risk

factor to determine the presence of multicollinearity for the risk factors in each adjusted group-based trajectory model.

In complex models such as the ones included in this research work, the interpretation of the odds ratios in influencing the membership to non-adherent trajectories compared to perfectly adherent trajectories is inexorably complicated. Thus, dummy observations were created to represent risk scenarios and appended to each respective dataset. Since there are no dummy or imputed PDC measurements, these observations had no influence in the estimation of the risk factor trajectory model. However, the traj add on still outputs probability of trajectory membership for each dummy observation. The comparison of the imputed probabilities with the trajectory membership probabilities of the models estimated in previously was interpreted as the cumulative effects of each risk factor.

Modifiable individual features

A multi-trajectory group-based model was implemented to identify how time varying features influence the probability of membership in each medication adherence trajectory of medications for all three diseases: hypertension, diabetes, and high blood cholesterol. The multi-trajectory group-based model includes the developmental trajectory the previously identified medication adherence and plots the changes of the time-varying predictor simultaneously. The difference between this approach and the previous group-based trajectory analysis is that once each trajectory group membership is established, conditional probabilities of membership are calculated for the second, third, and n^{th} predictor of the outcome of interest (Table 7). The models proposed in Specific Aim 1 and 2 were estimated using STATA® 17 MP and the traj add-on. [122,123]

Table 7 - Time-varying characteristics and respective data transformation features

Type of time-varying characteristics	Variables	Dichotomization / Scale
Enabling characteristics	Self-reported health status	5-point scale: 1 - Excellent 2 – Very good 3 - Good 4 - Fair 5 - Poor
	Limitations in work due to health	Yes (1) / No (0)
	Life Satisfaction	5-point scale: 1 – Completely satisfied 2 – Very satisfied 3 – Somewhat satisfied 4 – Not very satisfied 5 – Not at all satisfied
	Retirement Satisfaction	3-point scale: 1 – Very satisfying 2 – Moderately satisfying 3 – Not at all satisfying
	Mental Health	CESD-D 8 Item Scale 1 – lower frequency of depression symptoms 8 – higher frequency of depression symptoms
	Cognitive impairment	Cognitive Status Index 35-point scale (sum of scores): Word recall, Serial 7s test, Counting backwards, naming tasks, and vocabulary questions.
Need characteristics	Poverty threshold	Below (1) / Above (0)
	Family structure • Loss of spouse • Number of resident children • Widowhood Substance abuse • Alcohol consumption • Smoking status Assistance with activities • Instrumental Activities of Daily Living • Activities of Daily Living	 Yes (1) / No (0) Count of children in household Yes (1) / No (0) Number days/week w/ drinks Yes (1) / No (0) Number of activities requiring assistance/can't perform

3.3.3 SPECIFIC AIM 3: PREDICTIVE MODEL LINKING MEDICATION ADHERENCE TRAJECTORIES TO HEALTH OUTCOMES:

This specific aim focused on examining the statistical relationship between medication adherence trajectory groups and disease-specific outcomes. Myocardial infarction and stroke were considered outcomes for all three cohorts: participants taking hypertension medications, statins, and diabetes medications. Diabetes-specific outcomes were considered for the cohort taking diabetes medications. Participants who suffered from these outcomes were flagged by the identification of the corresponding diagnosis code. ICD-10 codes implementation was mandated by HIPAA requirements in 2015, so outcomes were identified by both ICD-9 codes (2008-2014) and ICD-10 codes (2015-2016). Expert elicitation informed the setting in which diagnoses codes could be retrieved from. Myocardial infarction and stroke are typically outcomes that require inpatient/emergency care, which justified flagging patients who have experienced these outcomes in the inpatient CMS file-only. Conversely, diabetes-specific outcomes such as ophthalmic complications, nephropathy, neuropathy, and peripheral angiopathy are clinical events that depending on the severity can be managed either in an inpatient or outpatient setting. Thus, participants who have experienced diabetes-specific outcomes were flagged using both inpatient and outpatient CMS-files. The specific diagnoses codes used to flag each outcome are listed in Table 8.

Table 8 – ICD-9 and ICD-10 codes used to flag outcomes

	Outcome	ICD-9 Codes	ICD-10 Codes
Hypertension medications / Statins / Diabetes medications	Myocardial infarction	410.x	I21.x I22.x
		430.x 431.x 433.x 434.x 435.x	I60.x I61.x I63.x I64.x G45.x
	Stroke	436.x	H34.1
Diabetes medications			E11.31x E11.36 E11.39 E10.31x
	Ophtalmic complications	250.5x	E10.36 E10.39
	Renal complications	250.4x	E11.29 E10.29 E11.21 E10.21
	Neuropathy	250.6x	E11.4 E10.4
	Peripeheral angiopathy (DFU w/ or w/o gangrene)	250.7x	E11.51 E10.51

One of the ways to investigate the link between medication adherence trajectories and health outcomes is based on regression models:

- Binomial logistic regression model: for categorical outcomes, such as complications, death.

  i. For each of the models computed, the parameter estimates (regression slopes, standard error, 95% confidence interval, and *p*-values) were reported for each covariate, indicating a positive or negative association between the covariate with the outcome. Odds ratios and 95% confidence intervals were presented to demonstrate the strength of the association between each predictor and the outcome.

  ii. The predictive accuracy of a binary model (identified using logistic regression or other procedures) is determined by the c-statistic (c-stat). For our study, the c-stat refers to the probability of a randomly selected patient who suffered an MI, for example, should have a higher predicted probability of having suffered an MI compared to another randomly selected patient who did not suffer an MI. [135,136] C-stat is equivalent to the area under the curve (AUC) of the receiver operating curve (ROC). Each value of the predicted probability of the outcome enables the identification of a threshold: in threshold for each probability (ranging from 0-1), a dichotomization is done (outcome or no outcome). Therefore, the ROC is the graphical representation of all thresholds according to their sensitivity vs 1-specificity. ROC and the respective c-stat was determined for all models. The model with largest c-stat value also indicates the highest predictive probability.

In addition to logistics regression model, several random forest models were computed to investigate the association of the medication adherence trajectories and health outcomes. To

assess external validity of the predictive models obtained via random forest algorithms, an 80:20 holdout sample was created, based on every 5th observation of the dataset until the equivalent of 20% of the total observations.

Original RF algorithm

The original RF algorithm computed all possible decision trees to predict the outcome of interest. We explored the optimal number of forest by assessing the number of trees after which there is no variation in the predictive ability. For the situations where the outcome is a continuous variable (e.g.: total medical costs), we computed the average out-of-bag predictions. The fit of the RF models predicting continuous outcomes was assessed by out-of-bag cross validation, which allows us to compute a coefficient of determination (R^2). The adequacy of the RF models investigating categorical outcomes was determined by the AUC.

Boosted RF models

Boosted RF algorithms (Extreme Gradient Boosting) was also computed to examine the association between medication adherence trajectories and health outcomes.

- **Extreme Gradient boosting**: We fine tuned each model by running models with different numbers of trees (500, 1000, 5000), define the number of splits in each tree (4, 5, 6, and 7), specify the learning rate at 0.1, or 0.3, and investigate the optimal number of iterations (500, 1000, or 2000).

The boosted RF algorithms were assessed for model fit by computing the AUC (c-stat). Binary logistic regressions were computed using STATA 14 MP and the random forest

algorithm-based models were estimated using the caret package and dplyr, ranger, magrittr, tidyverse, xgboost, and caTools dependencies for R. [116,123,124]

3.4 SAMPLE SIZE AND POWER CALCULATIONS

The sample size for this study was based on the difference test specified on Specific Aim 2, differences observed in different medication adherence trajectories. This analysis was powered to identify a minimum of 20% difference between the trajectory groups, assuming an error probability of 5%. Thus, when medication adherence longitudinal patterns yield a minimum of two trajectory groups, each group should be comprised of at least 542 participants. A plot showing minimum sample size per level of statistical power for an effect size of at least 20% difference is shown in Figure 11.

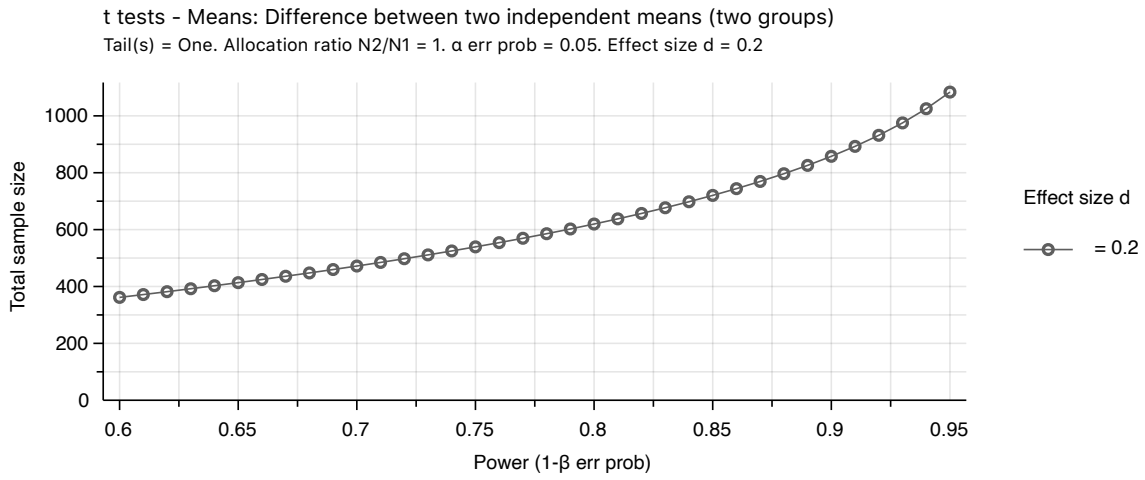

Figure 11 - Sample size calculation per statistical power, assuming an effect size difference of at least 20%.

Previous research in group-based trajectory modelling showed that a total sample of at least 500 participants is deemed ideal for these types of methods. [137] Since the statistical power calculation yields a sample size that is over the minimum identified for a compelling GBTM

approach, we should have included a sample size of at least (2x542) = 1,084 participants for each drug group (hypertension medications, statins, and diabetes medications) .

3.5 ETHICAL CONSIDERATIONS

This research project was submitted to the Virginia Commonwealth University Internal Review Board (IRB) for approval. An additional IRB review request was submitted to the University of Michigan. Both IRB approvals were provided to the Institute of Social Research of the University of Michigan, who is the holder and record keeper of the public access data to the HRS, who authorized merging publicly available to linked administrative health records (Restricted Data Agreement 2021-027). Linked administrative health claims data from Medicare were obtained through CMS's 3rd party data provider ResDAC and MedRIC (CMS Data Use Agreement RSCH-2021-57440).

Electronic data records pertaining to the Public Access component of the HRS were accessed via a virtual desktop infrastructure. Data records were stored in the computer owned by the author of this research work. Adequate measures, including folder encryption and protection of files with passwords were put in place to impede the loss or sharing of data to unrelated people or entities. As established in the CMS Data User Agreement, the restricted pertaining to CMS data will be eliminated from the Graduate Student computer at the end of the DUA license.

4. RESULTS

4.1 SAMPLE POPULATION CHARACTERIZATION

Of the total 42,235 HRS participants in the years 2008-2016, only 11,068 participants were included in the study for having claims for at least one drugs listed in Table 5. The sociodemographic and the other characteristics included as antecedents are presented in Table 8. Regarding sociodemographic characteristics, the participants were described according to their age, sex, race (white or non-white), Hispanicity, marital status, level of education, place of birth (US or non-US), beneficiary status of additional insurance from Medicaid or CHAMPUS/VA, poverty threshold status, type of home, homeownership, nursing home status, autonomy and functionality status (as measured by Activities of Daily Living and Instrumental Activities of Daily Living respectively), cognitive impairment, and depression status. All characteristics listed in Table 9 were described using the latest (or last) observation of each participant.

Missing data was present for the majority of characteristics listed in Table 9. Consequently, the number of observations (n) is indicated below the name of each characteristic. The average participant in the sample (N=11,068) was 76.5 years old (S.E.: 0.095). Moreover, the sample was predominantly female (60.75%), white (76%), non-Hispanic (88.23%), born in US soil (88.61%), and non-college educated (79.55%). Less than the half (44.80%) of the participants were married, 14.89% were either separated or divorced, 35.41% were widowed, and 4.61% were never married. Additional insurance provided by Medicaid was provided to 20.48% of the sample (N=11,068), while complementary benefits for veterans was made available to 2.35%. Of the 1,420 participants (14.84%) whose household income is below the poverty threshold, 643 (45.28%) were Medicaid beneficiaries ($\chi^2 = 1.3 \times 10^3$, p-value < 0.000). A very small percentage of participants lived in mobile homes (8.58%), while the vast majority

lived in homes or apartments. Those who still live independently, a significant majority (66.96%) own their home, while 25.37% rent. Only 856 participants (8.71%) lived in nursing homes. In what concerns comorbidities, Table 9 shows that the vast majority (69.42%) is either overweight or obese (as informed by self-reported weight and height), 23.15% have had cancer in the past, and only 10.11% are currently smokers. Finally, the Activities of Daily Living (ADL) score refers to the ability of an individual to independently take care for oneself by being able to eat, bathe, and be mobile. In addition, the Instrumental Activities of Daily Living (IADL) encompass those requiring more complex thinking and organizational skills. For the ADL component, the HRS includes five tasks: bathing, eating, dressing, walking across the room, and getting in and out of bed. Using a telephone, taking medication, and handling money are the tasks considered in the IADL component of the HRS survey. Accordingly, a significant majority of the participants reported to be completely independent (ADL=0 64.30%), and even greater proportion of participants reported to be highly functional (IADL=0 75.93%). Finally, the components measuring cognitive status and depression symptoms showed that only 5.30% of the participants display impairment, while 20.35% of the participants report clinical depression.

Table 9 - Sample sociodemographic characteristics

Characteristic N = 11,068	**n**	**Freq. %**	**Mean**	**Standard error**	**95 % CI**	
					Upper Limit	*Lower Limit*
Age (n=9,826)			76.499	.095	76.312	76.686
Sex (n=11,068)						
Female	6,724	60.75				
Race (n=11,057)						
White	8,460	76.51				
Non-white	2,597	23.49				
Hispanicity (n=11,058)						
Hispanic	1,302	11.77				

Characteristic N = 11,068	**n**	**Freq. %**	**Mean**	**Standard error**	**95 % CI** *Upper Limit*	*Lower Limit*
Marital status (n=9,824)						
Married	4,401	44.80				
Separated	224	2.28				
Divorced	1,248	12.70				
Widowed	3,479	35.41				
Never Married	453	4.61				
Other	19	0.19				
College educated (n=11,068)						
Has college degree or higher	2,263	20.45				
US Born (n=9,564)						
Yes	8,475	88.61				
Medicaid beneficiary (n=9,798)						
Yes	2,007	20.48				
CHAMPUS/VA beneficiary (n=9,815)						
Yes	231	2.35				
Poverty Index (n=9,609)						
Household income lower than poverty threshold	1,426	14,84				
Type of Home (n=9,678)						
Mobile home	830	8.58				
House/Apartment	8,848	91.42				
Ownership Status of Home (n=9,191)						
Own	6,154	66.96				
Rent	2,332	25.37				
Rent-free with others	593	6.45				
Other	112	1.22				
Currently in nursing home (n=9,826)						
Yes	856	8.71				
Cancer (excluding skin) *(n=9,810)*						
Yes	2,271	23.15				
BMI status *(n=9,340)*						
Underweight	224	2.40				
Healthy Weight	2,632	28.18				
Overweight	3,331	35.66				
Obese	3,153	33.76				

Characteristic N = 11,068	n	Freq. %	Mean	Standard error	95 % CI	
					Upper Limit	*Lower Limit*
Smoking (n=9,749)						
Yes	986	10.11				
Activities of Daily Living Score (n=9,822)						
0 (Completely independent)	6,316	64.30				
1	1,160	11.81				
2	735	7.48				
3	504	5.13				
4	486	4.95				
5 (Totally dependent)	621	6.32				
Instrumental Activities of Daily Living Score (n= 9,822)						
0 (Highly functional)	7,458	75.93				
1	1,035	10.54				
2	605	6.16				
3 (Not functional)	724	7.37				
Cognition Score* (n=9,176)			19.34	0.062	19.221	19.465
With cognitive impairment	486	5.30				
Without cognitive impairment	8,690	94.70				
Clinical depression (CESD-8)** (n=9,432)			1.841	0.022	1.798	1.885
With clinical depression	1,919	20.35				
Without clinical depression	7,513	79.65				

** The cognition score implemented in the HRS ranges from 0 – 35. The dichotomization pertaining to cognitive impairment is based on Herzog and Wallace, who suggested that a score of 8 or less in the composite score should be indicative of cognitive impairment.* [138,139]

*** The CESD-8 (Center for Epidemiologic Studies Depression 8-item) scale is a validated instrument to measure depressive symptoms. Per Steffick and colleagues, a score > 3 is indicative of clinical depression*[140]

4.2 SPECIFIC AIM 1: TRAJECTORIES GROUP MODEL IN CHRONIC PATIENTS INITIATING ORAL-DOSAGE FORMS

Patients taking statins represented the largest cohort (n=8,221), followed by patients taking anti-hypertension medication (n=7,727). The smallest group was the cohort of patients taking

diabetes medications (n=3,146). The number of participants that were included per drug class is presented in Table 10.

Table 10 - Number of participants included in study per drug class

Drug class	Total number of patients
Hypertension medications	**7,727 unique observations**
Angiotensin-converting-enzyme inhibitors	5,170
Angiotensin II receptor blockers	4,198
Hypercholesterolemia	**8,221 unique observations**
Statins medications	8,211
Diabetes medications	**3,214 unique observations**
Biguanides	2,591
DPP-4 inhibitors (single or in combination)	102
GLP-1 receptor agonists	107
Meglitinides	82
SGLT2 inhibitors	66
Sulfonylureas	1,640

DPP-4: Dipeptidyl peptidase 4; GLP-1: Glucagon-like peptide 1; SLGT2: Sodium-glucose cotransporter-2

Furthermore, most of the patients were using medication for more than one disease. In fact, 5,075 (46.68%) of patients of a total 10,873 unique patients were taking all three types of medications and the vast majority of patients taking diabetes medications were also taking statins (n=3,146, 97.88% of the diabetes medications cohort). Figure 12 displays the relative frequencies of concomitant drug use in patients included in the analysis.

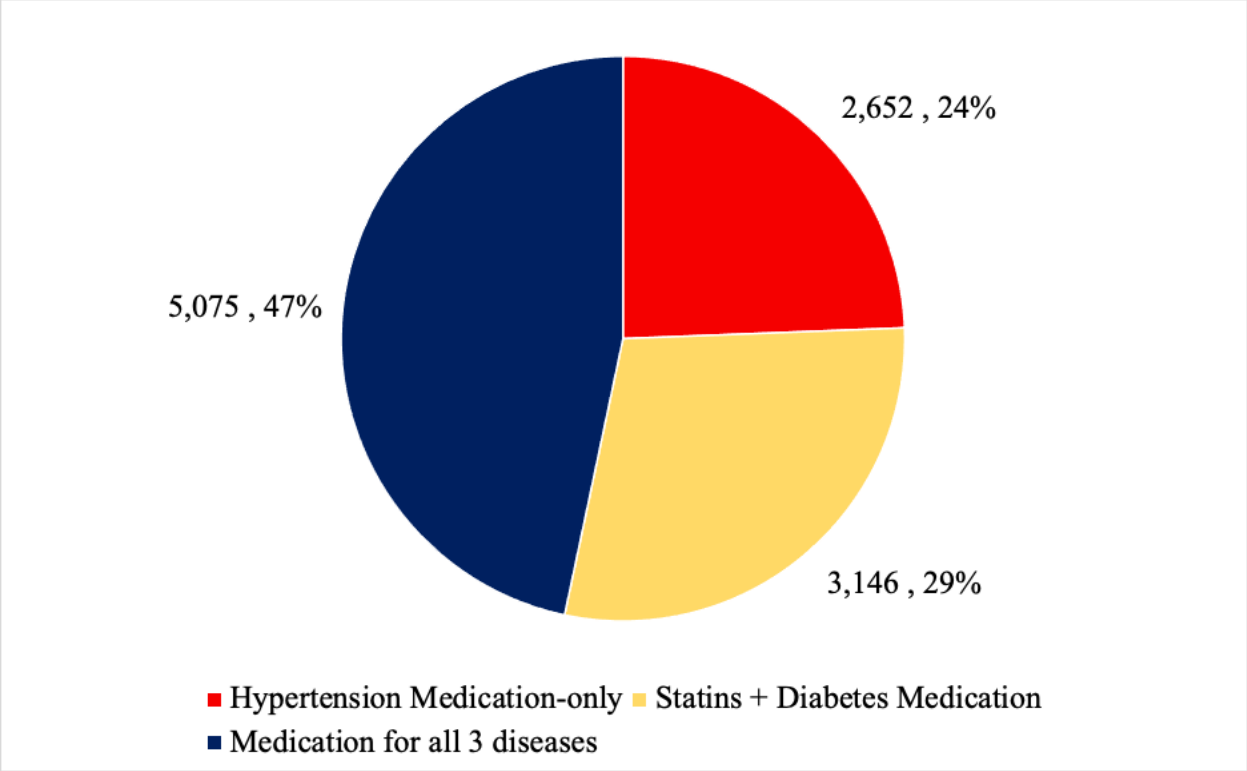

Figure 12 - Proportion of patients taking concomitant drugs

4.2.1 TRAJECTORY MODEL ELICITATION

After the estimation of monthly adherence after the first fill was identified, several models were estimated with trajectories ranging from 2-7, assuming all trajectories followed a quadratic function. Model adequacy was estimated by identifying the model with the smallest BIC value. [56,57] The BICs for each model by disease are listed in Table 11. The model with 3 trajectories yielded the smallest BIC value for hypertension, 5 trajectories for statins, and 6 trajectories for diabetes medications. In some instances, smaller BIC values would be estimated, but the model would falsely converge. Therefore, only the models with true convergence were considered.

Table 11 – Bayesian Information Criteria of the different trajectory models estimated for hypertension medications, statins, and diabetes medications

Disease	Number of trajectories in the model	Bayesian Information Criterion (BIC)
Hypertension medications	2	-662,982.75
	3	**-693,157.13**
	4*	-818,779.21
	5*	-836,981.77
	6*	-799,605.88
Statins	2	-778,043.94
	3	-753,251.00
	4	-792,072.04
	5	**-852,278.33**
	6	-769,680.19
	7*	-778043.94
Diabetes medications	2	-377,122.35
	3	-397,947.11
	4	-387,001.15
	5*	-387,011.61
	6	**-412,841.22**
	7*	-347,113.78

* Model falsely converged.

After the number of trajectories was identified, the function of each trajectory was adjusted. For the hypertension model, all trajectories were found to be statistically significant ($p<0.000$) for quadratic functions. The same was observed for the trajectory model for statins, in which all trajectories were significant ($p<0.000$) assuming quadratic functions. Given the overall shape of observed datapoints in the diabetes model, cubic functions were also tested. Consequently, four trajectories were found to be statistically significant in cubic functions, one assuming quadratic function, and another one assuming 0 order function. Each trajectories' parameters, estimates, standard errors, and *p*-values for all models is presented in Table 12.

Table 12 - Trajectory model parameters

Group	Parameter	Estimate	S.E.	T for H0	*p*-value
Hypertension medications					

Group	Parameter	Estimate	S.E.	T for H0	*p*-value
1	Intercept	1.248	0.017	-73.717	< 0.000
(*Rapid decline*)	Linear	-0.064	0.001	-45.029	< 0.000
	Quadratic	0.000	0.000	11.124	< 0.000
2	Intercept	0.841	0.014	60.144	< 0.000
(Sow decline)	Linear	-0.002	0.001	-4.589	< 0.000
	Quadratic	-0.000	0.000	-18.105	< 0.000
3	Intercept	1.203	0.012	103.12	< 0.000
(*Near-perfect adherence*)	Linear	-0.001	0.000	-1.944	0.052
	Quadratic	0.000	0.000	14.416	< 0.000
Statins					
1	Intercept	1.387	0.016	85.005	< 0.000
(*Rapid decline*)	Linear	-0.094	0.001	-97.999	< 0.000
	Quadratic	0.001	0.000	63.336	< 0.000
2	Intercept	0.820	0.017	49.468	< 0.000
(*Moderate decline*)	Linear	0.028	0.001	31.376	< 0.000
	Quadratic	-0.001	0.000	-70.724	< 0.000
3	Intercept	0.746	0.017	45.106	< 0.000
(*Slow decline*)	Linear	0.037	0.001	53.591	< 0.000
	Quadratic	-0.000	0.000	-76.989	< 0.000
4	Intercept	1.290	0.011	116.102	< 0.000
(*Near-perfect adherence*)	Linear	0.001	0.000	1.244	0.214
	Quadratic	0.000	0.000	11.235	< 0.000
5	Intercept	0.475	0.013	37.266	< 0.000
(Low then increasing	Linear	-0.015	0.001	-28.120	< 0.000
adherence)	Quadratic	0.000	0.000	43.178	< 0.000
Diabetes medications					
1	Intercept	0.512	0.031	16.654	< 0.000
	Linear	-0.026	0.001	-23.364	< 0.000
	Quadratic	0.000	0.000	24.461	< 0.000
2	Intercept	1.001	0.035	28.539	< 0.000
	Linear	-0.051	0.003	-13.525	< 0.000
	Quadratic	-0.007	0.000	-7.610	< 0.000
	Cubic	0.000	0.000	12.669	< 0.000
3	Intercept	1.223	0.035	34.608	< 0.000
	Linear	-0.032	0.004	-8.984	< 0.000
	Quadratic	0.001	0.000	10.892	< 0.000
	Cubic	-0.000	0.000	-18.461	< 0.000
4	Intercept	0.940	0.027	34.336	< 0.000
	Linear	-0.037	0.002	-17.603	< 0.000
	Quadratic	0.001	0.000	15.633	< 0.000
	Cubic	-0.000	0.000	-10.758	< 0.000
5	Intercept	1.289	0.031	41.317	< 0.000
	Linear	-0.037	0.002	-13.380	< 0.000
	Quadratic	0.001	0.000	19.524	< 0.000
	Cubic	-0.000	0.000	-27.419	< 0.000
6	Intercept	1.570	0.003	368.526	< 0.000

The adequacy of the final models was determined by the estimating the average posterior probabilities of each trajectory, the respective odds of correct classification, and observed proportion when compared to the expected classification proportions. All trajectories in all models resulted in average posterior probabilities higher than 70%, and odds of correct classification higher than 5. Table 13 contains the absolute model fit statistics for all the final models.

Table 13 - Absolute group-based trajectory model fit statistics

Trajectory	Group Average Posterior Probability	Odds of Correct Classification	Odds of Correct Classification (weighted by Posterior Probability)	Estimated Proportion of each trajectory	Observed proportion of each trajectory
Hypertension medications					
1	99.21%	522.88	522.30	19.45%	19.47%
2	97.68%	85.72	85.50	32.93%	32.99%
3	98.60%	77.48	77.70	47.62%	47.55%
Statins					
1	98.94%	744.23	739.14	11.14%	11.20%
2	98.23%	382.76	383.70	12.64%	12.62%
3	96.10%	118.77	119.37	17.19%	17.12%
4	96.90%	56.12	56.73	35.73%	35.49%
5	95.41%	68.42	67.37	23.30%	23.58%
Diabetes Medications					
1	99.61%	2,512.09	2,500.27	9.23%	9.27%
2	90.42%	27.29	27.48	25.69%	25.56%
3	93.90%	98.64	97.94	13.50%	13.58%
4	95.16%	96.78	97.14	16.89%	16.84%
5	98.73%	659.13	658.06	10.59%	10.60%
6	92.49%	38.77	38.68	24.10%	24.15%

4.2.2 MEDICATION ADHERENCE TRAJECTORY MODELS

With the model elicitation completed, we used the estimated monthly PDC and actual observed PDC to graphically represent each medication adherence trajectory. Generally,

common features were identified in all medication adherence trajectory models. These include one consistently high adherence trajectory group, a slowly declining adherence group, and a rapidly declining adherence group.

For hypertension medications, the 3-trajectory model suggests the occurrence of one trajectory with relatively high adherence over time (close to 80%), one that slowly declines over time, and a third one with rapidly declining adherence (Figure 13). The *near-perfect adherence* trajectory accounted for the largest group (47.5%, n=3,670), followed by the slow declining adherence group (33%, n=2,550), and low then increasing adherence (19.5%, n=1,507).

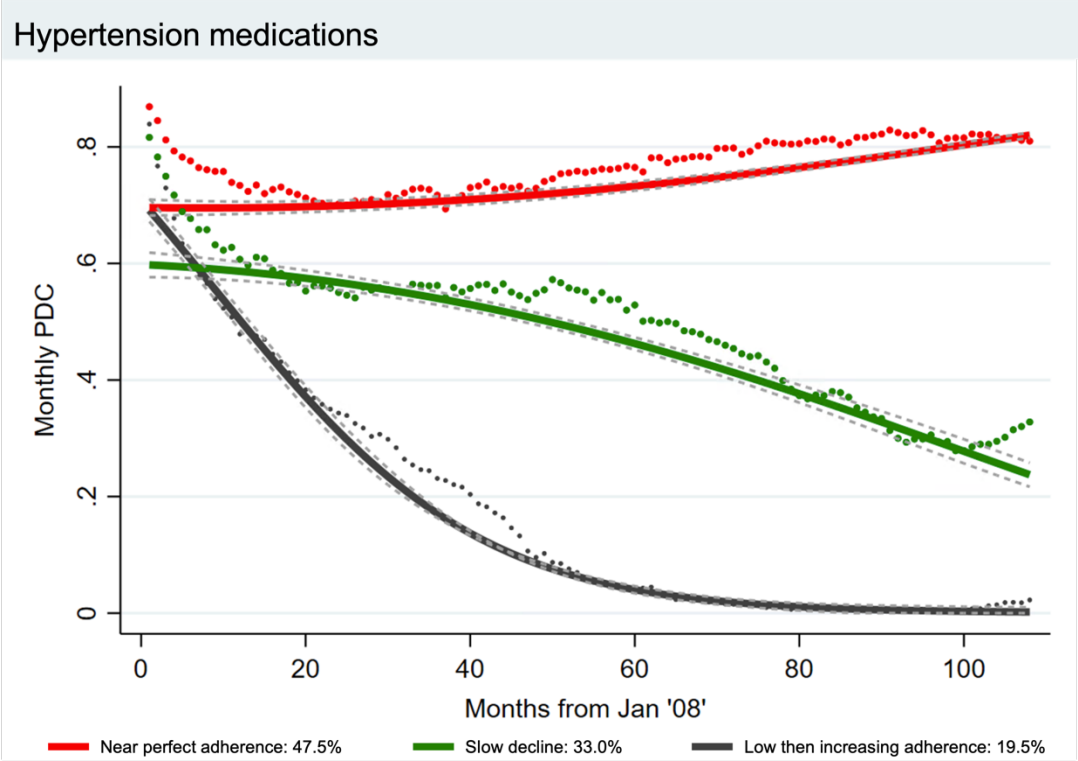

Figure 13 - Hypertension medications: Medication Adherence Group-based trajectory model

In the statins model, the 5 trajectories included the groups common to all models, *near-perfect adherence* (35.5%, n=2,918), slowly declining adherence (17.1%, n=1,406), and rapid declining adherence groups (11.2%, n=921). However, two additional groups were obtained: a

moderately declining adherence (12.6%, n=1,036), and one with a seemingly declining adherence, but that rebounds over time (23.6%, n=1,940) – low then increasing adherence trajectory group (Figure 14). Albeit a smaller percentage, the *near-perfect adherence* trajectory is still the largest group in the model, while the low then increasing adherence trajectory was the second largest.

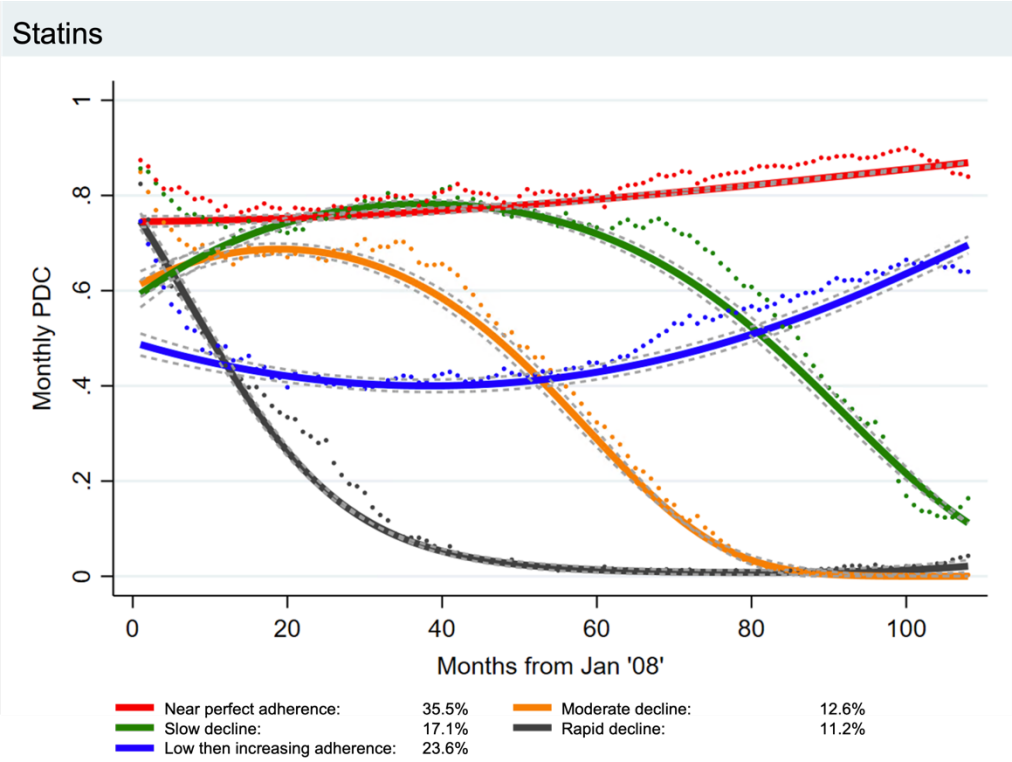

Figure 14 - Statins: Medication Adherence Group-based trajectory model

Like in the statins trajectory model, the diabetes medication model included a *near-perfect adherence* (24.2%, n=778), slowly declining adherence (16.9%, n=543), rapid declining adherence (9.3%, n=299), *moderate decline* (10.7%, n=344), a lower then increasing adherence (13.8%, n=444), and a trajectory that like the previous declined but rebounds to PDC values much closer to the *near-perfect adherence* trajectory (25.1%, n=807). Contrary to the hypertension medication and statins models, this latter trajectory – higher then increasing

adherence, is the largest trajectory group in the model followed by the *near-perfect adherence* group (Figure 15).

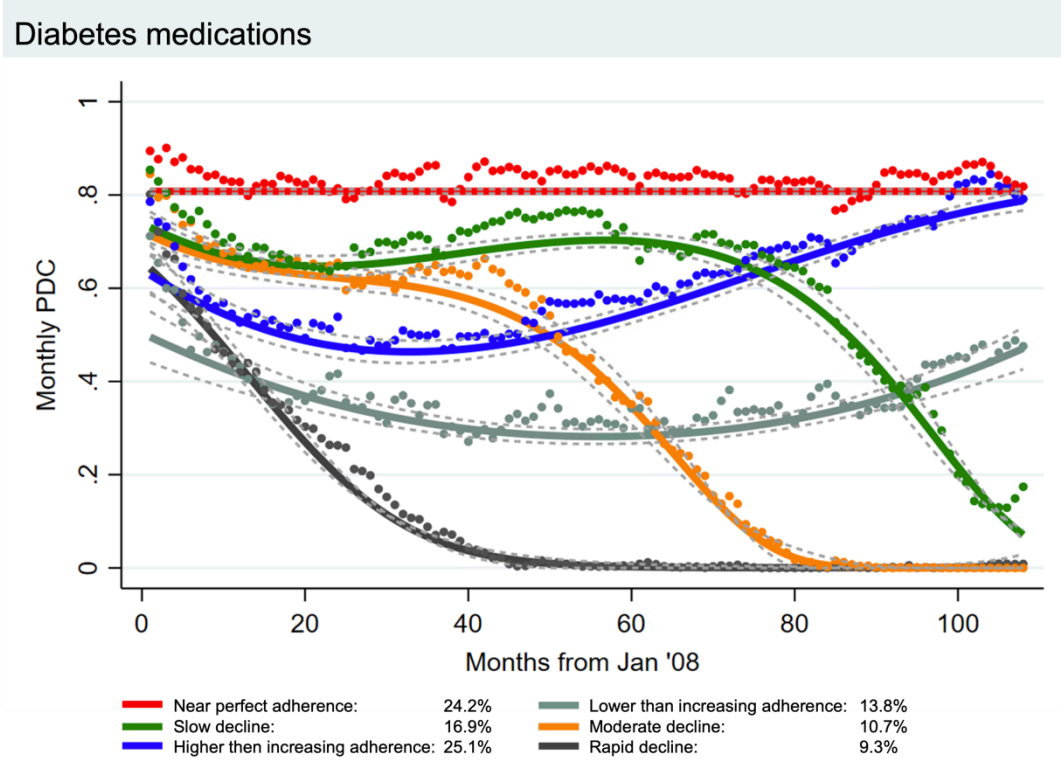

Figure 15 - Diabetes: Medication Adherence Group-based trajectory model

4.2.3 COMPARISON OF GBTM MODELS WITH TRADITIONAL PDC CALCULATION

Using the traditional methods of classification of non-adherence, total PDC estimates for the entire follow-up period 2008-2016 were computed. The proportion of patients traditionally categorized as adherent (PDC > 80%) was 22.44% for hypertension medications, 23.99% for statins, and 18.53% for diabetes medication.

Figure 16 compares the proportion of patients classified as adherent and non-adherent with the proportion of patients following near-perfect adherent trajectories in each model.

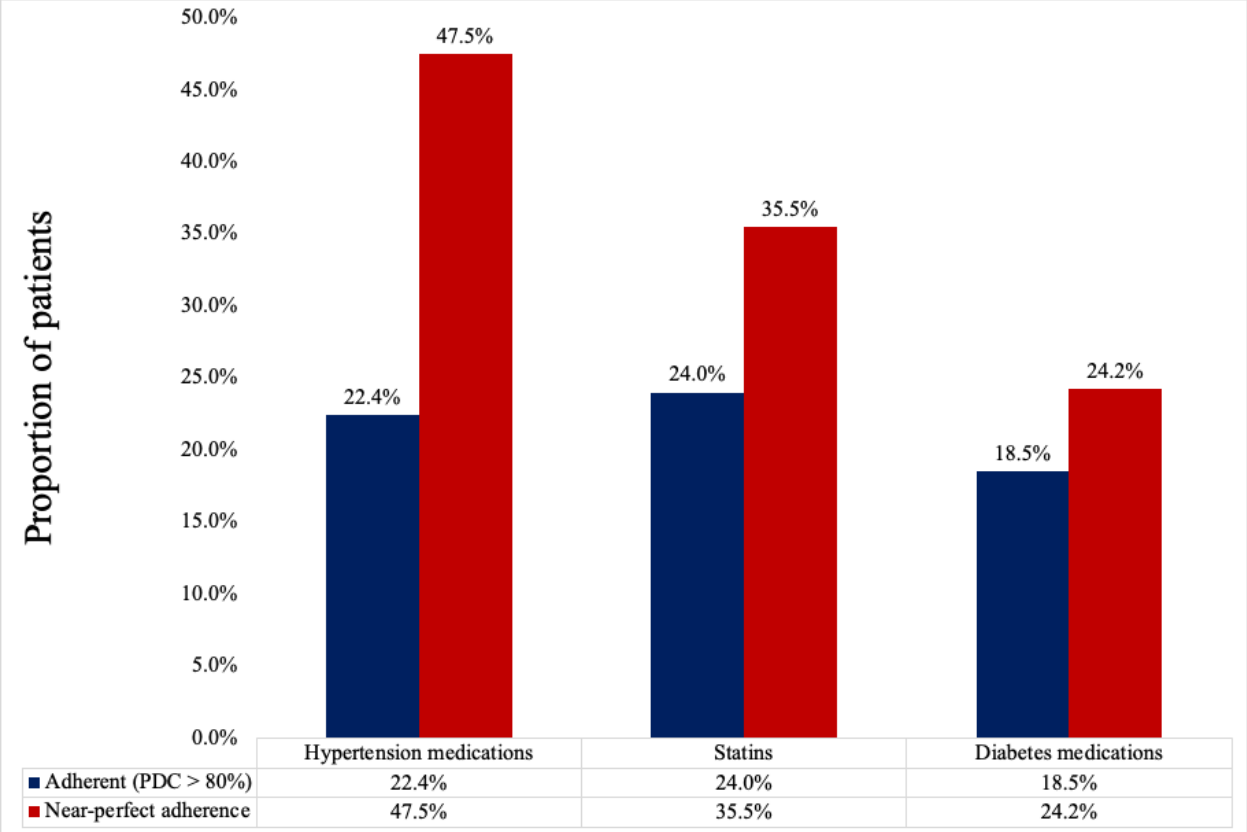

Figure 16 – Comparison of the proportion of patients categorized as adherent using a traditional approach and group-based trajectory models

4.3 SPECIFIC AIM 2: IDENTIFICATION OF TIME-FIXED AND TIME-VARYING PREDICTORS OF MEDICATION ADHERENCE TRAJECTORIES

A risk factor variation was implemented in the estimation of each disease's group-based trajectory model of medication adherence, assuming the *near-perfect adherence* as the reference group.

4.3.1 UNADJUSTED TIME-FIXED RISK FACTOR MODELS

In the unadjusted models investigating the association between covariates and trajectory membership individually, being a female was found to be associated with the *low then increase* trajectory in the statins model ($\pi_{low\ then\ increase}$=0.299, *p*-value<0.001) and with the *low then*

increase ($\pi_{low\ then\ increase}$=0.411, *p*-value<0.001) and the *high then increase* ($\pi_{high\ then\ increase}$=0.248, *p*-value=0.005) trajectories of the diabetes medications model.

Being non-white was found the be associated with the *slow decline* trajectory of the hypertension medications model ($\pi_{slow\ decline}$=0.402, *p*-value<0.001), with all trajectories of the statins model ($\pi_{rapid\ decline}$=0.292, *p*-value<0.001, $\pi_{moderate\ decline}$=0.240, *p*-value=0.002, $\pi_{slow\ decline}$=0.350, *p*-value<0.001, and $\pi_{low\ then\ increase}$=0.779, *p*-value<0.001), and with *low then increase* and the *high then increase* trajectories of the diabetes medications model ($\pi_{low\ then\ increase}$=0.745, *p*-value<0.001, $\pi_{high\ then\ increase}$=0.322, *p*-value=0.001).

The association between trajectories and Hispanic ethnicity was found to be statistically significant for the *rapid decline* trajectory of the hypertension medications model ($\pi_{rapid\ decline}$=-0.349, *p*-value<0.001), *low then increase* trajectory of the statins model ($\pi_{low\ then\ increase}$=0.489, *p*-value<0.001), and the *low then increase* and *high then increase* trajectories of the diabetes medications model ($\pi_{low\ then\ increase}$=0.366, *p*-value=0.003, $\pi_{high\ then\ increase}$ =0.293, *p*-value=0.008).

Trajectory membership and being born outside of the US was an association found to be statistically significant for the *rapid decline* trajectory of the hypertension medications model ($\pi_{rapid\ decline}$=-0.295, *p*-value=0.003), *low then increase* of the statins model ($\pi_{low\ then\ increase}$=0.470, *p*-value<0.001), and *high then increase* trajectory of the diabetes medications model ($\pi_{high\ then\ increase}$=0.342, *p*-value=0.006).

Being unmarried was found to be statistically associated with the *rapid decline* and *slow decline* trajectories of the hypertension medications model ($\pi_{rapid\ decline}$=0.468, *p*-value<0.001, $\pi_{slow\ decline}$=0.350, *p*-value<0.001), *rapid decline*, *moderate decline*, *slow decline*, and *low then increase* trajectories of the statins model ($\pi_{rapid\ decline}$=0.439, *p*-value<0.001, $\pi_{moderate\ decline}$=0.510, *p*-value<0.001, $\pi_{slow\ decline}$=0.294, *p*-value<0.001, $\pi_{low\ then\ increase}$ =0.260, *p*-value<0.001), and the

rapid and *slow decline* trajectories of the diabetes medications model ($\pi_{rapid\ decline}$=0.316, *p*-value=0.009, $\pi_{slow\ decline}$=0.315, *p*-value=0.005).

Not being college educated was found to be statistically associated with belonging to all declining trajectories of the hypertension medications model ($\pi_{rapid\ decline}$=0.566, *p*-value<0.001, $\pi_{slow\ decline}$=0.352, *p*-value<0.001), all of the declining trajectories of the statins model ($\pi_{rapid\ decline}$=0.544, *p*-value<0.001, $\pi_{moderate\ decline}$=0.581, *p*-value<0.001, $\pi_{slow\ decline}$=0.383, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.352, *p*-value<0.001), and all but the *high then increase* trajectory of the diabetes medications model ($\pi_{low\ then\ increase}$=0.284, *p*-value=0.003, $\pi_{rapid\ decline}$=0.536, *p*-value<0.001, $\pi_{moderate\ decline}$=0.534, *p*-value<0.001, $\pi_{slow\ decline}$=0.328, *p*-value=0.008).

Compared to the *near-perfect adherence* trajectories in each model, living below the poverty threshold was found to be associated with membership to all trajectories in the hypertension medications model ($\pi_{rapid\ decline}$=0.333, *p*-value<0.001, $\pi_{slow\ decline}$=0.473, *p*-value<0.001), all of the trajectories in the statins model ($\pi_{rapid\ decline}$=0.622, *p*-value<0.001, $\pi_{moderate\ decline}$=0.411, *p*-value<0.001, $\pi_{slow\ decline}$=0.547, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.650, *p*-value<0.001), and all but the *slow decline* trajectory in the diabetes medications model ($\pi_{low\ then\ increase}$=0.572, *p*-value<0.001, $\pi_{rapid\ decline}$=0.468, *p*-value=0.003, $\pi_{moderate\ decline}$=0.498, *p*-value<0.001, $\pi_{high\ then\ increase}$ =0.333, *p*-value=0.008).

When compared to the *near-perfect adherence* trajectories, the association between trajectory membership and being a Medicaid beneficiary was found to be statistically significant for all non-perfect adherence trajectories in all three models: a) hypertension medications ($\pi_{rapid\ decline}$=0.333, *p*-value<0.001, $\pi_{slow\ decline}$=0.473, *p*-value<0.001), b) statins ($\pi_{rapid\ decline}$=0.802, *p*-value<0.001, $\pi_{moderate\ decline}$=0.699, *p*-value<0.001, $\pi_{slow\ decline}$=0.440, *p*-value<0.001, $\pi_{low\ then}$

$_{increase}$=0.605, *p*-value<0.001), and c) diabetes medications ($\pi_{low\ then\ increase}$=0.570, *p*-value<0.001, $\pi_{rapid\ decline}$=0.732, *p*-value<0.001, $\pi_{moderate\ decline}$=0.353, *p*-value=0.002, $\pi_{high\ then\ increase}$ =0.404, *p*-value<0.001, $\pi_{slow\ decline}$ =0.630, *p*-value<0.001).

Beneficiaries of CHAMPUS/VA were found be more likely to be associated with *rapid decline* trajectory in the hypertension medications model ($\pi_{rapid\ decline}$=0.386, *p*-value=0.049) and the *rapid decline* trajectory of the statins model ($\pi_{rapid\ decline}$=0.829, *p*-value<0.001), and less likely in the *low then increase* trajectory of the statins model, $\pi_{low\ then\ increase}$=-0.721, *p*-value=0.005).

Living in a mobile home was found to be more likely to belong to the *rapid decline* trajectory in the hypertension model ($\pi_{rapid\ decline}$=0.520, *p*-value<0.001) and the rapid and *moderate decline* trajectories in the statins model ($\pi_{rapid\ decline}$=0.321, *p*-value=0.007, $\pi_{moderate\ decline}$=0.323, *p*-value=0.004). Individuals who were not homeowners were found more likely to belong to all declining trajectories in the hypertension medications model ($\pi_{rapid\ decline}$=0.544, *p*-value<0.001, $\pi_{slow\ decline}$=0.442, *p*-value<0.001), all non-perfect adherence trajectories in the statins model ($\pi_{rapid\ decline}$=0.632, *p*-value<0.001, $\pi_{moderate\ decline}$=0.637, *p*-value<0.001, $\pi_{slow\ decline}$=0.416, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.506, *p*-value<0.001), and all non-perfect adherence trajectories of the diabetes medications model ($\pi_{low\ then\ increase}$=0.539, *p*-value<0.001, $\pi_{rapid\ decline}$=0.511, *p*-value<0.001, $\pi_{moderate\ decline}$=0.737, *p*-value=0.002, $\pi_{high\ then\ increase}$ =0.418, *p*-value<0.001, $\pi_{slow\ decline}$ =0.486, *p*-value<0.001).

The association between trajectory membership and living in a nursing home was found to be statistically significant for all non-perfect adherence trajectories in all three models, except for the *low then increase* and *high then increase* trajectories in the diabetes medications model: a)

hypertension medications ($\pi_{rapid\ decline}$=1.549, *p*-value<0.001, $\pi_{slow\ decline}$=1.043, *p*-value<0.001), b) statins ($\pi_{rapid\ decline}$=1.440, *p*-value<0.001, $\pi_{moderate\ decline}$=1.788, *p*-value<0.001, $\pi_{slow\ decline}$=1.128, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.541, *p*-value<0.001), and c) diabetes medications ($\pi_{rapid\ decline}$=1.251, *p*-value<0.001, $\pi_{moderate\ decline}$=0.888, *p*-value<0.001, $\pi_{slow\ decline}$ =1.868, *p*-value<0.001).

Overweight status was found to minimize the likelihood of belonging to non-adherent trajectories in all models. Specifically, the rapid a *slow decline* trajectories in the hypertension medications model ($\pi_{rapid\ decline}$=-0.875, *p*-value<0.001, $\pi_{slow\ decline}$=-0.222, *p*-value<0.001), rapid, moderate, and *slow decline* trajectories in the statins model ($\pi_{rapid\ decline}$=-0.781, *p*-value<0.001, $\pi_{moderate\ decline}$=-0.541, *p*-value<0.001, $\pi_{slow\ decline}$=-0.225, *p*-value<0.002), and rapid and *moderate decline* trajectories in the diabetes medications model($\pi_{rapid\ decline}$=-0.752, *p*-value<0.001, $\pi_{moderate\ decline}$=-0.557, *p*-value<0.001).

Cognitive impairment was found to be statistically associated with belonging to all non-adherent trajectories in all three models: a) hypertension medications ($\pi_{rapid\ decline}$=1.013, *p*-value<0.001, $\pi_{slow\ decline}$=0.709, *p*-value<0.001), b) statins ($\pi_{rapid\ decline}$=0.862, *p*-value<0.001, $\pi_{moderate\ decline}$=1.030, *p*-value<0.001, $\pi_{slow\ decline}$=0.591, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.571, *p*-value<0.001), and c) diabetes medications ($\pi_{low\ then\ increase}$=0.376, *p*-value<0.001, $\pi_{rapid\ decline}$=0.758, *p*-value<0.001, $\pi_{moderate\ decline}$=1.042, *p*-value=0.002, $\pi_{high\ then\ increase}$ =0.264, *p*-value=0.014, $\pi_{slow\ decline}$ =0.557, *p*-value<0.001).

Clinical depression was found to be statistically associated with membership to all non-adherent trajectories in all 3 models. These include membership to the rapid and *slow decline* trajectories in the hypertension model ($\pi_{rapid\ decline}$=0.860, *p*-value<0.001, $\pi_{slow\ decline}$=0.288, *p*-

value<0.001), *rapid*, *moderate*, and *slow decline* and the *low then increase* trajectories of the statins model ($\pi_{rapid\ decline}$=0.944, *p*-value<0.001, $\pi_{moderate\ decline}$=0.581, *p*-value<0.001, $\pi_{slow\ decline}$=0.500, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.372, *p*-value<0.001), and the *low then increase*, *rapid*, *moderate*, and *slow decline*, and *high then increase* trajectories in the diabetes medications model ($\pi_{low\ then\ increase}$=0.564, *p*-value<0.001, $\pi_{rapid\ decline}$=1.163, *p*-value<0.001, $\pi_{moderate\ decline}$=0.786, *p*-value<0.001, $\pi_{high\ then\ increase}$ =0.302, *p*-value<0.001, $\pi_{slow\ decline}$ =0.364, *p*-value<0.001).

Finally, survivorship status to specific health conditions registered slightly mixed results. In hypertension medication models, survivorship of cancer, stroke, and heart problems were risk factors for belonging to all non-adherent trajectories (cancer: $\pi_{rapid\ decline}$=0.298, *p*-value<0.001, $\pi_{slow\ decline}$=0.150, *p*-value=0.015; stroke: $\pi_{rapid\ decline}$=0.686, *p*-value<0.001, $\pi_{slow\ decline}$=0.492, *p*-value<0.001, heart problems: $\pi_{rapid\ decline}$=0.382, *p*-value<0.001, $\pi_{slow\ decline}$=0.274, *p*-value<0.001). In the statin's models, cancer was found to be a predictor for all trajectories but the *slow decline* trajectory ($\pi_{rapid\ decline}$=0.281, *p*-value<0.001, $\pi_{moderate\ decline}$=0.274, *p*-value<0.001, $\pi_{low\ then\ increase}$=-0.250, *p*-value<0.001). Conversely, stroke and heart problems were found to statistically influence membership to all non-adherent trajectories of the diabetes medications trajectory models: (stroke: $\pi_{rapid\ decline}$=0.581, *p*-value<0.001, $\pi_{moderate\ decline}$=0.649, *p*-value<0.001, $\pi_{slow\ decline}$=0.421, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.251, *p*-value<0.001; heart problems: $\pi_{rapid\ decline}$=0.581, *p*-value<0.001, $\pi_{moderate\ decline}$=0.649, *p*-value<0.001, $\pi_{slow\ decline}$=0.421, *p*-value<0.001, $\pi_{low\ then\ increase}$=0.251, *p*-value<0.001).

4.3.2 ADJUSTED TIME-FIXED RISK FACTOR MODELS

Prior to estimating the adjusted risk models for all 3 models, we computed the Variance Inflation Factor and respective R^2 for all risk factors. All risk factors included in each model displayed a VIF < 5, suggesting negligible evidence of multicollinearity. In fact, the average VIF for each model was 1.26 in the hypertension model, 1.27 in the statins model, and 1.30 in the diabetes medications model. The VIF of each covariate for all 3 trajectory models is presented in Table 14.

Table 14 - Variance Inflation Factors and R^2 values of each trajectory model

Hypertension medications			Statins			Diabetes medications		
Variable	**VIF**	**R^2**	**Variable**	**VIF**	**R^2**	**Variable**	**VIF**	**R^2**
Foreign born	1.45	0.31	Being female	1.16	0.14	Being female	1.2	0.17
Non-white	1.19	0.16	Foreign born	1.46	0.32	Foreign born	1.47	0.32
Hispanic	1.53	0.35	Non-white	1.2	0.17	Non-white	1.17	0.15
Not married	1.2	0.17	Hispanic	1.54	0.35	Hispanic	1.59	0.37
Not college educated	1.82	0.45	Not married	1.3	0.23	Not married	1.35	0.26
Below poverty index	1.3	0.23	Not college educated	1.86	0.46	Not college educated	1.77	0.44
Medicaid beneficiary	1.35	0.26	Below poverty index	1.28	0.22	Below poverty index	1.32	0.24
CHAMPUSVA beneficiary	1.01	0.01	Medicaid beneficiary	1.35	0.26	Medicaid beneficiary	1.34	0.25
Lives in mobile home	1.02	0.02	CHAMPUSVA beneficiary	1.01	0.01	Not homeowner	1.22	0.18
Not homeowner	1.26	0.20	Lives in mobile home	1.02	0.02	Living in nursing home	1.12	0.11
Living in nursing home	1.11	0.10	Not homeowner	1.26	0.21	Overweight	1.06	0.06
Overweight	1.06	0.05	Living in nursing home	1.11	0.10	Cognitive impairment	1.99	0.50
Cognitive impairment	1.98	0.50	Overweight	1.05	0.05	Cancer survivor	1.02	0.02
Cancer survivor	1.01	0.01	Cognitive impairment	2.07	0.52	Stroke survivor	1.08	0.08
Stroke survivor	1.07	0.07	Cancer survivor	1.02	0.02	Heart problems survivor	1.07	0.07
Heart problems survivor	1.04	0.04	Stroke survivor	1.07	0.06	Clinical depression	1.07	0.07
Clinical depression	1.06	0.06	Heart problems survivor	1.05	0.05			
			Clinical depression	1.07	0.07			
Mean VIF	1.26		Mean VIF	1.27		Mean VIF	1.30	

The adjusted models showed that the size and statistical significance of each risk factor altered when each covariate is included in the generalized logistic regression for each model.

Hypertension medication risk factors adjusted model

Of the 17 covariates that seemed to predict trajectory membership in the hypertension model, only 11 remain statistically significant in the adjusted model: being non-white, Hispanic, Medicaid or CHAMPUS/VA beneficiary, not homeowner, living in nursing home, having excessive weight, cognitive impairment, being a survivor of cancer, stroke or heart problems, and clinical depression diagnosis. Table 15 includes the estimates, standard errors, adjusted Odds Ratios, and *p*-values of all risk factors included in the adjusted hypertension medications trajectory model. Being foreign born, unmarried, not college educated, living below the poverty threshold, and living in a mobile home were no longer statistically significant in the adjusted model.

However, some characteristics seem prevent membership to non-adherent trajectories. Compared to *near-perfect adherence* trajectory, Hispanic individuals are 0.62 times (S.E. 0.138, *p*-value=0.001) *less* likely to belong to the *rapid decline* trajectory. Similarly, overweight individuals are 0.52 times less likely to belong to the *rapid decline* trajectory than overweight individuals in *near-perfect adherence* (Table 15).

Risk factors that in the unadjusted model appeared to be predictors of both non-adherent trajectories of the hypertension model, were now predictors of a single trajectory. These include being non-white and being a survivor of heart problems, which in the adjusted model only predict membership to the *slow decline* trajectory ($aOR_{non\text{-}white} = 1.62$, p-value<0.001, $aOR_{heart\ problems\ survivor}=1.17$, p-value $= 0.009$). Being Medicaid or CHAMPUS/VA beneficiary were risk factors statistically significant for predicting membership to *rapid decline* trajectory only ($aOR_{Medicaid} = 1.24$, p-value$=0.035$, $aOR_{CHAMPUS/VA}=1.94$, p-value $= 0.003$). Finally, living in a nursing home, having cognitive impairment, being a cancer or stroke survivor or having clinical

depression were risk factors shown to statistically predict membership to both the slow and *rapid decline* trajectories, when compared to the *near-perfect adherence* trajectories.

Table 15 - Time-fixed predictors in the hypertension medications trajectory model

	Slow decline				Rapid decline			
	Estimate	**S.E.**	**aOR**	***p*-value**	**Estimate**	**S.E.**	**aOR**	***p*-value**
Foreign born	-0.036	0.105	0.96	0.731	-0.174	0.135	0.84	0.197
Non-white	0.480	0.072	1.62	**0.000**	-0.027	0.090	0.97	0.762
Hispanic	0.123	0.104	1.13	0.235	-0.474	0.138	0.62	**0.001**
Not married	0.052	0.063	1.05	0.410	0.011	0.077	1.01	0.881
Not college educated	0.115	0.094	1.12	0.223	0.099	0.117	1.10	0.393
Below poverty index	-0.015	0.095	0.99	0.874	-0.081	0.112	0.92	0.469
Medicaid beneficiary	0.115	0.089	1.12	0.197	0.218	0.104	1.24	**0.035**
CHAMPUSVA beneficiary	-0.185	0.232	0.83	0.426	0.664	0.220	1.94	**0.003**
Lives in mobile home	0.273	0.155	1.31	0.078	0.290	0.169	1.34	0.086
Not homeowner	0.037	0.070	1.04	0.600	0.290	0.081	1.34	**0.000**
Living in nursing home	0.561	0.149	1.75	**0.000**	0.802	0.153	2.23	**0.000**
Overweight	-0.110	0.068	0.90	0.105	-0.656	0.075	0.52	**0.000**
Cognitive impairment	0.245	0.092	1.28	**0.008**	0.541	0.106	1.72	**0.000**
Cancer survivor	0.140	0.070	1.15	**0.047**	0.405	0.080	1.50	**0.000**
Stroke survivor	0.162	0.080	1.18	**0.042**	0.295	0.090	1.34	**0.001**
Heart problems survivor	0.156	0.059	1.17	**0.009**	0.108	0.071	1.11	0.129
Clinical depression	0.309	0.078	1.36	**0.000**	0.450	0.089	1.57	**0.000**

Cells in bold with grey background denote statistically significant *p*-values (*p*-value < 0.05)
Cells with orange background denote higher likelihood of membership than near perfect adherence (aOR > 1)
Cells with green background denote higher likelihood of membership than near perfect adherence (aOR < 1)

Statins risk factors adjusted model

Fourteen risk factors remained statistically significant for predicting trajectory membership in the statins adjusted model: being female, non-white, Hispanic, not college educated, Medicaid or CHAMPUS/VA beneficiary, living in mobile home or nursing home, not being a homeowner, overweight, having cognitive impairment, being a survivor of cancer, stroke, or heart problems, and having clinical depression. Table 16 includes the estimates, standard errors, adjusted Odds Ratios, and *p*-values of all risk factors included in the adjusted statins trajectory model. Like in the hypertension medications adjusted model, some risk factors were shown to prevent membership to non-adherent trajectories when compared to the *near-perfect adherence* trajectory. These included being female (except for the *low then increase* trajectory) and being

overweight. Specifically, females who take statins are 0.84 times less likely to belong to the *slow decline* and *moderate decline* trajectories when compared to the *near-perfect adherence* trajectory (Table 16). The *low then increase* trajectory was found to be the exception to this trend, as females were 1.20 times more likely to belong to this trajectory, when compared to *near-perfect adherence* trajectory. Overweight individuals taking statins were found to be 0.78 times less likely to belong to the *slow decline* trajectory, 0.66 times less likely to belong to the *moderate decline* trajectory, and 0.48 times less likely to belong to the *rapid decline* trajectory when compared to the *near-perfect adherence* trajectory.

Four risk factors were shown to increase probability of membership to all non-adherent trajectories, when compared to the *near-perfect adherence* trajectory. Living in a nursing home results in a 1.68 times higher probability of belonging to the *low then increase* trajectory than *near-perfect adherence* ($aOR_{low\ then\ increase}$=1.68, *p*-value = 0.002). The increased probability almost duplicates for the *slow*, *moderate*, and *rapid decline* trajectories ($aOR_{slow\ decline}$ = 2.44, *p*-value<0.001, $aOR_{moderate\ decline}$ = 3.45, *p*-value<0.001, $aOR_{rapid\ decline}$ = 2.22, *p*-value<0.001). Cognitive impairment seemed to increase the probability of belonging to non-adherent trajectories by approximately the same extent ($aOR_{slow\ decline}$ = 1.26, *p*-value=0.034, $aOR_{low\ then\ increase}$=1.33, *p*-value = 0.005, $aOR_{moderate\ decline}$ = 1.73, *p*-value<0.001, $aOR_{rapid\ decline}$ = 1.51, *p*-value=0.001). Lastly, clinical depression was shown to be a risk factor than approximately doubled the probability of belonging to all non-adherent trajectories ($aOR_{slow\ decline}$ = 1.94, *p*-value<0.001, $aOR_{low\ then\ increase}$=1.77, *p*-value<0.001, $aOR_{moderate\ decline}$ = 2.02, *p*-value<0.001, $aOR_{rapid\ decline}$ = 2.04, *p*-value<0.001).

Other risk factors sparsely influenced membership to non-adherent trajectories. Being non-white only seemed to increase probability of belonging to the *low then increase* ($aOR_{low\ then\ increase}$

= 1.87, *p*-value<0.001) and *rapid decline* trajectory ($aOR_{moderate\ decline}$ = 1.33, *p*-value=0.009). Hispanic individuals, on the other hand, only appeared to have increased probability of belonging to *low then increase* trajectory ($aOR_{low\ then\ increase}$ = 1.38, *p*-value=0.004) when compared to the *near-perfect adherence* trajectory. College education was found to increase membership to the *moderate* and *rapid decline* trajectories by approximately 30% compared to *near-perfect adherence* ($aOR_{moderate\ decline}$ = 1.33, *p*-value=0.022, $aOR_{rapid\ decline}$ = 1.30, *p*-value=0.050). Like in the hypertension model, being a Medicaid or CHAMPUS/VA beneficiary only increased probability of membership to the *rapid decline* trajectory (Medicaid: $aOR_{rapid\ decline}$ = 1.42, *p*-value=0.003, CHAMPUS/VA: $aOR_{rapid\ decline}$ = 2.46, *p*-value<0.001). Individuals who were not homeowners or cancer survivors were approximately 30% more likely to belong to the moderate and *rapid decline* trajectories, when compared to the *near-perfect adherence* (Not homeowners: $aOR_{moderate\ decline}$ = 1.35, *p*-value=0.001, $aOR_{rapid\ decline}$ = 1.28, *p*-value=0.011; Cancer survivors: $aOR_{moderate\ decline}$ = 1.29, *p*-value=0.003, $aOR_{rapid\ decline}$ = 1.38, *p*-value=0.001). As observed in the hypertension medications model, being a heart problems survivor meant having increased probability of membership to all non-adherent trajectories except the *rapid decline* trajectory ($aOR_{slow\ decline}$ = 1.28, *p*-value<0.001, $aOR_{low\ then\ increase}$ = 1.19, *p*-value=0.006, $aOR_{moderate\ decline}$= 1.23, *p*-value=0.008).

Table 16 - Time-fixed predictors in the statins trajectory model

	Slow decline				Low then increase				Moderate decline				Rapid decline			
	Estimate	**S.E.**	**aOR**	***p*-value**	**Estimate**	**S.E.**	**aOR**	***p*-value**	**Estimate**	**S.E.**	**aOR**	***p*-value**	**Estimate**	**S.E.**	**aOR**	***p*-value**
Being female	-0.176	0.075	0.84	**0.018**	0.184	0.069	1.20	**0.008**	-0.179	0.083	0.84	**0.032**	-0.154	0.090	0.86	0.086
Foreign born	-0.056	0.128	0.95	0.660	0.105	0.109	1.11	0.340	-0.228	0.150	0.80	0.129	-0.003	0.153	1.00	0.987
Non-white	0.154	0.092	1.17	0.095	0.625	0.079	1.87	**0.000**	0.100	0.102	1.11	0.325	0.284	0.107	1.33	**0.008**
Hispanic	0.024	0.129	1.02	0.850	0.324	0.111	1.38	**0.004**	-0.053	0.148	0.95	0.721	-0.237	0.162	0.79	0.143
Not married	0.021	0.078	1.02	0.791	-0.088	0.071	0.92	0.213	0.158	0.087	1.17	0.069	0.012	0.094	1.01	0.895
Not college educated	0.219	0.113	1.24	0.054	0.169	0.103	1.18	0.100	0.288	0.126	1.33	**0.022**	0.266	0.136	1.30	**0.050**
Below poverty index	0.214	0.115	1.24	0.062	0.090	0.104	1.09	0.387	-0.064	0.128	0.94	0.619	0.111	0.133	1.12	0.403
Medicaid beneficiary	0.059	0.106	1.06	0.581	0.056	0.095	1.06	0.554	0.095	0.114	1.10	0.404	0.353	0.120	1.42	**0.003**
CHAMPUSVA beneficiary	0.207	0.227	1.23	0.362	-0.527	0.272	0.59	0.053	0.370	0.241	1.45	0.125	0.899	0.221	2.46	**0.000**
Lives in mobile home	-0.058	0.181	0.94	0.747	0.269	0.155	1.31	0.083	0.221	0.181	1.25	0.222	0.268	0.196	1.31	0.172
Not homeowner	0.119	0.083	1.13	0.151	0.118	0.075	1.13	0.114	0.304	0.089	1.35	**0.001**	0.246	0.097	1.28	**0.011**
Living in nursing home	0.894	0.162	2.44	**0.000**	0.521	0.165	1.68	**0.002**	1.239	0.156	3.45	**0.000**	0.796	0.179	2.22	**0.000**
Overweight	-0.248	0.078	0.78	**0.001**	-0.073	0.073	0.93	0.313	-0.420	0.084	0.66	**0.000**	-0.732	0.088	0.48	**0.000**
Cognitive impairment	0.232	0.109	1.26	**0.034**	0.282	0.100	1.33	**0.005**	0.549	0.118	1.73	**0.000**	0.412	0.129	1.51	**0.001**
Cancer survivor	0.100	0.080	1.11	0.214	-0.142	0.077	0.87	0.065	0.257	0.087	1.29	**0.003**	0.325	0.093	1.38	**0.001**
Stroke survivor	0.062	0.089	1.06	0.482	0.030	0.083	1.03	0.716	0.222	0.094	1.25	**0.018**	0.129	0.103	1.14	0.211
Heart problems survivor	0.248	0.070	1.28	**0.000**	0.176	0.064	1.19	**0.006**	0.205	0.078	1.23	**0.008**	0.104	0.084	1.11	0.214
Clinical depression	0.665	0.091	1.94	**0.000**	0.573	0.085	1.77	**0.000**	0.704	0.098	2.02	**0.000**	0.713	0.105	2.04	**0.000**

Cells in bold with grey background denote statistically significant *p*-values (*p*-value < 0.05)
Cells with orange background denote higher likelihood of membership than near perfect adherence (aOR > 1)
Cells with green background denote higher likelihood of membership than near perfect adherence (aOR < 1)

Diabetes medications risk factors adjusted model

Only ten risk factors retained statistical significance as moderators of trajectory group membership in the diabetes medications model. Table 17 displays the results of the generalized logistic regression, including estimates, standard errors, adjusted Odds Ratios, and *p*-values of all risk factors included in the adjusted diabetes medications trajectory model

Akin to what was observed in the two previous models, being female and overweight appear to reduce the likelihood for belonging to *moderate decline* and *rapid decline* trajectories, respectively (being female: $aOR_{moderate\ decline}$= 0.62, *p*-value=0.001; overweight: $aOR_{rapid\ decline}$= 0.69, *p*-value=0.041). However, in the diabetes medications model, the reduced likelihood of belonging to the *slow decline* group no longer remained statistically significant ($aOR_{slow\ decline}$= 0.90, *p*-value=0.376), contrary to what was observed in the previous models. Notably, females were more likely to belong to the *higher then increasing* adherence compared to the *near-perfect adherence* ($aOR_{higher\ then\ increasing\ decline}$= 0.90, *p*-value=0.376). Despite not being considered adherent (using the 80% threshold cutoff), the *higher then increasing* trajectory registers a significant increase in PDC over time converging with the PDC of the *near-perfect adherence* (Figure 16).

Like in the hypertension, Hispanic individuals are 47% less likely to belong to the *rapid decline* adherence trajectory compared to the *near-perfect adherence* ($aOR_{rapid\ decline}$= 0.53, *p*-value=0.011). However, Hispanics are also 1.43 times more likely to belong to the lower than increasing adherence trajectory ($aOR_{lower\ then\ increasing}$= 1.43, *p*-value=0.046) than to the *near-perfect adherence*. Non-whites follow this trend, despite the lack of statistical significance in the hypertension model ($aOR_{rapid\ decline}$= 0.69, *p*-value=0.043, $aOR_{lower\ then\ increasing}$= 1.65, *p*-value<0.001).

College education was shown to influence only the *moderate decline* trajectory, as non-college educated individuals were 1.79 times more likely to follow this trajectory than the *near-perfect adherence* ($aOR_{moderate\ decline}$= 1.79, *p*-value=0.013).

Homeownership and clinical depression were the only two risk factors that were associated with differences in membership in four non-adherent trajectories. Not being a homeowner or having clinical depression increased likelihood of belonging to either non-adherent trajectories when compared to *near-perfect adherence* trajectory. *Lower then increasing* and *higher then increasing adherence* trajectories were the exceptions for homeownership and clinical depression respectively. (not being a homeowner: $aOR_{slow\ decline}$= 1.34, *p*-value=0.019, $aOR_{\text{higher then increasing decline}}$= 1.33, *p*-value=0.016, $aOR_{moderate\ decline}$= 1.68, *p*-value=0.001, $aOR_{rapid\ decline}$= 1.52, *p*-value=0.009; clinical depression: $aOR_{slow\ decline}$= 1.62, *p*-value=0.001, $aOR_{\text{lower then increasing decline}}$= 1.63, *p*-value=0.001, $aOR_{moderate\ decline}$= 1.84, *p*-value<0.001, $aOR_{rapid\ decline}$= 1.62, *p*-value=0.006).

Living in a nursing home was the strongest risk factor for individuals to follow a declining adherence trajectory. Individuals who live in a nursing home were consistently 2 times more likely to follow a *slow, moderate, or declining adherence* than a *near-perfect adherence* trajectory ($aOR_{slow\ decline}$= 2.17, *p*-value=0.003, $aOR_{moderate\ decline}$= 3.73, *p*-value<0.001, $aOR_{rapid\ decline}$= 2.11, *p*-value=0.012).

Contrary to what was observed in the two previous models, the results of the diabetes medications model show that cancer and stroke survivorship only increased the likelihood of belonging to the *rapid decline* trajectory (cancer survivor: $aOR_{rapid\ decline}$= 1.83, *p*-value<0.001; stroke survivor: $aOR_{rapid\ decline}$= 1.45, *p*-value=0.033).

Table 17 - Time-fixed predictors in the diabetes medications trajectory model

	Slow decline				Higher then increasing adherence				Lower then increasing adherence				Moderate decline				Rapid decline			
	Estimate	S.E.	aOR	*p*-value	Estimate	S.E.	aOR	*p*-value	Estimate	S.E.	aOR	*p*-value	Estimate	S.E.	aOR	*p*-value	Estimate	S.E.	aOR	*p*-value
Being female	-0.104	0.117	0.90	0.376	0.262	0.116	1.30	**0.024**	0.299	0.132	1.35	**0.024**	-0.478	0.146	0.62	**0.001**	0.118	0.158	1.13	0.457
Foreign born	0.079	0.171	1.08	0.643	0.073	0.163	1.08	0.653	-0.165	0.185	0.85	0.373	-0.001	0.227	1.00	0.997	0.051	0.243	1.05	0.833
Non-white	0.031	0.130	1.03	0.813	0.109	0.126	1.11	0.390	0.500	0.137	1.65	**0.000**	-0.173	0.164	0.84	0.290	-0.371	0.183	0.69	**0.043**
Hispanic	-0.101	0.170	0.90	0.550	0.122	0.162	1.13	0.450	0.354	0.177	1.43	**0.046**	-0.407	0.223	0.67	0.067	-0.638	0.251	0.53	**0.011**
Not married	-0.020	0.123	0.98	0.872	-0.126	0.121	0.88	0.297	-0.156	0.136	0.86	0.251	0.044	0.156	1.04	0.779	-0.167	0.166	0.85	0.316
Not college educated	0.199	0.186	1.22	0.285	-0.045	0.181	0.96	0.803	-0.197	0.207	0.82	0.343	0.584	0.234	1.79	**0.013**	0.120	0.257	1.13	0.640
Below poverty index	-0.153	0.165	0.86	0.354	-0.182	0.165	0.83	0.271	-0.244	0.179	0.78	0.172	0.104	0.196	1.11	0.597	0.157	0.215	1.17	0.464
Medicaid beneficiary	0.014	0.153	1.01	0.926	0.221	0.150	1.25	0.141	0.189	0.164	1.21	0.249	0.192	0.180	1.21	0.285	0.170	0.198	1.19	0.391
Not homeowner	0.289	0.123	1.34	**0.019**	0.288	0.120	1.33	**0.016**	0.222	0.136	1.25	0.101	0.518	0.149	1.68	**0.001**	0.420	0.162	1.52	**0.009**
Living in nursing home	0.773	0.257	2.17	**0.003**	-0.491	0.344	0.61	0.153	0.539	0.300	1.71	0.073	1.315	0.265	3.73	**0.000**	0.749	0.299	2.11	**0.012**
Overweight	-0.051	0.145	0.95	0.725	0.132	0.150	1.14	0.379	-0.225	0.156	0.80	0.150	-0.278	0.170	0.76	0.102	-0.367	0.180	0.69	**0.041**
Cognitive impairment	0.085	0.160	1.09	0.596	-0.032	0.158	0.97	0.838	-0.302	0.176	0.74	0.086	0.218	0.194	1.24	0.262	0.244	0.213	1.28	0.252
Cancer survivor	0.006	0.130	1.01	0.966	0.051	0.126	1.05	0.687	-0.042	0.147	0.96	0.777	0.294	0.153	1.34	0.055	0.607	0.157	1.83	**0.000**
Stroke survivor	0.113	0.141	1.12	0.424	-0.031	0.143	0.97	0.825	0.078	0.158	1.08	0.619	0.181	0.167	1.20	0.277	0.373	0.175	1.45	**0.033**
Heart problems survivor	0.172	0.110	1.19	0.120	0.014	0.108	1.01	0.895	0.109	0.122	1.11	0.374	0.159	0.137	1.17	0.248	0.246	0.147	1.28	0.094
Clinical depression	0.480	0.138	1.62	**0.001**	0.137	0.140	1.15	0.328	0.490	0.148	1.63	**0.001**	0.607	0.163	1.84	**0.000**	0.484	0.177	1.62	**0.006**

Cells in bold with grey background denote statistically significant *p*-values (*p*-value < 0.05)
Cells with orange background denote higher likelihood of membership than near perfect adherence (aOR > 1)
Cells with green background denote higher likelihood of membership than near perfect adherence (aOR < 1)

Cumulative Risk Effects in Probability Trajectory Membership

By including the risk factors that were shown to have statistical significance for each trajectory model, several scenarios were created to impute predicted trajectory membership probabilities. These were considered as the cumulative risk factor effect in trajectory probability membership. The scenarios for the hypertension medications, statins, and diabetes medications are available in Appendix 2.

Even in the total absence of risk factors, the highest proportion possible of near perfect adherers is never greater than 53.3%. Specifically, assuming the absence of risk factors in a sample population meaningfully altered the probability of individuals following a *near-perfect adherence*: 53.3%, 47.21%, and 36.12% for the hypertension medications, statins, and diabetes medications models respectively (Figures 17, 18, and 19). Interestingly, the proportion of near-perfect adherers decreases with the number of trajectories identified in the model. Moreover, the effect of all accumulated risk factors changes the proportions of the trajectory groups in a dramatic fashion: the *rapid decline* adherence trajectory was typically the smallest group in the original estimated model (base case). In the hypertension medications and statins trajectory models, the accumulation of risk factors transforms the *rapid decline* into the largest group, while the *near-perfect adherence* group became the smallest group. In the diabetes model, the risk factor accumulation effect transformed the *near-perfect adherence* group into the smallest in the model, the most prominent groups were the *slow decline* and *moderate decline* trajectories.

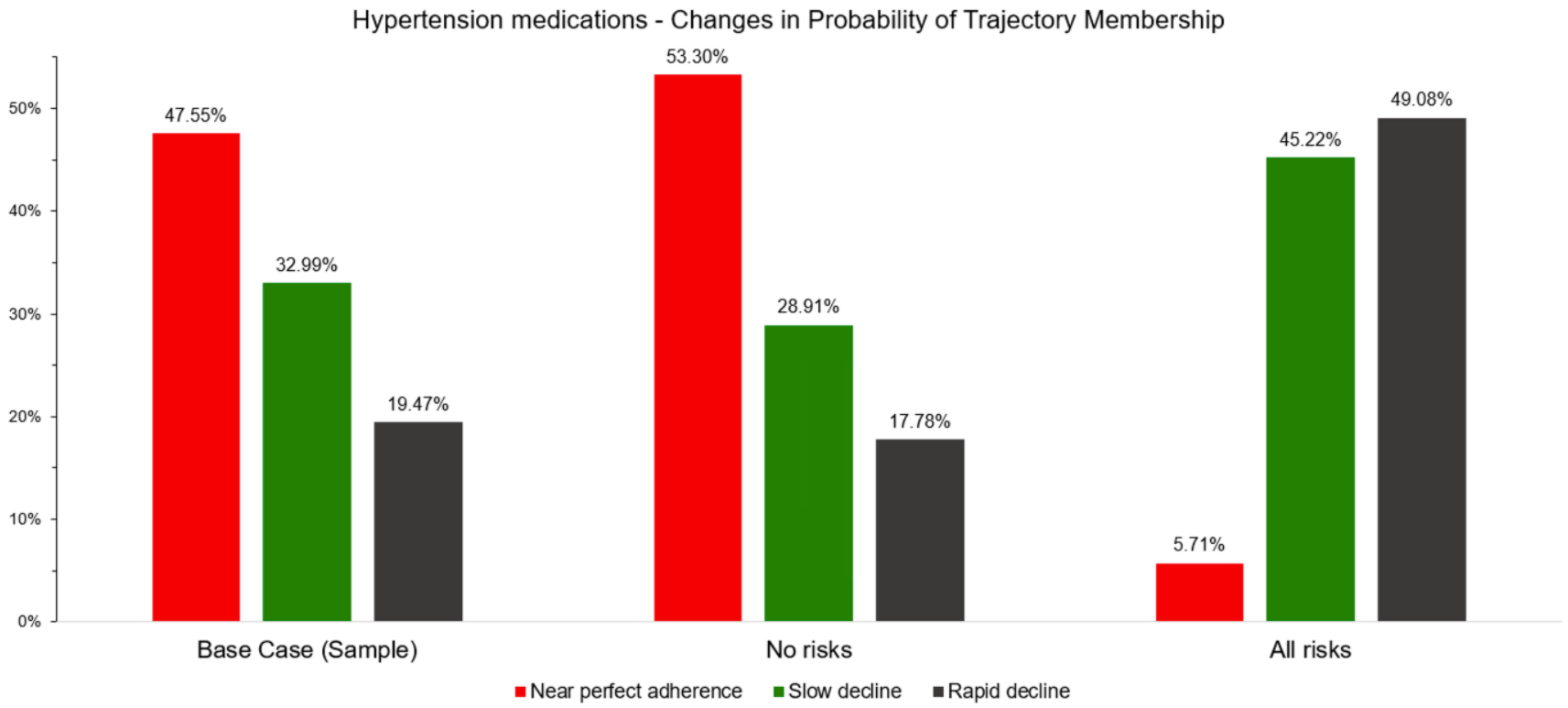

Figure 17 - Hypertension medications: Changes in probability membership due to risk factors

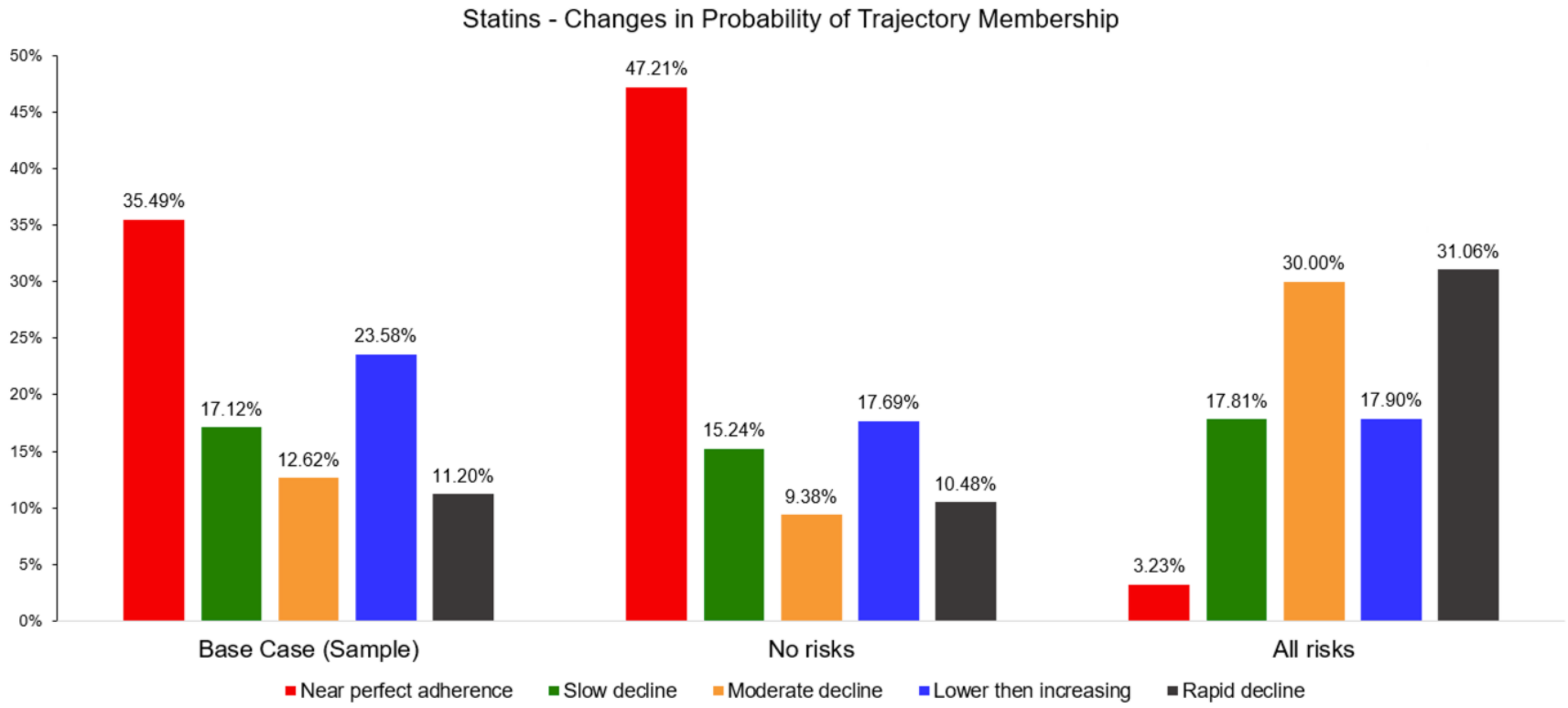

Figure 18 - Statins: Changes in probability membership due to risk factors

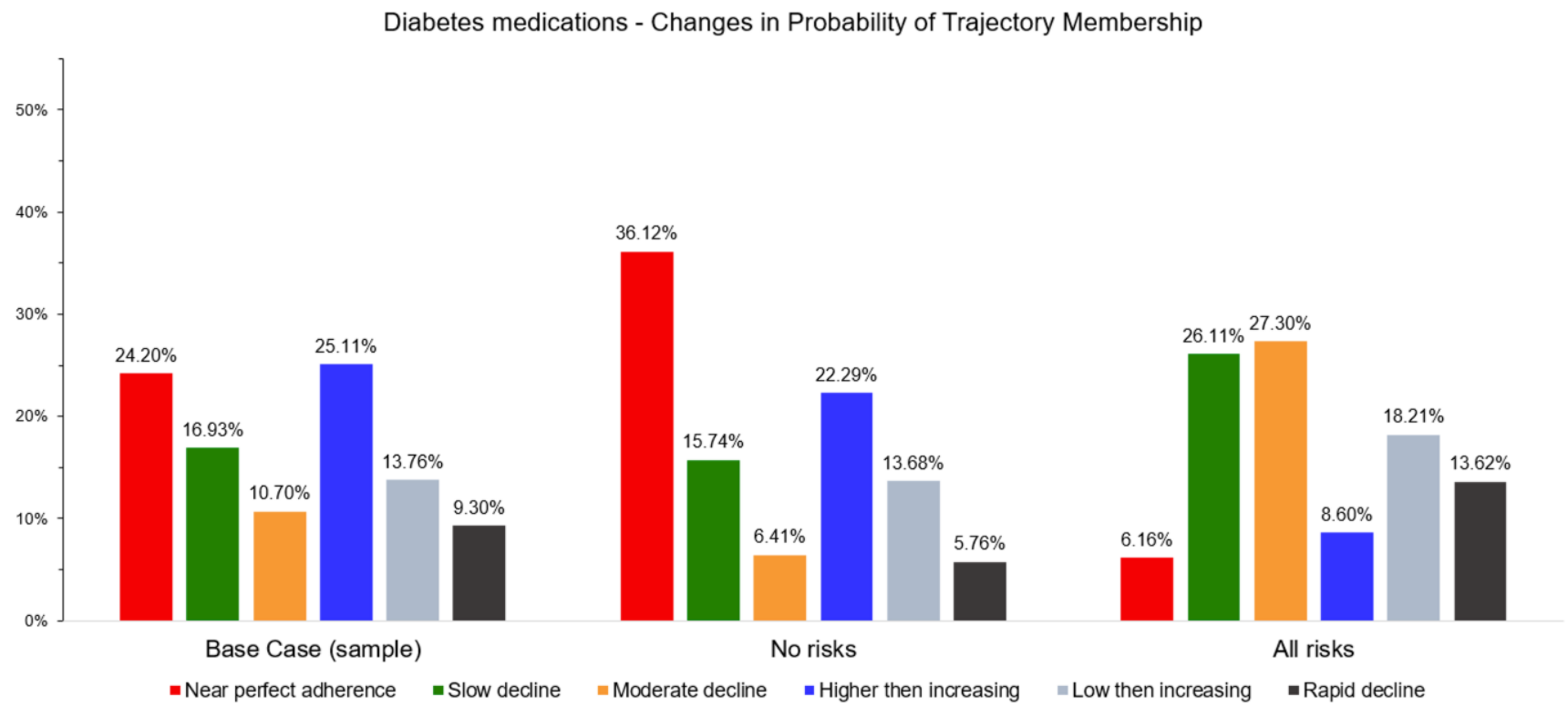

Figure 19 - Diabetes medications: Changes in probability membership due to risk factors

4.3.3 TIME-VARYING MEDICATION ADHERENCE TRAJECTORY RISK FACTORS

As described in previous sections, it is possible that medication adherence behavior is influenced by patient's characteristics that remain can either remain stable or change with time. A multi group-based trajectory analysis was implemented to investigate if and to what extent each of the time-varying *Enabling* and *Need* characteristics are associated with changes in medication adherence trajectories (as hypothesized in the Conceptual Framework, section 2.7). Table 18 lists the covariates included in the conceptual framework observed to change with time. While the traj function for STATA 17 MP allows estimating multi-trajectory models with the time-stable risk factors, that was not accomplished in this research work because of lack of computational power. Each multi trajectory model fitting the several time-varying risk factors took at least 10 hours to run to completion.

Table 18 - Time-varying characteristics included in the multi group-based trajectory models of medication adherence

Enabling	Need
Poverty threshold	Self-reported health status
Family structure: a) Number of resident children b) Number of living children	Cancer survivorship
Loss of spouse	Activities of Daily Living
Medicaid beneficiary status	Instrumental Activities of Daily Living
Additional health insurance coverage	Smoking status
	Number of drinking days/week
	Clinical depression
	Life satisfaction
	Retirement satisfaction
	Limitations in work due to health

The multi group-based trajectory models showed how time-varying predictors change along with each medication adherence trajectories. It is important to clarify that variations in covariates do not imply a causal relationship but rather a longitudinal description of how each adherence trajectories and covariate trajectories progressed with time.

Enabling characteristics

The enabling characteristics with the most significant changes throughout the medication adherence measurement period were living below the poverty threshold, loss of spouse, and additional health insurance. Near perfect adherers of hypertensive medications and statins typically saw a reduction of the probability of living below the poverty threshold. Conversely, *slow decline*rs of all models saw either a maintenance or an increase in the probability of living below the poverty threshold. *Moderate decline* trajectory in the statins model was the adherence trajectory with significant changes in the poverty threshold characteristic, in which patients were increasingly likely to live below the poverty threshold. In general, loss of spouse increased in every model, reflecting likely the natural likelihood that as individuals age, the more likely they are to die. However, in some adherence trajectories, individuals seemed to evolve from a situation in which the spouse (had) died to one with a spouse. This was observed in the *near-perfect adherence* and *slow decline* trajectories in the hypertension and diabetes medications models. Resident children in the household did not seem to register any significant changes with time: on average, the count of resident children was less than 1, indicating how rare it was for individuals to still have a child living in the household. In the hypertension and diabetes models, being a Medicaid beneficiary did not change to a great extent, which was possibly related with the lack of change in the poverty threshold trajectories (Figure 20). However, in the statins model, when individuals were no longer living below the poverty threshold, they were not also

Medicaid beneficiaries. In general, it appeared that improvements in the socioeconomic status of individuals taking statins was consistent with maintaining high adherence or increases in medication adherence (Figure 21). Moreover, the use of additional health insurance seemed to decrease with time, which bodes with the assumption that at age 65, individual might still be working, or have enough income to supplement Medicare benefits. Notably, individuals in the rapid or *moderate decline* trajectories registered consistently lower probability of using additional health insurance (Figure 20, 21, 22). Only exception was in the perfect adherer's trajectory in the statins model, in which the probability of having additional insurance increased dramatically with time.

Need characteristics

While self-perceived health status did not register meaningful changes throughout the period of analysis, clinical depression (dichotomized variable in 0/1, Table 9) displayed important trends. Declining trajectories in virtually all models observed increases in the probability of having clinical depression (Figures 23, 25, 27). No obvious trends were observed with life satisfaction. The more significant changes with time were observed in perfect adherers of hypertension medications and statins and in slow and *moderate decline*rs of statins. In all models, limitations in work due to health seem to increase, particularly in declining trajectories. Additionally, the probability of suffering from limitations in work due to health was, in most cases, the lowest in the *near-perfect adherence* trajectory. Despite limitations in work increasing for perfect adherers, the rate of increase seemed to be larger for all other non-perfect trajectories. Retirement satisfaction did not display any meaningful trends in association with medication adherence trajectories.

Smoking status as a covariate appeared to be meaningful only for the perfect adherers of hypertension medication (Figure 24). For all other models, smoking status was generally low, frequently well below 11% of probability of smoking (Figures 26 and 28). High adherence then increasing adherers of diabetes medications were the only exception to this trend: smoking probability increased with time. The number of drinking days per week did not display any relevant trends, remaining constant for most trajectories in all three models.

The level of autonomy provided one of the most interesting findings. While all adherence trajectories were observed to follow increases in the level of dependency (both in IADL and ADL scores), the increases were less prominent for the perfect adherers (Figure 24, 26, 28). *Rapid decline*rs of all models registered some decrease in ADL. The HRS survey coded responses from respondents who did not require any assistance the IADL and ADL tasks as zero, even for those activities respondents stated they do not perform.

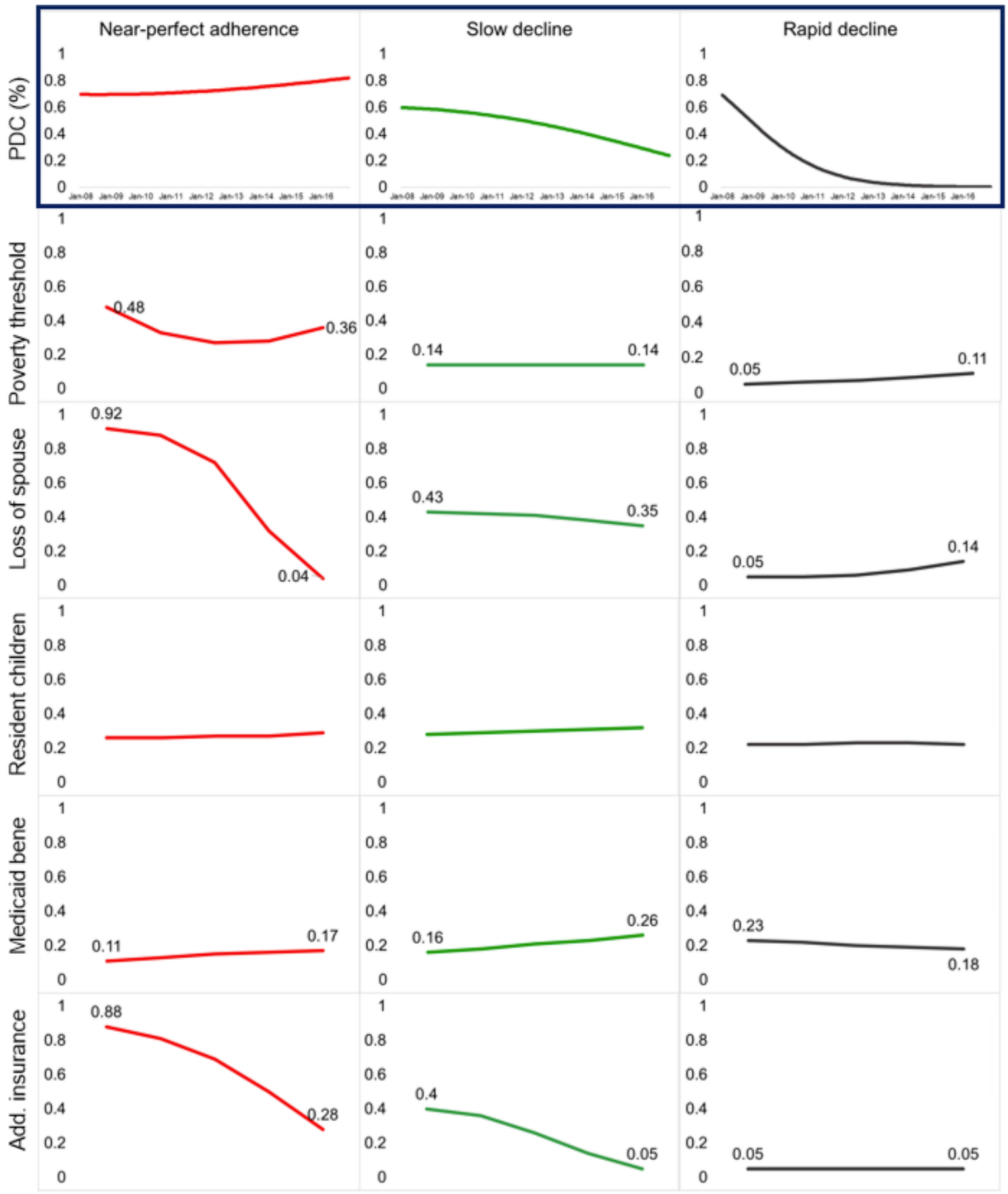

Figure 20 - Multi-trajectories of Enabling Characteristics - hypertension medications model

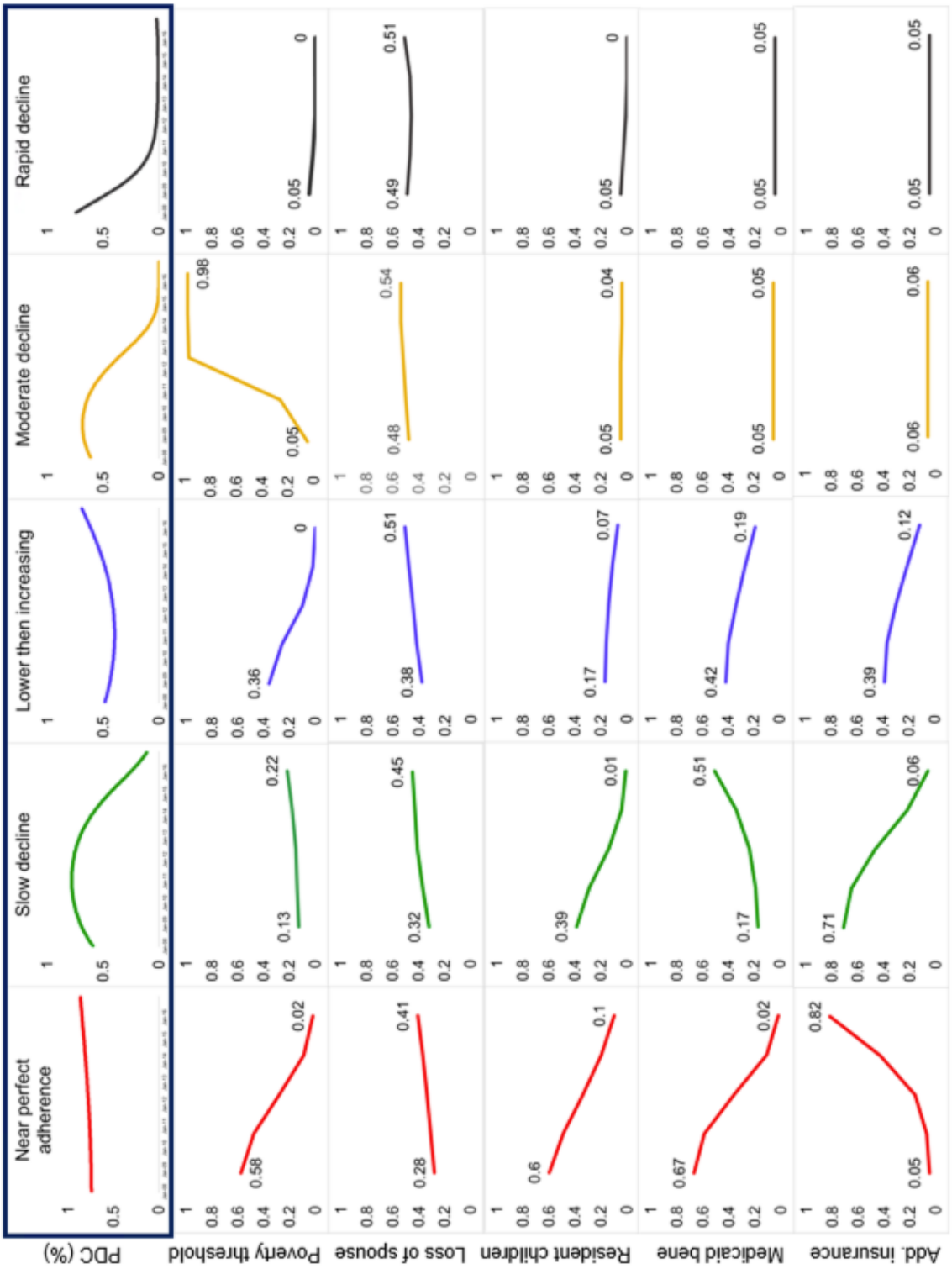

Figure 21 - Multi-trajectories of Enabling Characteristics – Statins model

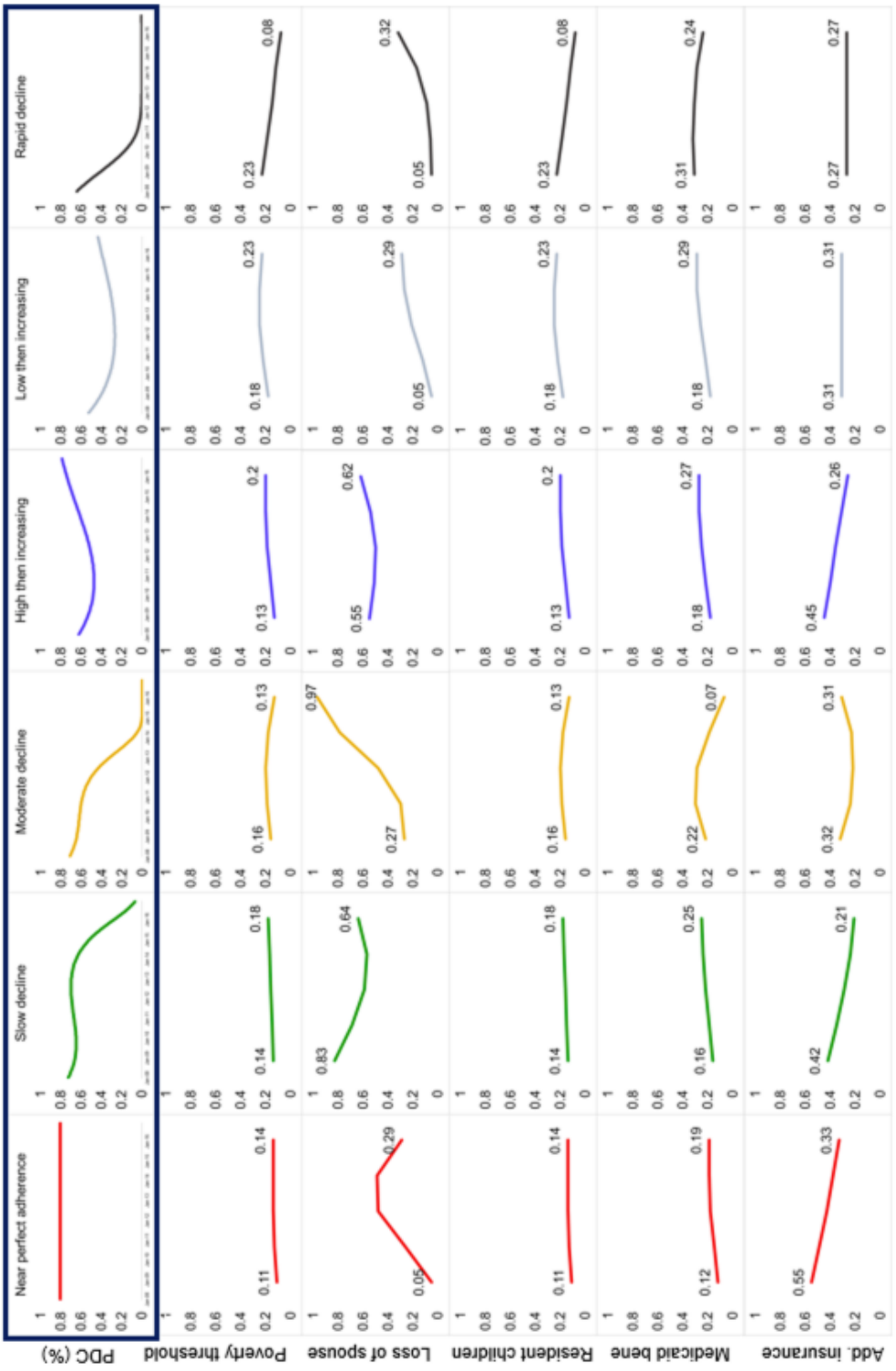

Figure 22 - Multi-trajectories of Enabling Characteristics – Diabetes medication model

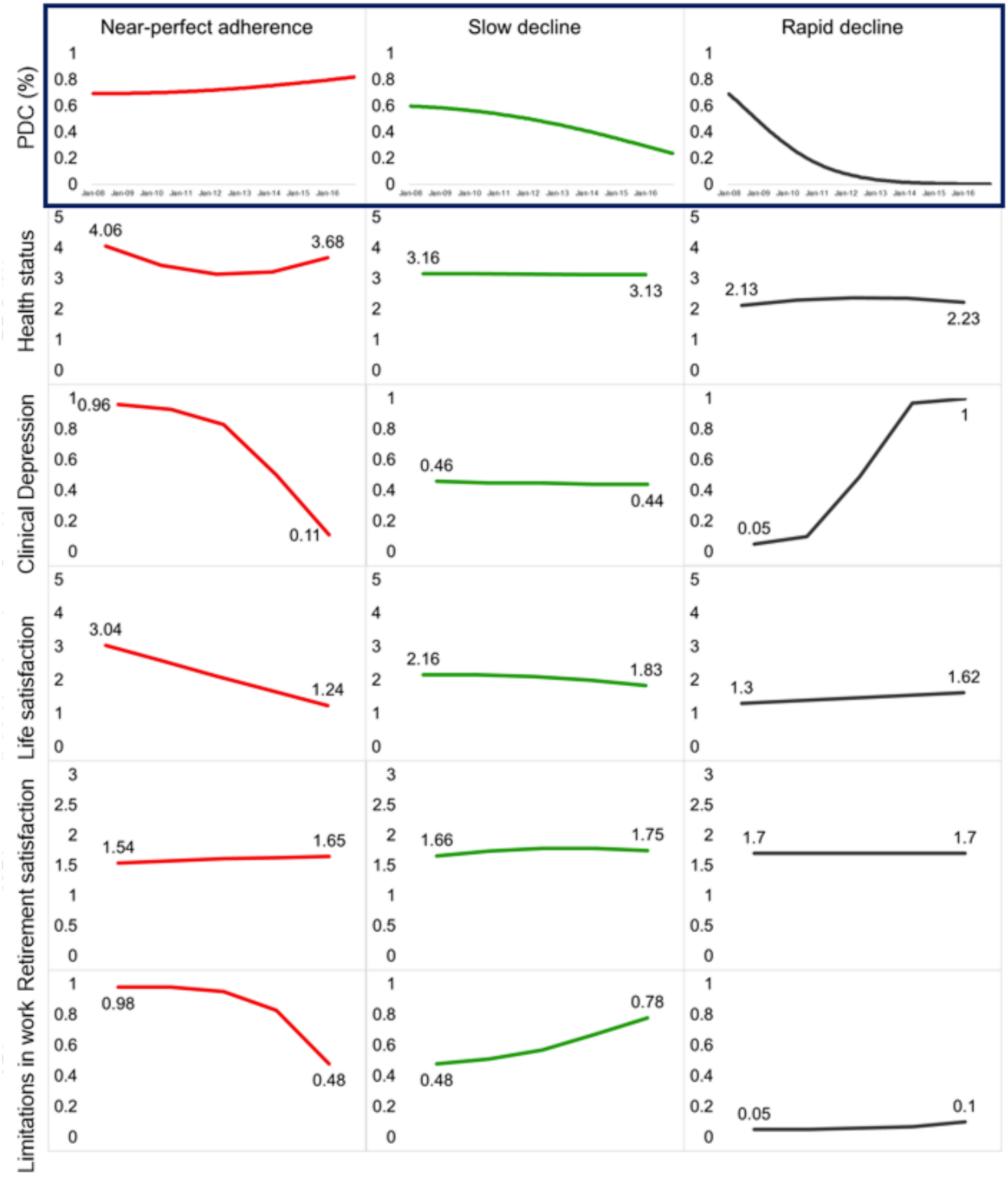

Figure 23 - Multi-trajectories of Need Characteristics - hypertension medications model

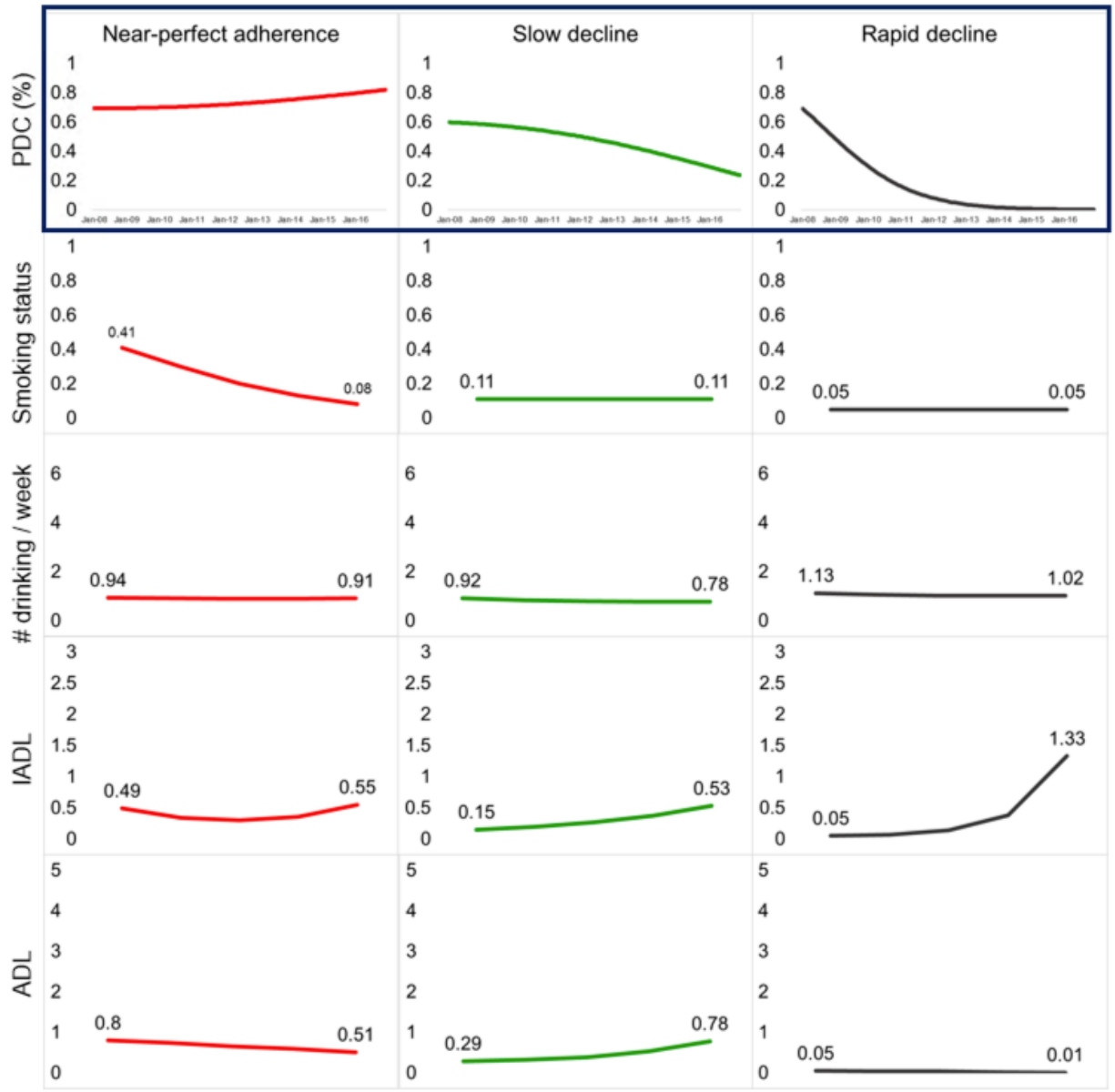

Figure 24 - Multi-trajectories of Need Characteristics - hypertension medications model (continued)

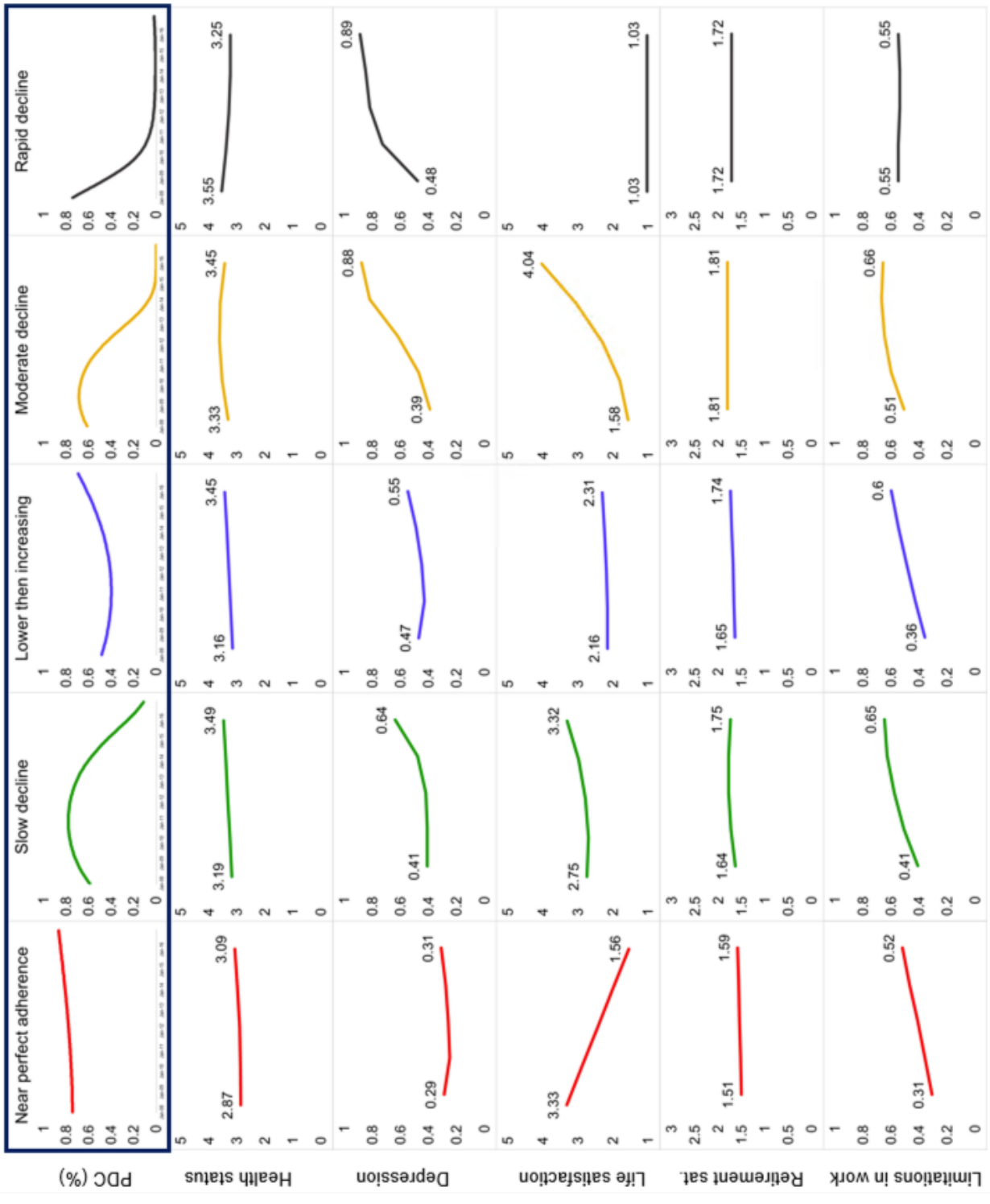

Figure 25 - Multi-trajectories of Need Characteristics - Statins medications model

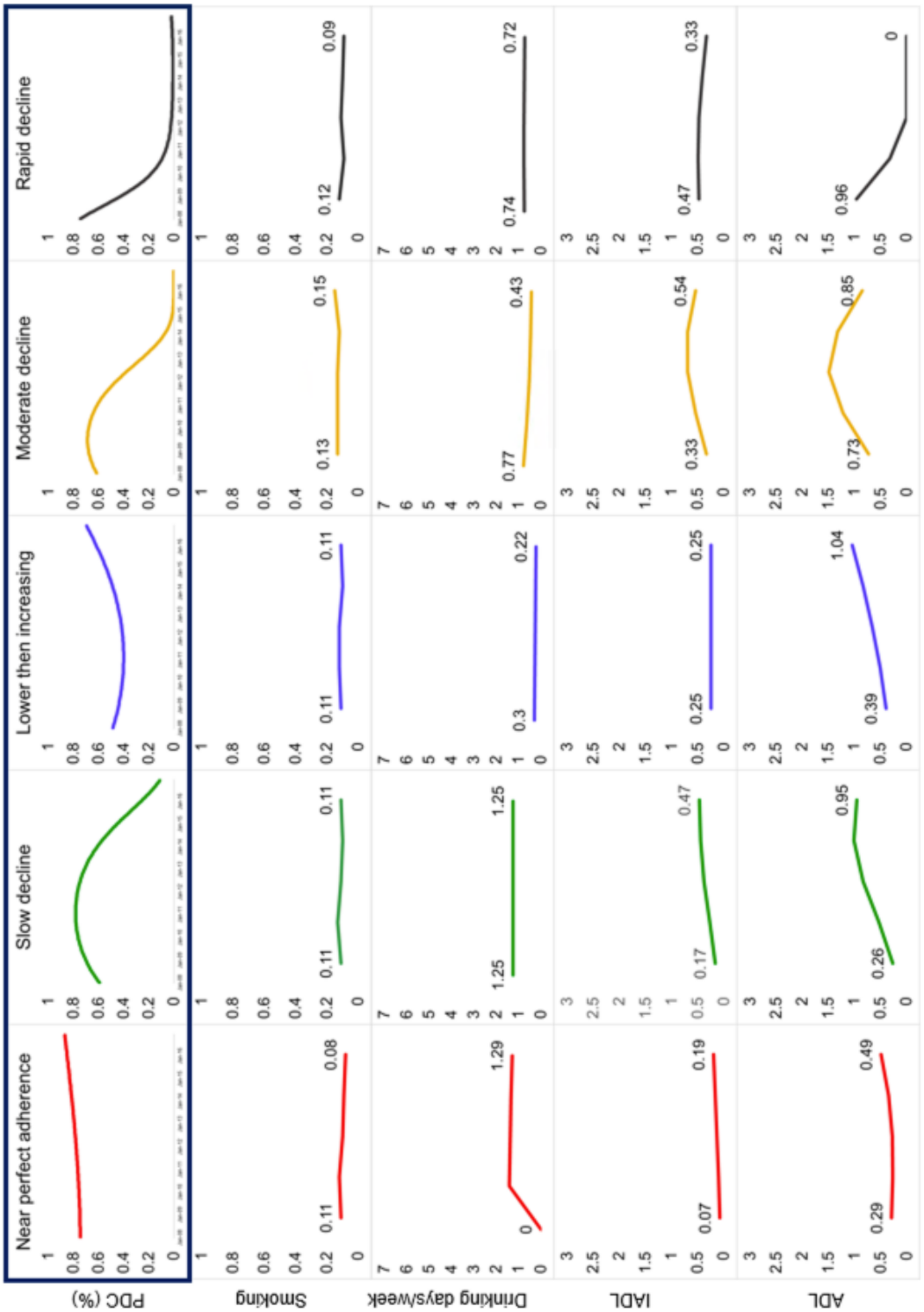

Figure 26 - Multi-trajectories of Need Characteristics - Statins medications model (continued)

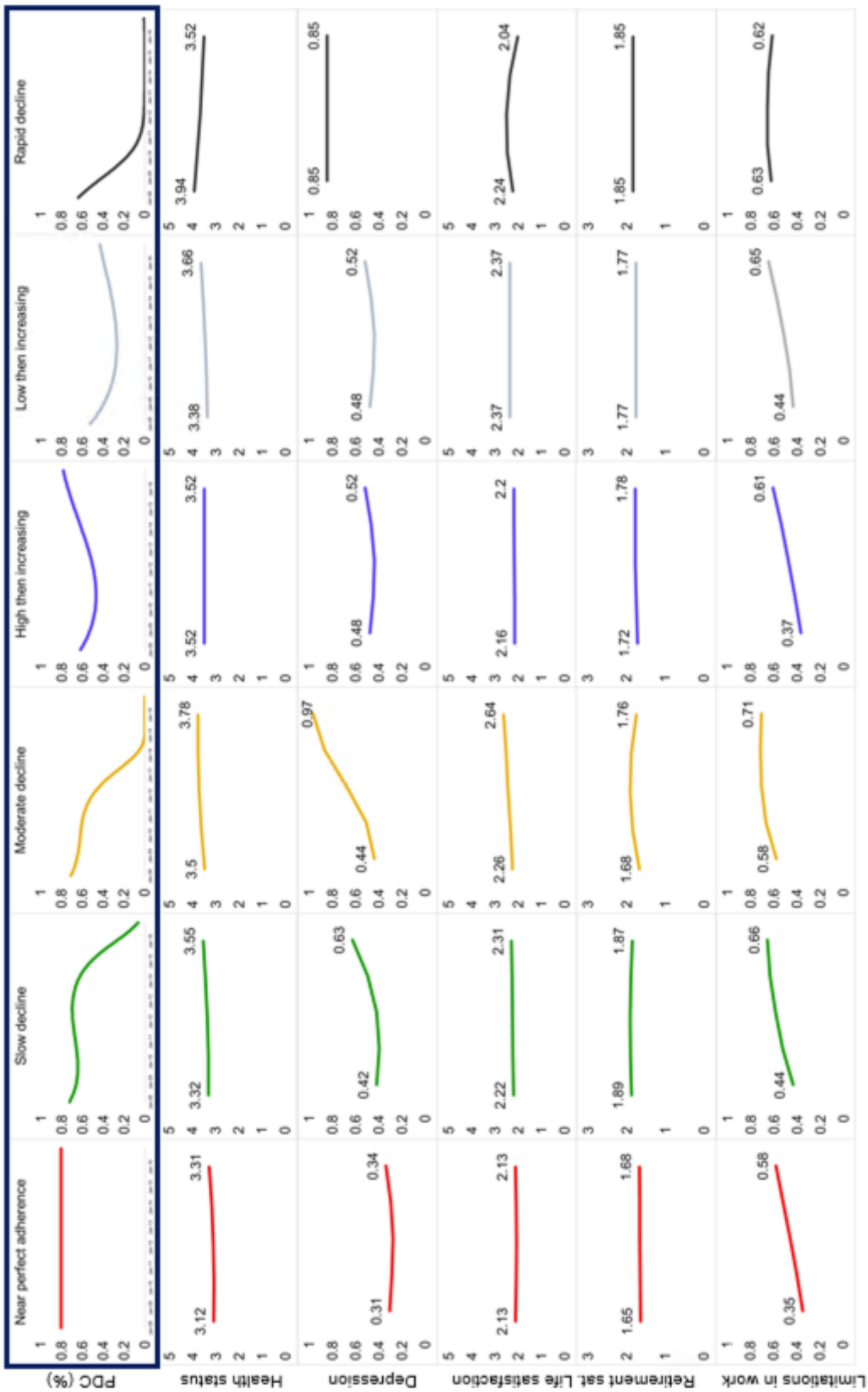

Figure 27 - Multi-trajectories of Need Characteristics – Diabetes medication model

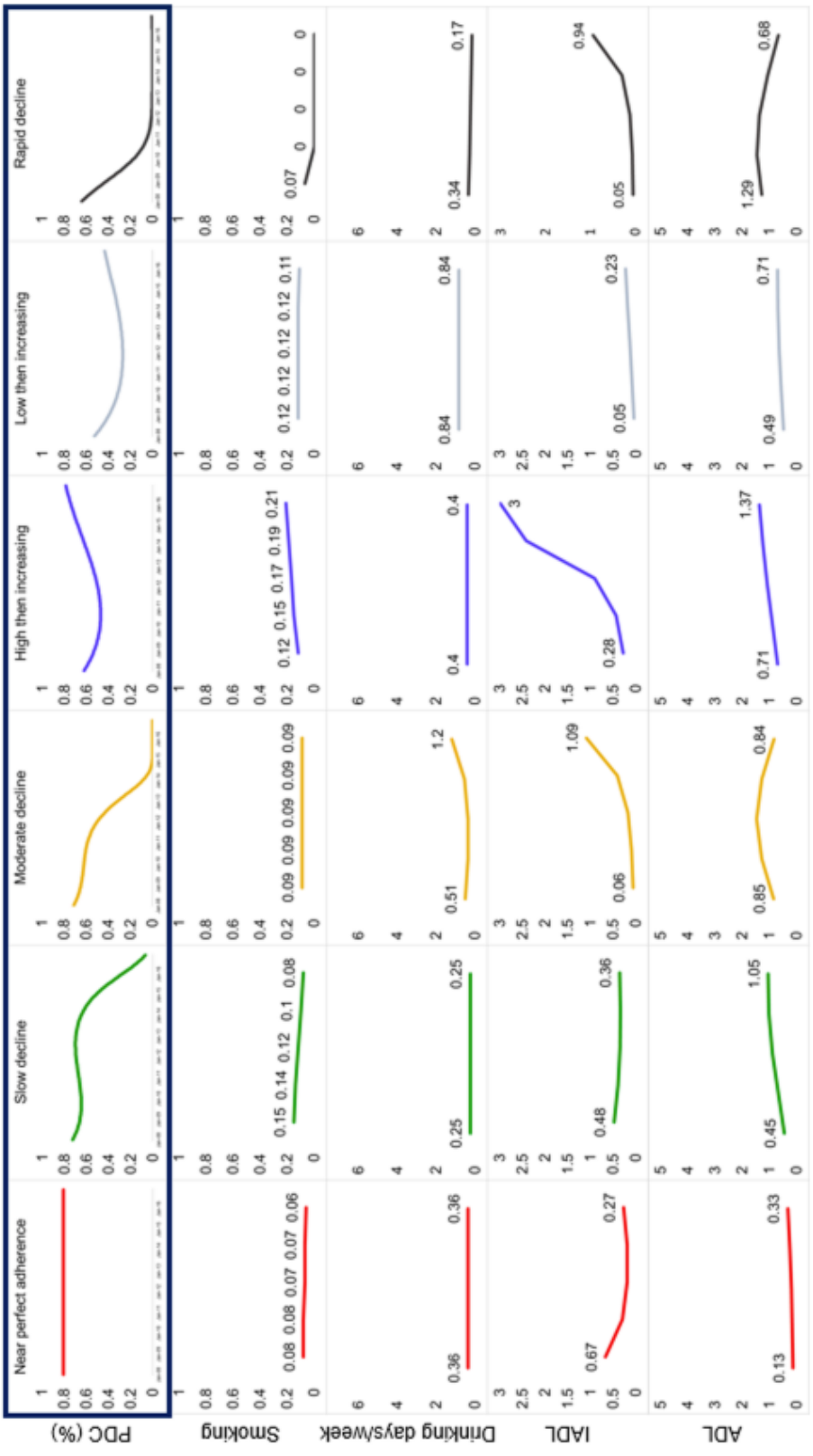

Figure 28 - Multi-trajectories of Need Characteristics – Diabetes medication model (continued)

4.4 SPECIFIC AIM 3: PREDICTIVE MODEL LINKING MEDICATION ADHERENCE TRAJECTORIES TO HEALTH OUTCOMES

4.4.1 MEDICATION ADHERENCE TRAJECTORIES AND OUTCOMES

In order to investigate the association between medication trajectory membership and outcomes, a logistic regression was computed. Three individual unadjusted logistic regression models for hypertension medications, statins, and diabetes medications were estimated using the trajectory membership variables as independent variables and outcome variables as dependent variables. The outcomes considered model were specific to the type of medication (Table 19).

Table 19 - Outcomes considered by medication adherence trajectory models

Trajectory model	**Outcomes**
Hypertension medications	Myocardial infarction Stroke
Statins	Myocardial infarction Stroke
Diabetes medications	Myocardial infarction Stroke *Diabetes-specific outcomes* Ophthalmic complications Nephropathy Neuropathy Diabetic peripheral angiopathy

The *near-perfect adherence* trajectories found in every trajectory model were used as the reference category for estimation purposes. The likelihood ratio suggested that all three unadjusted models were statistically significant. Table 20 encompasses the unadjusted models' coefficients, standard errors, 95% confidence intervals for MI and stroke, while Table 21 displays the same statistics for the outcomes specific to diabetes medications.

Table 20 – Myocardial infarction and stroke predictive models: unadjusted model logistic regressions

Model	Coefficient	S.E.	95% C.I.	Likelihood ratio $\chi 2$	*p*-value
Outcome: Myocardial infarction					
Trajectory Model: Hypertension medications				**55.71**	**< 0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Rapid decline	0.648	0.108	0.473 – 0.896		< 0.001
Slow decline	0.595	0.096	0.407 – 0.783		< 0.001
Trajectory Model: Statins				**74.95**	**< 0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Rapid decline	0.641	0.117	0.411 – 0.871		< 0.001
Moderate decline	0.720	0.111	0.503 – 0.936		< 0.001
Slow decline	0.734	0.101	0.536 – 0.932		< 0.001
Low then increase	0.393	0.100	0.197 – 0.590		< 0.001
Trajectory Model: Diabetes medications				**59.93**	**<0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Low then increase	0.342	0.225	-0.099 – 0.783		0.128
Rapid decline	0.978	0.216	0.555 – 1.401		<0.001
Moderate decline	1.303	0.196	0.919 – 1.687		<0.001
Higher *low then increase*	0.405	0.192	0.029 – 0.781		0.035
Slow decline	0.860	0.192	0.484 – 1.236		<0.001
Outcome: Stroke					
Trajectory Model: Hypertension medications				**94.73**	**< 0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Rapid decline	0.747	0.087	0.577 – 0.917		< 0.001
Slow decline	0.596	0.078	0.444 – 0.749		< 0.001

Model	Coefficient	S.E.	95% C.I.	Likelihood ratio χ2	*p*-value
Outcome: Stroke					
Trajectory Model: Statins				**72.47**	**< 0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Rapid decline	0.544	0.100	0.348 – 0.741		< 0.001
Moderate decline	0.728	0.092	0.548 - .909		< 0.001
Slow decline	0.416	0.090	0.240 – 0.592		< 0.001
Low then increase	0.304	0.084	0.139 – 0.470		< 0.001
Trajectory Model: Diabetes medications				**26.99**	**<0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Low then increase	0.446	0.160	0.132 – 0.759		0.005
Rapid decline	0.529	0.176	0.183 – 0.875		0.003
Moderate decline	0.521	0.169	0.190 – 0.851		0.002
High then increase	0.002	0.150	-0.292 – 0.296		0.988
Slow decline	0.495	0.150	0.201 – 0.789		0.001

Table 21 – Diabetes-specific predictive outcomes: unadjusted model logistic regressions

Model	Coefficient	S.E.	95% C.I.	Likelihood ratio χ2	*p*-value
Outcome: Ophthalmic complications					
Trajectory Model: Diabetes medications				**12.83**	**0.025**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Low then increase	0.208	0.265	-0.311 – 0.727		0.432
Rapid decline	-0.141	0.338	-0.803 – 0.521		0.676
Moderate decline	0.503	0.263	-0.012 – 1.017		0.055
High then increase	-0.374	0.260	-0.884 – 0.137		0.151
Slow decline	0.310	0.243	-0.167 – 0.787		0.203
Outcome: Nephropathy					

Model	Coefficient	S.E.	95% C.I.	Likelihood ratio $\chi 2$	*p*-value
Trajectory Model: Diabetes medications				**34.14**	**<0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Low then increase	0.222	0.232	-0.233 – 0.677		0.339
Rapid decline	0.734	0.228	0.287 – 1.182		0.001
Moderate decline	0.946	0.209	0.539 – 1.358		<0.001
High then increase	0.084	0.204	-0.317 – 0.484		0.682
Slow decline	0.669	0.199	0.279 – 1.058		0.001
Outcome: Neuropathy					
Trajectory Model: Diabetes medications				**20.46**	**0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Low then increase	0.432	0.131	0.174 – 0.689		0.001
Rapid decline	0.271	0.154	-0.031 – 0.572		0.078
Moderate decline	0.339	0.144	0.057 – 0.622		0.018
High then increase	-0.030	0.122	-0.269 – 0.209		0.806
Slow decline	0.283	0.128	0.033 – 0.533		0.026
Outcome: Diabetic peripheral angiopathy					
Trajectory Model: Diabetes medications				**49.06**	**<0.001**
Near-perfect adherence	*reference*	*reference*	*reference*	*reference*	*reference*
Low then increase	0.393	0.306	-0.208 - 0.994		0.200
Rapid decline	1.264	0.277	0.721 – 1.808		<0.001
Moderate decline	1.320	0.266	0.799 – 1.841		<0.001
High then increase	-0.040	0.291	-0.611 – 0.531		0.890
Slow decline	0.725	0.270	0.195 – 1.255		0.007

The unadjusted models show a positive association between trajectory membership and MI or stroke in both hypertension medications and statins. Contrastingly, in the diabetes medications models, a positive association between trajectory membership and outcomes is only observed for

the *slow*, *moderate*, and *rapid decline* trajectories for renal complications, *low then increase*, moderate and *slow decline* for neuropathy, and *rapid*, *moderate*, and *slow decline* for diabetic peripheral angiopathy (Table 21).

The adjusted models were estimated including sociodemographic characteristics and other variables that were associated with trajectory membership for each type of medication. These included:

- *Hypertension medications*: foreign born, race, Hispanic, marital status, college education, poverty threshold, Medicaid or CHAMPUS/VA beneficiary, type of home, homeownership, living in nursing home, excessive weight, cognitive impairment, survivor of cancer or heart problems.
- *Statins*: sex, foreign born, race, Hispanic, marital status, college education, poverty threshold, Medicaid or CHAMPUS/VA beneficiary, type of home, homeownership, living in nursing home, excessive weight, cognitive impairment, survivor of cancer or heart problems.
- *Diabetes medications*: sex, foreign born, race, Hispanic, marital status, college education, poverty threshold, Medicaid beneficiary, homeownership, living in nursing home, excessive weight, cognitive impairment, survivor of cancer or heart problems.

As seen on Table 22, the likelihood ratios indicated that the adjusted models were still statistically significant in predicting outcomes per trajectory membership. However, some of the associations between trajectory membership and outcomes that were found to be statistically significant in the unadjusted models no longer remained so in the adjusted models. These include the association between stroke and the *moderate* and *slow decline* and the *low then increase* trajectories of statins users, and the *low then increase*, *rapid*, and *moderate decline* trajectories of

diabetes medications. Only the association between *low then increase* trajectory and MI was not found as statistically significant. The adjusted model for neuropathy showed no statistically significant association with trajectory membership of diabetes medications users. The odds ratios, 95% confidence intervals and statistical significance for the outcomes by medications groups are shown in Figures 29, 30, and 31 for MI and stroke, and Figure 32 for diabetes-specific outcomes.

Table 22 – Likelihood ratio and *p*-values of the adjusted models predictive of outcomes

Model	Likelihood Ratio $\chi 2$	*p*-value
Outcome: Myocardial infarction		
Trajectory Model: Hypertension medications	331.29	<0.001
Trajectory Model: Statins	512.51	<0.001
Trajectory Model: Diabetes medications	275.28	<0.001
Outcome: Stroke		
Trajectory Model: Hypertension medications	922.79	<0.001
Trajectory Model: Statins	1218.61	<0.001
Trajectory Model: Diabetes medications	477.64	<0.001
Outcome: Ophthalmic complications		
Trajectory Model: Diabetes medications	91.91	<0.001
Outcome: Nephropathy		
Trajectory Model: Diabetes medications	139.99	<0.001
Outcome: Neuropathy		
Trajectory Model: Diabetes medications	155.08	<0.001
Outcome: Diabetic peripheral angiopathy		
Trajectory Model: Diabetes medications	90.69	<0.001

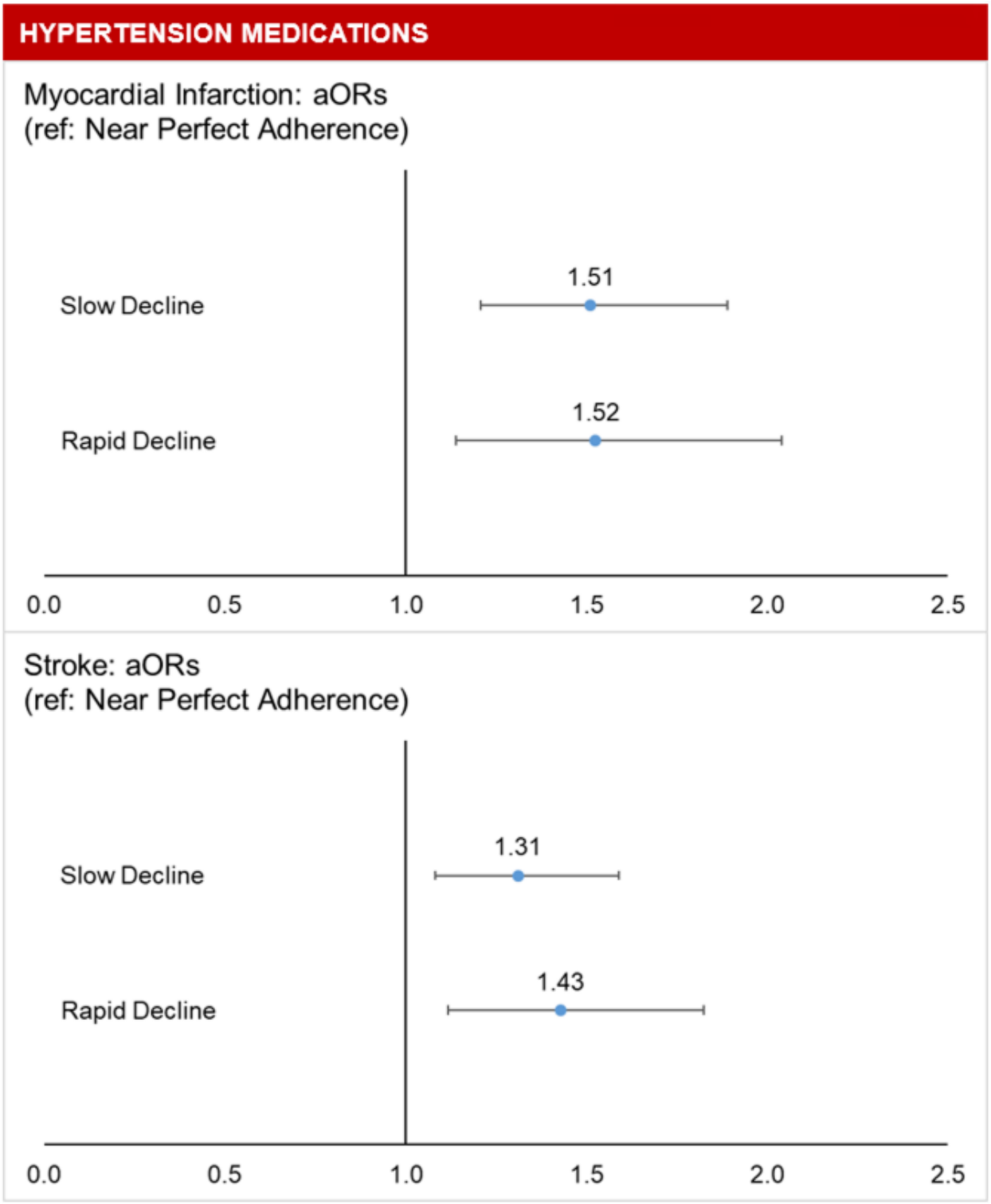

Figure 29 - Predictive aORs of medication adherence trajectories for hypertension medications and myocardial infarction and stroke

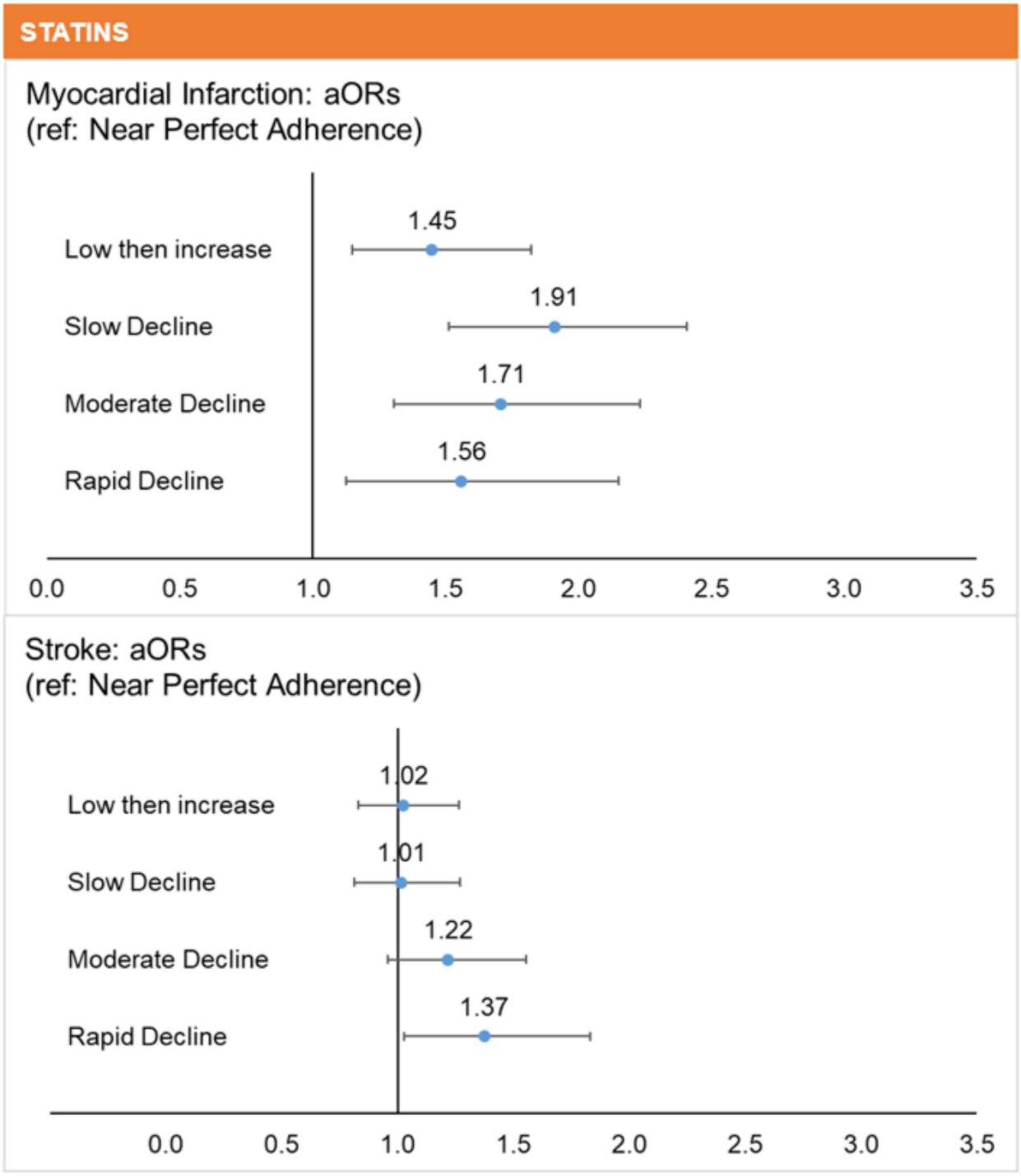

Figure 30 - Predictive aORs of medication adherence trajectories for statins and myocardial infarction and stroke

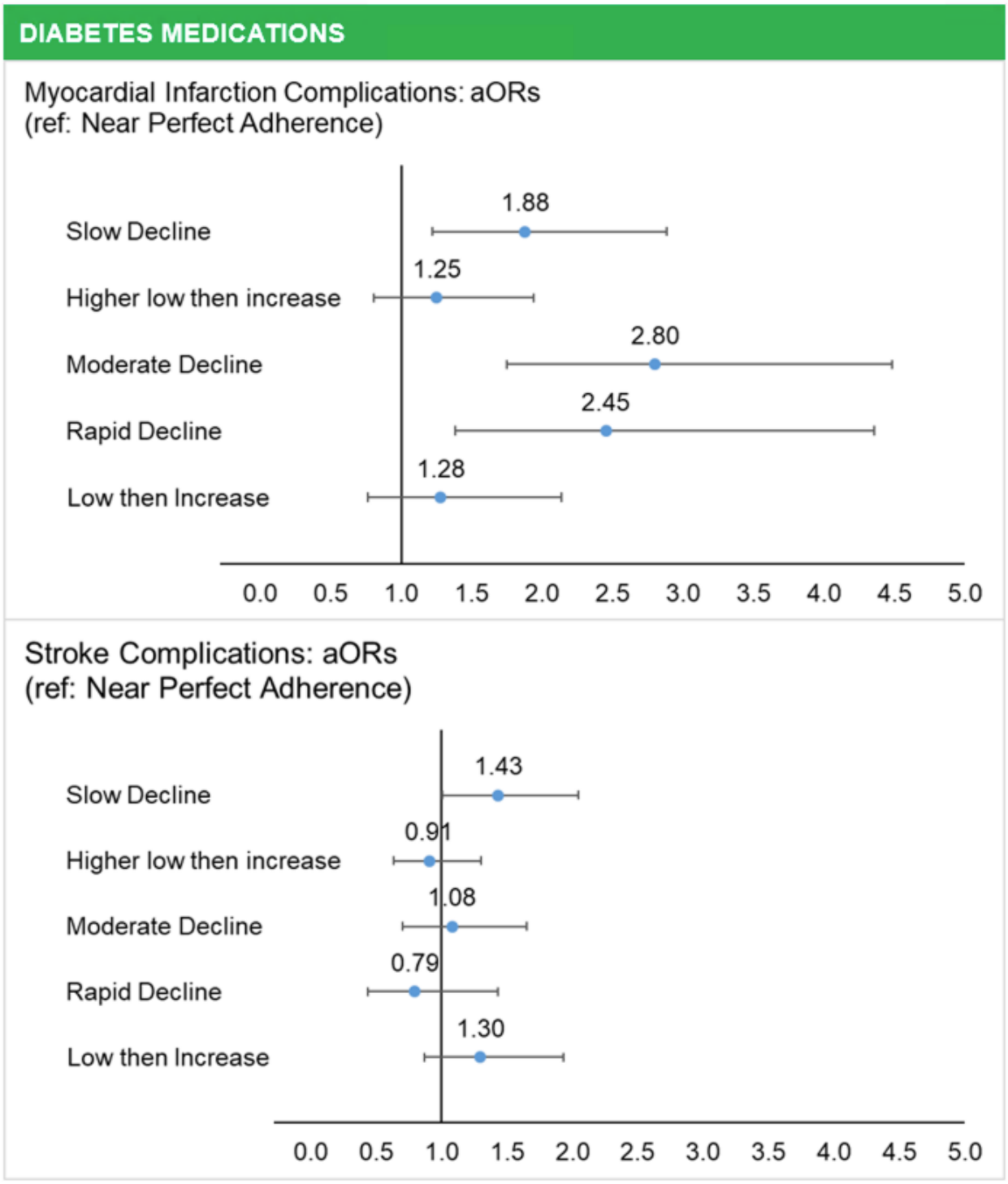

Figure 31 – Predictive aORs of medication adherence trajectories for diabetes medications and myocardial infarction and stroke

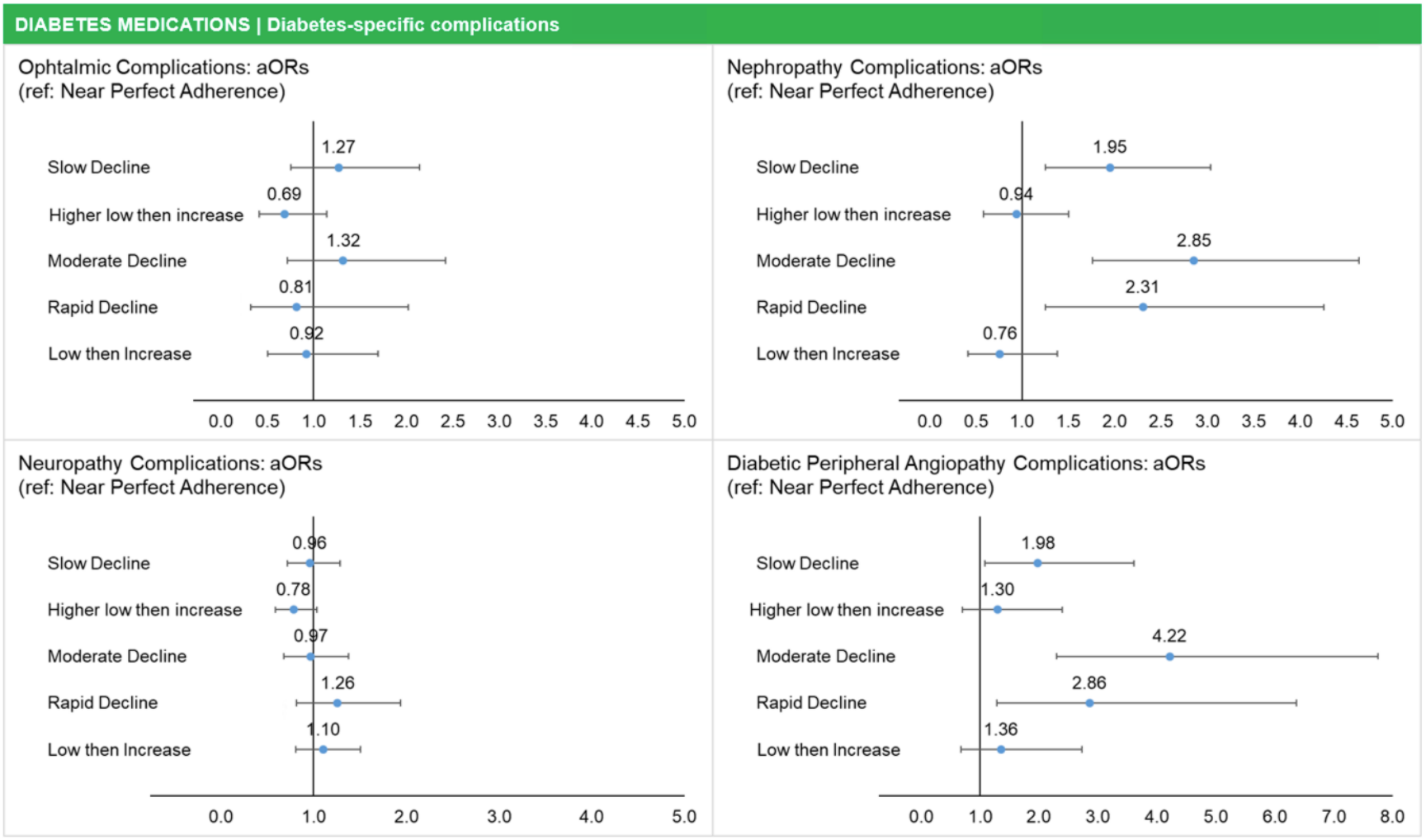

Figure 32 – Predictive aORs of medication adherence trajectories for diabetes medications and diabetes-specific outcomes: ophthalmic complications, nephropathy, neuropathy, and peripheral angiopathy

The adjusted ORs of the several logistic regressions exhibited clinically meaningful trends. For instance, anything less than *near-perfect adherence* to hypertension medications or statins resulted in a higher likelihood of MI, when compared to *near-perfect adherence* (hypertension medications: $aOR_{rapid\ decline} = 1.52$, p-value= 0.005, $aOR_{slow\ decline} = 1.51$, p-value<0.001; statins: $aOR_{rapid\ decline} = 1.56$, p-value=0.007, $aOR_{moderate\ decline} = 1.71$, p-value<0.001, $aOR_{slow\ decline} = 1.91$, p-value<0.001, $aOR_{low\ then\ increase} = 1.45$, p-value=0.002) (Figures 29 and 30). The same was observed for the declining adherence trajectories of diabetes medications patients ($aOR_{rapid\ decline} = 2.45$, p-value=0.002, $aOR_{moderate\ decline} = 2.80$, p-value<0.001, $aOR_{slow\ decline} = 1.88$, p-value=0.004) (Figure 31). However, the trajectories that observed an increase over time, such as the *low then increase* and *high then increase* trajectories in the diabetes medications model did not show any statistically significant difference in the likelihood of suffering from an MI compared to near-perfect adherers ($aOR_{low\ then\ increase} = 1.28$, p-value=0.352, $aOR_{high\ then\ increase} = 1.25$, p-value=0.314).

All declining adherence trajectories of hypertension medications were more likely to suffer from a stroke than near perfect adherers ($aOR_{rapid\ decline} = 1.43$, p-value=0.004, $aOR_{slow\ decline} = 1.31$, p-value=0.006) (Figure 29). Nevertheless, only selected declining adherence trajectories displayed statistically significant differences of the likelihood of stroke in the statins and diabetes medications trajectory models (Figures 30 and 31). Only *rapid decline*rs of adherence to statins seemed to be more likely to suffer from a stroke compared to near perfect adherers ($aOR_{rapid\ decline} = 1.37$, p-value=0.031). Similarly, only *slow decline*rs of adherence to diabetes medications were more likely to suffer from a stroke compared to near perfect adherers ($aOR_{slow\ decline} = 1.43$, p-value=0.047).

The results of the diabetes-specific outcomes showed that all non-perfect adherence trajectories to diabetes medications did not have statistically significant differences in the likelihood of ophthalmic complications or neuropathy complications compared to perfect adherers (Figure 32). As observed for MI, the adherence trajectories to diabetes medications that were followed by an increase in adherence with time did not exhibit any statistically significant differences in any of the diabetes-specific outcomes compared to perfect adherers. Only those that were consistently declining showed more propensity for nephropathy and angiopathy complications compared to perfect adherers (nephropathy: $aOR_{rapid\ decline} = 2.31$, p-value=0.008, $aOR_{moderate\ decline} = 2.85$, p-value<0.001, $aOR_{slow\ decline} = 1.95$, p-value=0.003; angiopathy: $aOR_{rapid\ decline} = 2.86$, p-value=0.010, $aOR_{moderate\ decline} = 4.22$, p-value<0.001, $aOR_{slow\ decline} = 1.98$, p-value=0.025). Remarkably, the declining trajectories that registered higher adherence for the longest time, such as the slow declining adherence trajectories are those with the least aORs for nephropathy and angiopathy.

4.4.2 PREDICTIVE ABILITY COMPARISON: LOGISTIC REGRESSION VERSUS MACHINE LEARNING ALGORITHMS

In addition to the logistic regressions, machine learnings algorithms based on random forests were implemented to investigate whether the predictive ability between medication adherence trajectories and outcomes could be improved. All machine learnings models included the same dependent and independent variables of the adjusted models implemented as logistic regressions.

The performance of machine learning algorithms like the ones implemented in this research work are known to be influenced by how imbalanced the outcome variable is. In other words, how rare the outcome event can be.

To provide a perspective of how imbalanced each outcome modelled was, relative frequencies of MI, stroke, and diabetes-specific complications by trajectory models are shown in Table 23.

Table 23 - Relative frequencies of outcomes per medication adherence trajectory model

	Hypertension medications		Statins		Diabetes medications					
Trajectories	MI (%)	Stroke (%)	MI (%)	Stroke (%)	MI (%)	Stroke (%)	Ophthalmic complications (%)	Neuropathy (%)	Nephropathy (%)	Angiopathy (%)
Rapid decline	9.00%	14.87%	10.04%	13.85%	9.32%	12.01%	2.48%	15.11%	7.45%	6.42%
Moderate decline	-	-	10.77%	16.19%	12.46%	11.92%	4.63%	16.01%	9.07%	6.76%
Slow decline	8.30%	13.06%	10.91%	12.38%	8.37%	11.65%	3.85%	15.27%	7.01%	3.85%
Low then increase	-	-	8.01%	11.22%	5.16%	11.16%	3.49%	17.29%	4.60%	2.79%
high then increase	-	-	-	-	5.48%	7.46%	1.98%	11.64%	4.03%	1.83%
Near perfect adherence	4.75%	7.64%	8.21%	8.53%	3.72%	7.44%	2.85%	11.96%	3.72%	1.90%

All random forest models were estimated using 10 cross validation folds, using a randomized 80:20 data split. The optimal number of combinations of predictors for each random forest models is available in Appendix 3, including model accuracies graphs, and the ranked variable importance factors. Likewise, the extreme gradient boosting random forest models were estimated with the same independent variables and similar random data split. For each extreme gradient boosting model, the maximum number of trees was set to 8 and a learning rate of 0.01. The training error plots are available in Appendix 3.

As seen on Table 24, machine learning algorithms like random forests produce high accuracy rates, with most models scoring above 90%. However, accuracy is described as an inadequate measure to assess model performance, given the high cost of misclassification.

Table 24 - Machine learning algorithms model accuracies

Model accuracies			
		ML Algorithm	
	Outcomes	*Random Forest*	*Extreme Gradient Boosting*
Hypertension medications	MI	0.936	0.929
	Stroke	0.896	0.897
Statins	MI	0.931	0.928
	Stroke	0.905	0.907
Diabetes medications	Opthalmic complications	0.977	0.961
	Neuropathy Complications	0.905	0.889
	Nephropathy Complications	0.958	0.962
	Angiopathy complications	0.979	0.975
	MI	0.936	0.949
	Stroke	0.928	0.937

Consequently, model performance was assessed by computing the c-statistic, since it provides the level of concordance between the true positivity and the false positivity rates. Table 25 displays a heat map of the c-statistic for all models estimated using logistic regressions and machine learnings algorithms, like random forests and extreme gradient boosting. In this heat map, the highest values are denoted in green, while the lowest values are highlighted in light orange.

Table 25 - Model fit assessment by comparison of the c-statistic: logistic regression versus random forest versus extreme gradient boosting models

	c-statistic			
	Outcomes	**Logistic regression**	**Random Forest**	**Extreme Gradient Boosting**
Hypertension medications	MI	0.748	0.500	0.528
	Stroke	0.801	0.507	0.544
Statins	MI	0.754	0.613	0.649
	Stroke	0.809	0.700	0.660
Diabetes medications	Ophthalmic complications	0.743	0.706	0.607
	Neuropathy Complications	0.667	0.734	0.640
	Nephropathy Complications	0.724	0.719	0.726
	Angiopathy complications	0.710	0.699	0.656
	MI	0.793	0.626	0.648
	Stroke	0.810	0.749	0.740

Except for the diabetes-specific outcomes of neuropathy and nephropathy, logistic regression models outperformed the machine learning algorithms, indicating that the latter had

poor predictive ability. Consequently, logistic regression models were selected as the better predictive models. The ROC curves for each of the logistic regression are displayed in Figures 33, 34, 35, and 36.

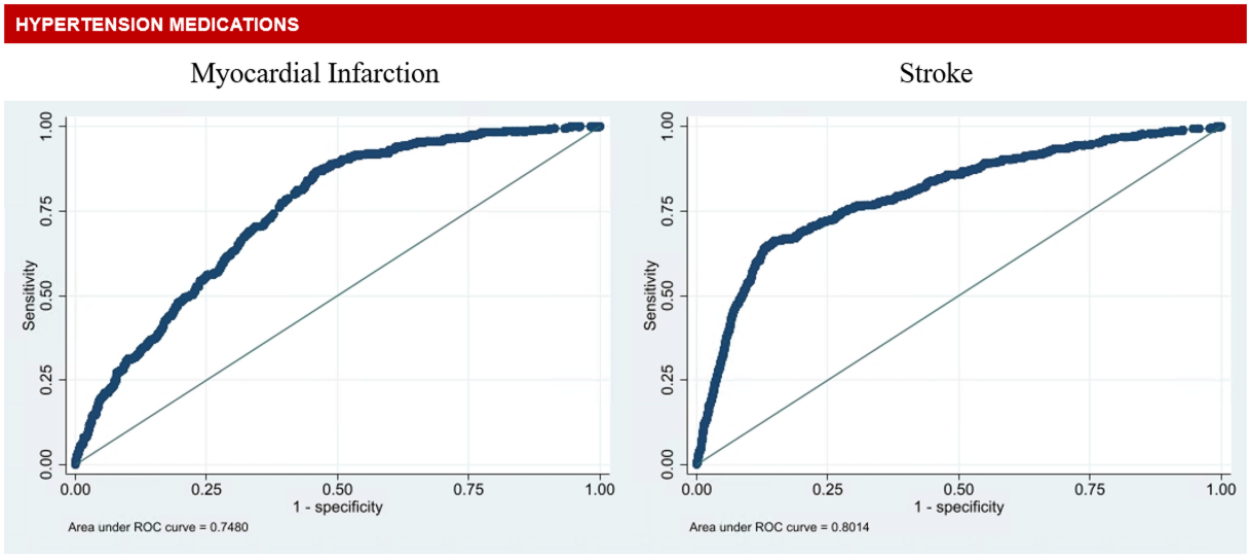

Figure 33 - ROC curves of logistic regression models: outcomes prediction of the hypertension medications adherence trajectory models

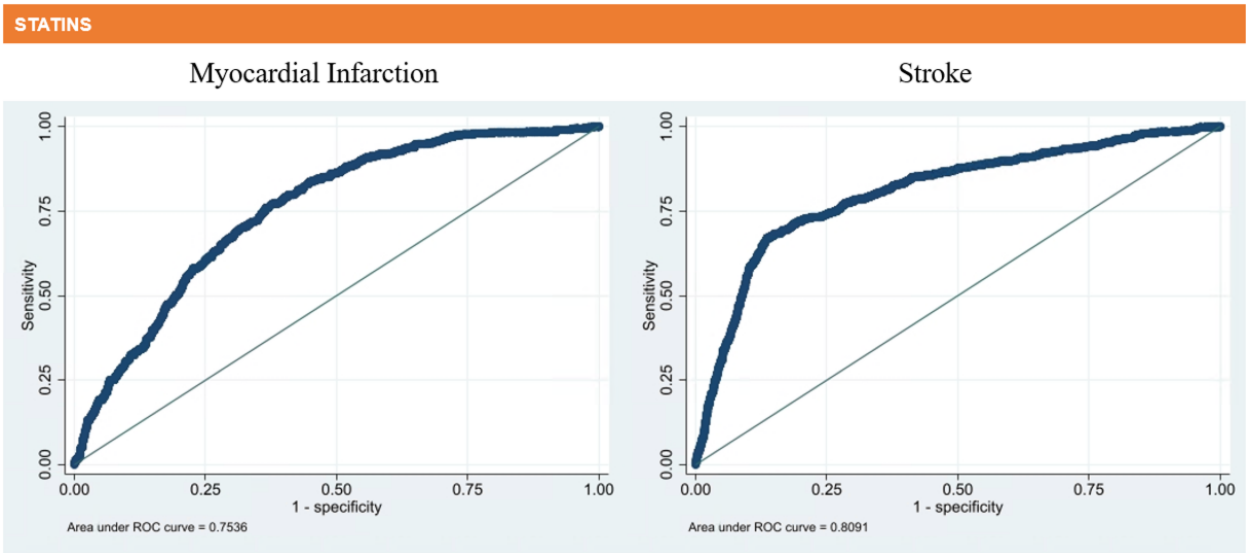

Figure 34 - ROC curves of logistic regression models: outcomes prediction of the statins adherence trajectory models

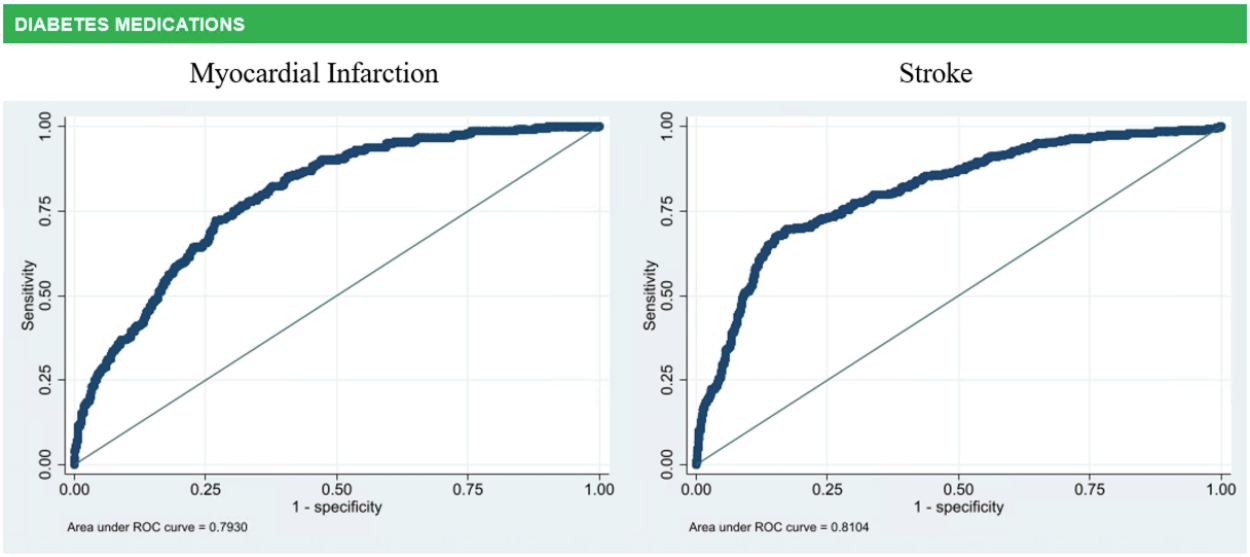

Figure 35 - ROC curves of logistic regression models: outcomes prediction of the diabetes medications adherence trajectory models

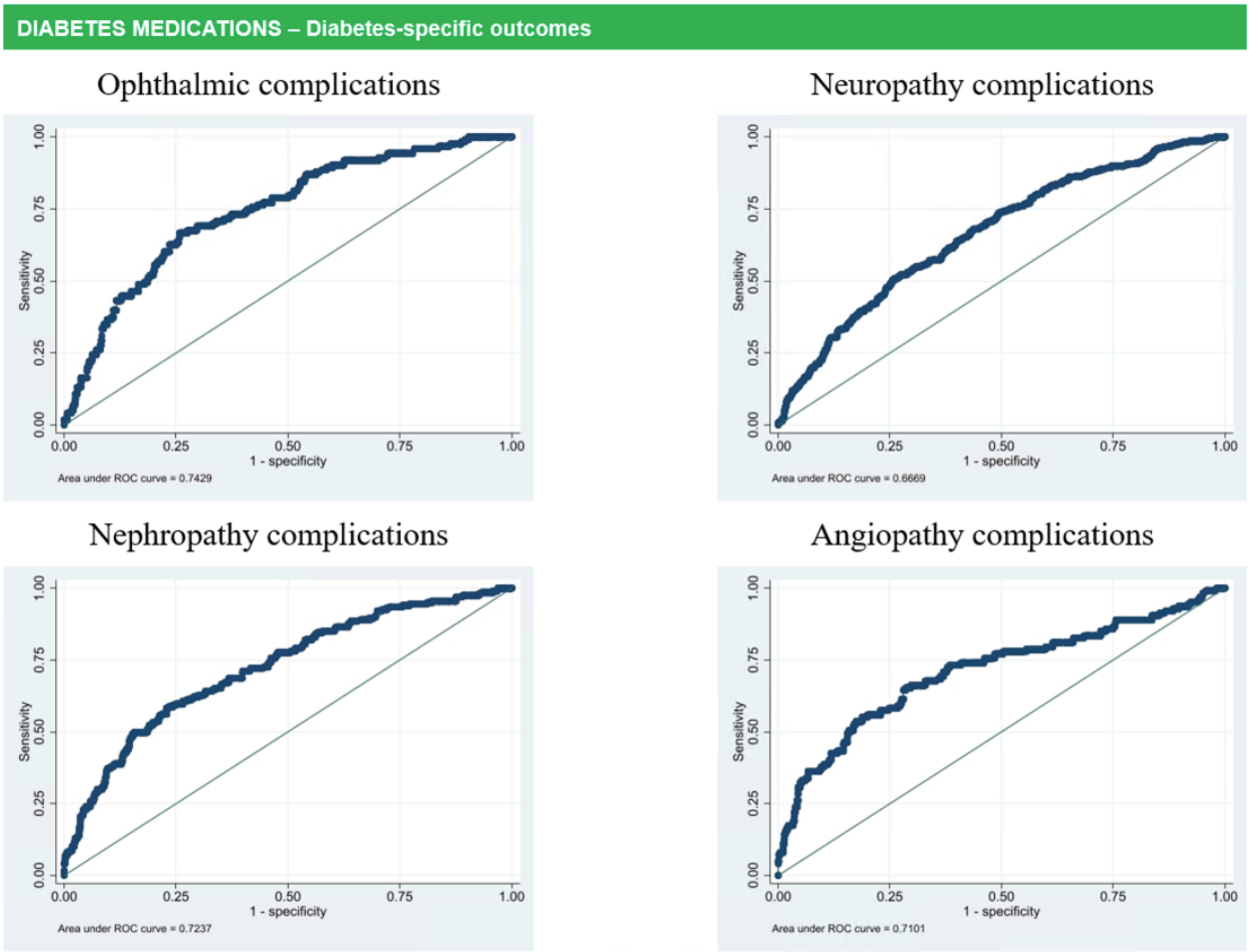

Figure 36 - ROC curves of logistic regression models: diabetes-specific outcomes prediction of the diabetes medications adherence trajectory models

5. DISCUSSION

5.1 SAMPLE MEDICARE POPULATION

The individuals included in this research study were Medicare beneficiaries, who were primarily participants in the HRS. To the extent of our knowledge, this study pioneered the characterization of the medication adherence trajectories and its time-varying predictors in patients taking the drugs included in the MA and Part D Plan Star Ratings Program, and its association with disease-specific outcomes. The results presented in this section have significant clinical and health policy implications in the United States. Despite HRS researchers' efforts to sample according to the US population benchmarks, the sample weights were not available for the linked Medicare participants. Consequently, the sample population had a similar proportion of Whites (77% vs 76%), but a slightly more female (61% vs 55%), Hispanic (12% vs 8%), and less likely to have a college degree (20% vs 27%), when compared to the overall Medicare population. [141] When considering the types of medications used by the participants in this study, approximately half of the sample (47%) was taking medication for hypertension, diabetes, and high blood cholesterol, while 29% was taking statins combined with diabetes medication. Twenty four percent of the participants in this study were being medicated for hypertension only. This suggests that the sample population is of high cardiovascular risk, especially for MI and stroke. [142-144] Therefore, it is possible that these differences can affect the generalizability of the interpretation of the results of this study.

5.2 MEDICATION ADHERENCE TRAJECTORIES

The group-based trajectory analysis conducted in this study yielded slightly different results in the cohorts taking hypertension and diabetes medications, and statins. The cohort

taking hypertension medications yielded three trajectories, five in the cohort taking statins, and six trajectories in the cohort taking diabetes medications. These differences have to be interpreted in the light of the probabilistic nature of group-based trajectory modeling: the number of trajectories in a given sample is not deterministic as they pertain to how likely that specific sample coalesces around unobserved trends over time. Thus, in the cohort taking hypertension, patients most likely follow into a *near-perfect adherence*, a *slow* and *rapid declining* trajectories. These trajectories are also observed in the cohorts taking diabetes medications and statins, in addition a *moderate declining* trajectory, and trajectories exhibit an "ebb and flow" shape. The number and shape of the trajectories identified in this study is consistent with previous literature and with the findings of the systematic literature review presented in Chapter 2. Typically, medication adherence to antihypertensives, statins, and oral antidiabetics is characterized by one trajectory of consistent high adherence, difference speeds of declining adherence, and trajectories with declining adherence that are followed by an increase.[95] This consistency strengthens the external validity of the results of this research study. All three trajectory models were highly significant, and the post-hoc analyses of average posterior probabilities and odds of correct classification indicated that all models were highly predictive of each participants actual PDC measurement (Tables 12 and 13).

While prior research has focused on individual analysis of medication adherence trajectories of specific drug classes, this research study is the first to conduct an exhaustive analysis of medication adherence in patients of high cardiovascular risk taking multiple drugs (initiators and previous users) for hypertension, high blood cholesterol, and diabetes over an 8 year period. While all patients taking statins were also taking antidiabetics or antihypertensives, approximately one third of the patients taking antihypertensives were only taking those drugs. It

is fair to assume that the medication regimen complexity is higher if the patients are taking drugs for hypertension and other comorbidities than for hypertension alone. [145,146] While medication regimen complexity can be a useful intervention for pharmacists to identify patients in need of MTM services, it is also possible that with higher treatment complexity, the more diverse medication adherence behavior can be. This would explain why in this particular research study the hypertension medications model yielded only 3 trajectories, while the statins and diabetes medications models yielded 5 and 6 trajectories, respectively. Our results using an alternative methodological approach to examining medication adherence add to the existing literature that hints at the connection between regimen complexity and pill burden and medication adherence. [147] While a significant body of evidence typically suggests that about 50% of the patients are adherent to their medication, our results show that as the number of probable adherence trajectories grows, the proportion of patients following *near-perfect adherence* trajectories diminishes. For example, in the diabetes medications model with 6 trajectories, only 24.2% of patients likely follow a *near-perfect adherence*. While the follow-up period was appreciably longer than in previous studies (8 years versus 1-2 years), which would increase the probability of non-adherence behaviors, the diseases of interest require chronic treatments, are responsible for a high lifetime burden, and require longer follow-up periods than 1 year. [148] Our results also show that when examining long-term medication adherence, the much common blanket statement "*patients are non-adherent about 50% of the time*" ignores the fluidity of non-adherence behavior demonstrated in this study and is inconsistent with the data. [149,150]

The results also suggest that the longitudinal trends of medication adherence display similar trajectories (in number and in shape), even in a meaningfully larger time frame. Furthermore, this research study estimated group-based trajectory models using a continuous

measurement of monthly PDC, which has been reported to yield better prediction and classification of patients according to their actual PDC, when compared to a binary categorization of monthly PDC into adherent and non-adherent. The higher predictive ability of using continuous measurements extends to examining the relationship between trajectories and outcomes. [151] This is of particular importance given that a significant number of the previous studies estimated trajectories by dichotomizing patients into adherent/non-adherent based on an 80% threshold. [60,62,74,79,86]

5.3 PREDICTORS OF MEDICATION ADHERENCE TRAJECTORIES

This research study focused on time-fixed and time-varying predictors of medication adherence trajectories, following a theoretical framework based on the ABM. The analysis of the time-fixed characteristics continued to support the notion that low socioeconomic status (as demonstrated by being non-white, living below poverty threshold, being Medicaid beneficiary, not being a homeowner) is predictive of belonging to declining medication adherence trajectories. However, the adjusted models suggest a higher preponderance of characteristics associated with comorbidities (high BMI, cancer survivorship, and clinical depression) than with indicators of low socioeconomic status, particularly for the hypertension medications and statins models. Furthermore, the effect of risk accumulation in hypothetical scenarios exhibited clinically meaningful results: in all three models, the *near-perfect adherence* trajectory transitioned from being the largest trajectory group in the model to the smallest group. This pattern is of special note in the statins and diabetes medications model, since the base-case analysis with the sample showed that much less than half of the patients follow the *near-perfect adherence* trajectory.

To further explore the causes of non-adherence, this study also pioneered the estimation of a multi trajectory group-based model, conjecturing that other factors, that change with time, can help explain changes and inflexions of individual medication adherence trajectories. Those characteristics included the enabling and need factors typically described in the ABM theoretical framework. [25,64,65,69] While the results did not categorically show which time-varying factors change with the medication adherence trajectories, some trends consistent with previous findings were observed. [62,152,153] For example, a decrease in the probability of living under the poverty threshold was found to occur in parallel with increases in medication adherence in the *moderate decline* trajectories of the statins model. Similarly, loss of spouse was generally found to occur in parallel with decreases in adherence, or the opposite, in which regaining a spouse was met with increases or maintenance of high adherence. These results are consistent with previous findings that revealed social support (e.g.: informal caregivers, spouses, and friends) as positive factors of medication adherence. [154-156] The level of autonomy of patients yielded the most significant demonstration of how multi group-based trajectory modeling can be used to identify patients at risk of following non-adherence trajectories. It is recognized that aging is associated with a decrease in autonomy, which usually translates in more difficulty in performing certain daily tasks. This type of autonomy is what is examined in scales such as ADL and IADL. While it is expected that ADL and IADL scores increased with age, the results show that declining adherence trajectories exhibit larger increases in IADL and ADL scores (Figures 23, 25, and 27). Interestingly, Mizokami and colleagues demonstrated that worsening IADL scores, particularly those in the shopping and self-medication tasks were associated with cognitive disfunction. [157] A significant body of research in geriatrics predictably reported worsening IADL and ADL with higher probability of non-adherence. [158-160]

Furthermore, the implementation of multi group-based trajectory modeling demonstrated the conceptual usefulness of the method in further investigating the time-varying experiences that patients go through and the limitations of the HRS as the source of data for this type of analysis. Additional aspects pertaining to social support and informal caregiving could have been included in this research study. These included, for example, the number of resident and non-resident children, number of children living within 10 miles of the respondent, help received from child, sibling, or other family member. [161] However, the level of data missingness in the HRS dataset made those types of variables unusable for this research study. Moreover, the HRS core survey is administered every two years, which limits the granularity of the data and how it can be modelled with monthly PDC using multi group-based trajectory methods. [162] Given the potential of using multi group-based trajectory methods to determine time-varying predictors of medication adherence longitudinal trends, the results of the research study highlight the importance of adding more questions to the HRS questionnaire. Those could include questions pertaining to an expanded medication adherence conceptual framework proposed by Krousel-Wood and colleagues, especially those concerned social factors like built environment, health care system, and disease and treatment factors[163]:

HEALTH CARE SYSTEM FACTORS:

- Communication between patient and provider
- Perceived discrimination (sexual, race, education, or income)
- Accessibility to care (waiting times, distance to care, transportation availability)

DISEASE AND TREATMENT FACTORS:

- Beliefs, mistrust, and preferences
- Experience of adverse effects

5.4 MEDICATION ADHERENCE TRAJECTORY OUTCOMES AND POLICY IMPLICATIONS

The third aim of this research study examined the association between medication adherence trajectories and health outcomes, which include MI, stroke, and four diabetes-specific outcomes, ophthalmic complications, nephropathy, neuropathy, and diabetic peripheral angiopathy complications. All outcome events were identified during the follow-up period, that is, during the same time period in which monthly PDC was estimated. Two alternative methods were used to examine the associations between medication adherence trajectories and outcomes. The methods using logistic regression models performed considerably better than random forests and extreme gradient boosting random forests. While in some circumstances it is common for machine learning algorithms to perform better than traditional regression models, our results highlight the limitations of these methods. Random forest algorithms tend to generate predictive models when the outcome variable is significantly imbalanced. [164] In this research study, the frequency of all outcomes was never higher than 16% in a given trajectory, and in most cases the outcomes relative frequency was below 10%. In other words, the proportion of patients without outcomes disproportionately higher than those with outcomes. The methods used in this research study used *off-the-shelf* machine learning packages, but strategies to overcome the limitations of imbalanced data have been proposed to date, like boostrapping with bagging classifiers and random Undersampling, or boosting methods using the SMOTEBoost algorithm. [164-166] However, it was not possible to verify whether those strategies have been validated in health services research and implementation of those strategies fell outside of the scope of this research study. Nevertheless, the complexity of implementing machine learning algorithms that adequately help researchers address unique research questions highlights the need for collaborative teams and

input from experts in computer science. This is exactly what the leaders of the ISPOR's Machine Learning working group recommend in the forthcoming PALISADE checklist for using machine learning algorithms in health outcomes research. [167]

The adjusted logistic regression models revealed that all declining trajectories are more likely to experience MI when compared to *near-perfect adherence* trajectories. These findings highlight the importance of supporting patients with their medication experience to improve outcomes. Nevertheless, the results also feature a more nuanced perspective on what should be considered ideal medication adherence and whether the traditional cut-off at PDC>80% is adequate for health care quality metrics. Outcomes such as stroke were observed to more likely to occur in patients following specific declining adherence trajectories when compared to *near-perfect adherence*: a) in all declining trajectories in the hypertension medications model, b) only the *rapid decline* trajectory in the statins model, and c) the *slow decline* in the diabetes medications model. Moreover, diabetes-specific outcomes, including nephropathy and diabetic peripheral angiopathy were only more likely in the declining trajectories (*rapid, moderate, and slow*) when compared to *near-perfect adherence*. Conversely, no statistically significant differences were found for ophthalmic complications and neuropathy between non-adherent trajectories and *near-perfect adherence* trajectories. These results suggest that whatever "optimal" should be is likely outcome and disease-specific. For example, trajectories displaying an ebb-and-flow shape like the *higher then increase* and *low then increase* trajectories in the diabetes medication model showed no greater likelihood of MI, stroke, and any of the diabetes-specific outcomes. The same was observed in the statins model for stroke. Therefore, from a clinical perspective to improve patient outcomes, the goal posts the ideal medication adherence should be flexible. These findings present a great opportunity for demonstrating the long-term

impact pharmacy practice interventions can have on patient outcomes. Because of how the Quality Bonus Payments are structured, improvements in medication adherence in hypertension medications, statins, and diabetes medications are only rewarded if exceeding a PDC > 80% during the Medicare coverage period. However, this research study demonstrated that trajectories of medication adherence can fluctuate over time and patients, who are not perfectly adherent can still experience certain outcomes with the same probability as those perfectly adherent over time, particularly patients taking statins and diabetes medication.

At a first glance, it may seem that the lack of unique association between certain outcomes and *near-perfect adherence* trajectories should diminish or even remove the impact of medication adherence quality metrics in the Star Ratings program, particularly those associated with statins use and diabetes. Notwithstanding the evidence presented in this research study, one cannot ignore how the implementation of Star Ratings program and the increase in weight of medication-based measures seems to be associated with improvements in medication adherence of the targeted and non-targeted medications of the program. [168] A different option is suggested instead.

Firstly, organizations responsible for developing quality metrics focused on medication adherence and medicines appropriate use (i.e., PQA) should clarify how each quality metric is directly tied to long-term economic and clinical outcomes. The studies presented by PQA as evidence for supporting the existing quality metrics of the Star Ratings program were limited to economic, including all-cause health care utilization and expenditures. [169-172] While dollars are an obviously important metric for providers and payers, when assessing the impacts of a measure such as medication adherence, one should consider the disease-specific outcomes of the therapeutic indication for that medication.

Secondly, quality metrics should be developed implementing methods that better elicit medication adherence behaviors and using longer periods of follow-up. For example, Menditto and colleagues suggest using persistence measurements instead as indicators of medication adherence. [173] Persistence refers to the time between treatment initiation and the last dose preceding discontinuation. [174] Using persistence for chronic medications that patients take for years, as opposed to binary observations of annual adherence logically produces a more compelling perspective of the medication-taking behavior. [62,173]

Surprisingly, the PQA-sponsored economic studies followed a 1-year long temporal perspective. [169-172] Given the evidence presented in this study, 1-year is a manifestly too short period to examine the long-term effects of adherence to chronic medication for chronic diseases like hypertension, high blood cholesterol, and diabetes. As evidenced in this research study and in previous literature, the simple dichotomization of medication adherence based on the 80% threshold is not an adequate methodological approach to describe adherence behavior over time and the association with clinical outcomes seems to be disease-dependent. [52,63,79,85] Consequently, the broader implementation of methods such as group-based trajectory modeling following well-defined criteria should be encouraged. [173] While providers in the United States experience a burden of measurement of quality metrics, the demonstration of interventions that define patient experience and outcomes should not be ignored simply because of how complex those methods are. [175] Farley and Urick suggested the expansion of the Star Ratings to other drugs classes and simplify the problem of drug-class adherence calculation by deploying the multiple medications for chronic conditions (MMCC) measure as an alternative to PDC. [53] In fact, this measure has been successfully implemented in a Medicaid alternative value-based model to reward community pharmacies in North Carolina. [176] Future research can compare the association

between medication adherence trajectories with clinical and economic outcomes using drug class monthly PDC measurements and MMCC.

5.5 LIMITATIONS

The Health and Retirement Study (HRS) is a nationally representative panel study to examine the intricacies of aging in the United States. However, several limitations were identified while conducting the present research study that could compromise the external validity of this research study. Firstly, only a subset of the total sample of the HRS allows linkage to the Medicare administrative health care claims. This reduced the potential sample from 42,235 to a total of 27,895. [177] The actual sample size included in this research study was significantly lower. Moreover, the nationally representativeness was lost when building the medication adherence trajectory models. To date, it is not yet possible to include sampling weights or stratification variables when building group-based trajectory models. Effectively, the sample included in this study was slightly more female and whiter than the overall Medicare population, which is indicative of potential selection bias. Furthermore, the examination of concomitant medications showed that the sample included in this study was probably of higher cardiovascular risk than the overall Medicare population.

GBTM is based on several assumptions, included those about population distribution of developmental trajectories. Consequently, this research study may have overlooked how individual differences in development are distributed in the sample. This is particularly important given that it does not occur in other finite mixture models, such as the growth curve model (GCM), and the HRS is a longitudinal panel that recruits new participants every two years from different demographic generations. In summary, this means that the number of trajectories that

result from a GBTM does not account for random effects, like in other methods, which threatens the internal validity of the method. The addition of random effects in the calculation of trajectories would likely allow for more within-group variability in individual-level trajectories. This would eventually result in a calculation of reduced number trajectories. In developmental theories, unobserved sub-groups of people are considered to follow individual trajectories. Despite the methodological limitations, GBTM can still be considered a better approach to measuring adherence over time because the trajectory groups are considered to be approximations of individuals who follow approximately [probabilistic] the same development trajectory (i.e.: the same medication adherence trajectory). As Haviland and colleagues put it, trajectory groups in GBTM are considered to be latent longitudinal strata that capture differences across groups in the shape and intensity of their trajectories. [178] Nevertheless, the results showed that even while using a typically different data source to estimate medication adherence trajectories, the number and shape of the those trajectories was consistent with prior research. Therefore, vigorous concerns with the external validity of this research study and each of the medication adherence trajectory models are not warranted.

Additionally, despite linking HRS to event-level data from Medicare, the actual HRS survey is only administered every two years. This means that the multi-trajectory models were estimated using monthly adherence measurement with time-varying predictors retrieved during only 5 timepoints: 2008, 2010, 2012, 2014, 2016. Moreover, it was assumed that the data collected in each biennium corresponded to first January of that period (i.e.: data from 2008 was assumed to correspond to January 2008). It is also possible that a more granular dataset from time-varying predictors could have yielded more compelling and consistent results. Further

research should be conducted using data collected during the same timepoints as medication adherence.

This research study presented a conceptual proof that time-varying predictors of medication adherence can be conducted. However, the core data from the HRS included only a subset of variables included in the Andersen's Behavior Model of health care services use. Specifically, it was not possible to include information pertaining to health care system-related factors to examine their relationship with medication adherence trajectories. Thus, future research investigating predictors of medication adherence trajectories should incorporate this information. Computational power was a major setback in this study – each group based trajectory model ran for several hours. Future research that includes samples of millions of patients should consider high performing computers in order to run the analysis and troubleshoot within a reasonable timeframe.

Finally, the machine learning (ML) algorithms used to examine the relationship between outcomes and medication adherence trajectories showed surprising poorer predictive ability than logistic regressions. As mentioned previously, the outcome imbalanced variables can cause random forests to perform poorer than other methods. Future studies deploying these methods should work to validate the use of these methods to draw statistical inferences, particularly the procedures used to fine tune the ML models (i.e.: number of cross-validations, depth of trees, learning rate, etc.). Given the complexity of these methods, research teams could benefit from input provided by computer scientists for model validation. The forthcoming ISPOR recommendations on ML use in health economics and outcomes research will also provide significant guidance to future studies using these methods.

6. CONCLUSION

In summary, this research study consisted in a comprehensive examination of the longitudinal medication adherence trajectories in patients taking medications ranging from antihypertensives, statins, and diabetes medications. Group-based trajectory models were implemented with a sample of participants in the Health and Retirement Study linked to Medicare administrative health care claims from 2008 through 2016. The hypertension medications model yielded 3 trajectories, *near-perfect adherence*, *slow*, and *rapid decline*. The statins models elicited 5 trajectories, including *near-perfect adherence*, *slow*, *moderate*, and *rapid decline,* and *low then increase adherence*. The diabetes medications model exhibited an extra trajectory, *high then increase*, similar to the *low then increase* adherence of the statins models, despite never decreasing as much as the latter in the beginning of the follow-up period. Furthermore, a pioneering approach using multi group-based trajectory modeling was implemented to investigate both time-fixed and time-varying predictors of medication adherence trajectories. While clear and consistent trends were not observed for the time-varying risk factors, this research study served as a conceptual proof that this method can be applied when examining predictors of medication adherence behavior, such as life-changing events or time-varying contextual factors in patients' lives. Moreover, it is possible that the absence of consistent results in the multi trajectory group-based models are more a function of the dataset used in this study than a limitation of the method itself. Consequently, suggestions of data variables that could be incorporated in future medication adherence studies, including in the health care utilization section of the Health and Retirement Study, are provided.

Finally, several predictive models were implemented to examine the relationship between medication adherence trajectories and health outcomes, which included myocardial infarction

and stroke for all three models, and diabetes-specific outcomes (ophthalmic complications, neuropathy, nephropathy, and diabetes peripheral angiopathy). Overall, logistic regression models exhibited better predictive validity than machine learning algorithms based on random forests. Data characteristics, including the imbalance of the outcome variables possibly explain why machine learning algorithms performed worse than logistic regression. The results of this study highlight the need for collaborative approach when implementing more complex machine learning-based methods. The predictive models revealed non-adherence trajectories are not necessarily more likely to experience stroke and diabetes-specific outcomes than near-perfect medication adherence trajectories. Given how the Medicare Advantage and Part D Plans Star Ratings program are structured around quality metrics heavily influenced by medication adherence measurements, policy and quality metrics implications are discussed in light of the results of this study. Future directions include two aspects: 1) development of measurement variables that capture the context in which patients use medication to determine the life-changing events, or time-varying circumstances that most influence medication adherence; and 2) determine the specific outcomes that are deemed valued for quantifying medication adherence quality metrics and the horizon used when establishing value-based payments.

7. REFERENCES

1. World Health Organization. Adherence to Long-Term Therapies: Evidence for action. World Health Organization Accessed 20/01/2020, https://apps.who.int/medicinedocs/pdf/s4883e/s4883e.pdf
2. Holdford DA, Inocencio TJ. Adherence and persistence associated with an appointment-based medication synchronization program. *J Am Pharm Assoc (2003)*. Nov-Dec 2013;53(6):576-83. doi:10.1331/JAPhA.2013.13082
3. McCarthy R. The price you pay for the drug not taken. *Bus Health*. Oct 1998;16(10):27-8, 30, 32-3.
4. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. Aug 4 2005;353(5):487-97. doi:10.1056/NEJMra050100
5. New England Healthcare Institute. Thinking Outside the Pillbox: A System-wide Approach to Improving Patient Medication Adherence for Chronic Disease. Accessed 03/14/2020, https://www.nehi.net/writable/publication_files/file/pa_issue_brief_final.pdf
6. Viswanathan M, Golin CE, Jones CD, et al. Interventions to Improve Adherence to Self-administered Medications for Chronic Diseases in the United States. *Annals of Internal Medicine*. 2012/12/04 2012;157(11):785-795. doi:10.7326/0003-4819-157-11-201212040-00538
7. National Archives. Medicare and Medicaid Act (1965). Updated 02/08/2022. Accessed 05/03/2022, https://www.archives.gov/milestone-documents/medicare-and-medicaid-act
8. U.S. Centers for Medicare and Medicaid Services. Medicare.gov: What's Medicare? Accessed 05/04/2022, https://www.medicare.gov/what-medicare-covers/your-medicare-coverage-choices/whats-medicare
9. U.S. Department of Health & Human Services. What is Medicare Part C? Updated 08/03/2021. Accessed 05/05/2022, https://www.hhs.gov/answers/medicare-and-medicaid/what-is-medicare-part-c/index.html
10. Pharmacy Quality Alliance. PQA Measure Use in CMS' Part D Quality Programs. 18.01.20, Accessed December 10, 2019, https://www.pqaalliance.org/medicare-part-d
11. U.S. Centers for Medicare & Medicaid Services. Medicare Program; Changes to the Medicare Advantage and the Medicare Prescription Drug Benefit Programs for Contract Year 2013 and Other Changes. In: U.S. Centers for Medicare & Medicaid Services, editor.: Federal Register.
12. Burnier M. Is There a Threshold for Medication Adherence? Lessons Learnt From Electronic Monitoring of Drug Adherence. *Front Pharmacol*. 2018;9:1540. doi:10.3389/fphar.2018.01540
13. Brown MT, Bussell JK. Medication adherence: WHO cares? *Mayo Clin Proc*. Apr 2011;86(4):304-14. doi:10.4065/mcp.2010.0575
14. Seixas A, Connors C, Chung A, Donley T, Jean-Louis G. A Pantheoretical Framework to Optimize Adherence to Healthy Lifestyle Behaviors and Medication Adherence: The Use of Personalized Approaches to Overcome Barriers and Optimize Facilitators to Achieve Adherence. *JMIR Mhealth Uhealth*. Jun 24 2020;8(6):e16429. doi:10.2196/16429

15. Amico KR, Mugavero M, Krousel-Wood MA, Bosworth HB, Merlin JS. Advantages to Using Social-Behavioral Models of Medication Adherence in Research and Practice. *J Gen Intern Med*. Feb 2018;33(2):207-215. doi:10.1007/s11606-017-4197-5
16. Bandura A. Health promotion by social cognitive means. *Health Educ Behav*. Apr 2004;31(2):143-64. doi:10.1177/1090198104263660
17. Fishbein M. A reasoned action approach to health promotion. *Med Decis Making*. Nov-Dec 2008;28(6):834-44. doi:10.1177/0272989X08326092
18. Rippetoe PA, Rogers RW. Effects of components of protection-motivation theory on adaptive and maladaptive coping with a health threat. *J Pers Soc Psychol*. Mar 1987;52(3):596-604. doi:10.1037//0022-3514.52.3.596
19. Prochaska JO, Velicer WF, Rossi JS, et al. Stages of change and decisional balance for 12 problem behaviors. *Health Psychol*. Jan 1994;13(1):39-46. doi:10.1037//0278-6133.13.1.39
20. Arafat Y, Mohamed Ibrahim MI, Awaisu A, et al. Using the transtheoretical model's stages of change to predict medication adherence in patients with type 2 diabetes mellitus in a primary health care setting. *Daru*. Jun 2019;27(1):91-99. doi:10.1007/s40199-019-00246-7
21. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. *Health Educ Q*. Summer 1988;15(2):175-83. doi:10.1177/109019818801500203
22. Willis E. Applying the Health Belief Model to Medication Adherence: The Role of Online Health Communities and Peer Reviews. *J Health Commun*. 2018;23(8):743-750. doi:10.1080/10810730.2018.1523260
23. Amico KR. A situated-Information Motivation Behavioral Skills Model of Care Initiation and Maintenance (sIMB-CIM): an IMB model based approach to understanding and intervening in engagement in care for chronic medical conditions. *J Health Psychol*. Oct 2011;16(7):1071-81. doi:10.1177/1359105311398727
24. Ryan P, Sawin KJ. The Individual and Family Self-Management Theory: background and perspectives on context, process, and outcomes. *Nurs Outlook*. Jul-Aug 2009;57(4):217-225 e6. doi:10.1016/j.outlook.2008.10.004
25. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J Health Soc Behav*. Mar 1995;36(1):1-10.
26. Bronfenbrenner U. *Making human beings human: Bioecological perspectives on human development*. Sage; 2005.
27. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Annu Rev Clin Psychol*. 2010;6:109-38. doi:10.1146/annurev.clinpsy.121208.131413
28. Porta M. *A dictionary of epidemiology*. Oxford university press; 2014.
29. Aschengrau A, Seage GR. *Essentials of epidemiology in public health*. Jones & Bartlett Publishers; 2013.
30. Levy D, Larson MG, Vasan RS, Kannel WB, Ho KK. The progression from hypertension to congestive heart failure. *JAMA*. May 22-29 1996;275(20):1557-62.
31. Centers for Medicare & Medicaid Services. CMS Fast Facts. Updated 03/16/2022. Accessed 05/10/2022, https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Fast-Facts
32. Centers for Medicare & Medicaid Services. National Health Expenditure Data Fact Sheet. Updated 12/15/2021. Accessed 05/10/2022, https://www.cms.gov/Research-

Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet
33. Vintage 2019 Population Estimates
34. French MT, Homer J, Gumus G, Hickling L. Key Provisions of the Patient Protection and Affordable Care Act (ACA): A Systematic Review and Presentation of Early Research Findings. *Health Serv Res*. Oct 2016;51(5):1735-71. doi:10.1111/1475-6773.12511
35. Medicare Advantage and Medicare Prescription Drug Plan Quality Strategy: A Framework for Improving Care for Beneficiaries (2012).
36. Centers for Medicare & Medicaid Services. Value-Based Programs. Updated 03/31/2022. Accessed 05/11/2022, https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs
37. Centers for Medicare & Medicaid Services. History. Updated 12/01/2021. Accessed 05/12/2022, https://www.cms.gov/About-CMS/Agency-Information/History
38. Centers for Medicare & Medicaid Services. How Original Medicare works. Accessed 05/12/2022, https://www.medicare.gov/what-medicare-covers/your-medicare-coverage-choices/how-original-medicare-works
39. H.R.4961 - Tax Equity and Fiscal Responsibility Act of 1982, H.R. 4961, 97th Congress (Rep. Stark FPD-C- 1982). https://www.congress.gov/bill/97th-congress/house-bill/4961
40. Porell FW, Wallack SS. Medicare risk contracting: determinants of market entry. *Health Care Financ Rev*. Winter 1990;12(2):75-85.
41. McGuire TG, Newhouse JP, Sinaiko AD. An Economic History of Medicare Part C. *Milbank Q*. Jun 2011;89(2):289-332. doi:10.1111/j.1468-0009.2011.00629.x
42. Medicare Payment Advisory Commission. *Report to the Congress: Medicare Payment Policy*. March 2009. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/march-2009-report-to-congress-medicare-payment-policy.pdf
43. Jacobson G, Damico A, Neuman T, Gold M. *Medicare Advantage 2015 Spotlight: Enrollment Market Update*. 2015. June 2015. https://files.kff.org/attachment/issue-brief-medicare-advantage-2015-spotlight-enrollment-market-update
44. L&M Policy Research LLC. *Evaluation of the Medicare Quality Bonus Payment Demonstration*. 2016. Accessed 05/12/2022. https://innovation.cms.gov/files/reports/maqbpdemonstration-finalevalrpt.pdf
45. Centers for Medicare & Medicaid Services. 5-Star Special Enrollment Period. Accessed 05/12/2022, https://www.medicare.gov/sign-up-change-plans/when-can-i-join-a-health-or-drug-plan/5-star-special-enrollment-period
46. Centers for Medicare & Medicaid Services. Fact Sheet - 2022 Part C and D Star Ratings. Accessed 03/02/2022, https://www.cms.gov/files/document/2022-star-ratings-fact-sheet1082021.pdf
47. Alliance PQ. PQA Measures Overview. Accessed 05/12/2022, https://www.pqaalliance.org/assets/Measures/PQA_Measures_Overview.pdf
48. U.S. Centers for Medicare and Medicaid Services. Part C and D Performance Data - 2022 Star Ratings Technical Notes. Accessed 2022/01/30, https://www.cms.gov/files/document/2022-star-ratings-technical-notes-oct-4-2022.pdf
49. Alliance PQ. About PQA. Accessed 05/10/2022, https://www.pqaalliance.org/our-story

50. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. An empirical basis for standardizing adherence measures derived from administrative claims data among diabetic patients. *Med Care*. Nov 2008;46(11):1125-33. doi:10.1097/MLR.0b013e31817924d2
51. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Good and poor adherence: optimal cut-point for adherence measures using administrative claims data. *Curr Med Res Opin*. Sep 2009;25(9):2303-10. doi:10.1185/03007990903126833
52. Baumgartner PC, Haynes RB, Hersberger KE, Arnet I. A Systematic Review of Medication Adherence Thresholds Dependent of Clinical Outcomes. *Front Pharmacol*. 2018;9:1290. doi:10.3389/fphar.2018.01290
53. Farley JF, Urick BY. Is it time to replace the star ratings adherence measures? *J Manag Care Spec Pharm*. Mar 2021;27(3):399-404. doi:10.18553/jmcp.2021.27.3.399
54. Steiner JF. Measuring adherence with medications: time is of the essence. *Pharmacoepidemiol Drug Saf*. Mar 2016;25(3):333-5. doi:10.1002/pds.3932
55. Wiecek E, Tonin FS, Torres-Robles A, Benrimoj SI, Fernandez-Llimos F, Garcia-Cardenas V. Temporal effectiveness of interventions to improve medication adherence: A network meta-analysis. *PLoS One*. 2019;14(3):e0213432. doi:10.1371/journal.pone.0213432
56. Nagin DS. *Group-based modeling of development*. Harvard University Press; 2005.
57. Nagin DS. Group-based trajectory modeling: an overview. *Ann Nutr Metab*. 2014;65(2-3):205-10. doi:10.1159/000360229
58. Tobin J. Estimation of relationships for limited dependent variables. *Econometrica: journal of the Econometric Society*. 1958:24-36.
59. Lo-Ciganic WH, Donohue JM, Jones BL, et al. Trajectories of Diabetes Medication Adherence and Hospitalization Risk: A Retrospective Cohort Study in a Large State Medicaid Program. *J Gen Intern Med*. Sep 2016;31(9):1052-60. doi:10.1007/s11606-016-3747-6
60. Hargrove JL, Pate V, Casteel CH, et al. Antihypertensive Adherence Trajectories Among Older Adults in the First Year After Initiation of Therapy. *Am J Hypertens*. Oct 1 2017;30(10):1015-1023. doi:10.1093/ajh/hpx086
61. Librero J, Sanfelix-Gimeno G, Peiro S. Medication Adherence Patterns after Hospitalization for Coronary Heart Disease. A Population-Based Study Using Electronic Records and Group-Based Trajectory Models. *PLoS One*. 2016;11(8):e0161381. doi:10.1371/journal.pone.0161381
62. Franklin JM, Krumme AA, Shrank WH, Matlin OS, Brennan TA, Choudhry NK. Predicting adherence trajectory using initial patterns of medication filling. *Am J Manag Care*. Sep 1 2015;21(9):e537-44.
63. Franklin JM, Shrank WH, Pakes J, et al. Group-based trajectory models: a new approach to classifying and predicting long-term medication adherence. *Med Care*. Sep 2013;51(9):789-96. doi:10.1097/MLR.0b013e3182984c1f
64. Andersen RM, Davidson PL, Baumeister SE. Improving access to care in America. *Changing the US health care system: key issues in health services policy and management*. 3rd ed. Jossey-Bass; 2007:3-31.
65. Alfaifi AA, Lai L, Althemery AU. Barriers in utilizing lipid-lowering agents in non-institutionalized population in the U.S.: Application of a theoretical framework. *PLoS One*. 2021;16(8):e0255729. doi:10.1371/journal.pone.0255729

66. Thorpe CT, Fowler NR, Harrigan K, et al. Racial and Ethnic Differences in Initiation and Discontinuation of Antidementia Drugs by Medicare Beneficiaries. *J Am Geriatr Soc*. Sep 2016;64(9):1806-14. doi:10.1111/jgs.14403
67. Deshpande M, Chewning B, Mott D, Thorpe JM, Young HN. Asthma medication use among U.S. adults 18 and older. *Res Social Adm Pharm*. Nov-Dec 2014;10(6):e113-e123. doi:10.1016/j.sapharm.2014.02.006
68. Bali V, Kamble PS, Aparasu RR. Predictors of concomitant use of antipsychotics and stimulants and its impact on stimulant persistence in pediatric attention deficit hyperactivity disorder. *J Manag Care Spec Pharm*. Jun 2015;21(6):486-98. doi:10.18553/jmcp.2015.21.6.486
69. Babitsch B, Gohl D, von Lengerke T. Re-revisiting Andersen's Behavioral Model of Health Services Use: a systematic review of studies from 1998-2011. *Psychosoc Med*. 2012;9:Doc11. doi:10.3205/psm000089
70. Marcum ZA, Walker RL, Jones BL, et al. Patterns of antihypertensive and statin adherence prior to dementia: findings from the adult changes in thought study. *BMC Geriatr*. Feb 14 2019;19(1):41. doi:10.1186/s12877-019-1058-6
71. Winn AN, Fergestrom NM, Neuner JM. Using Group-based Trajectory Models and Propensity Score Weighting to Detect Heterogeneous Treatment Effects: The Case Study of Generic Hormonal Therapy for Women With Breast Cancer. *Med Care*. Jan 2019;57(1):85-93. doi:10.1097/mlr.0000000000001019
72. MacEwan JP, Silverstein AR, Shafrin J, Lakdawalla DN, Hatch A, Forma FM. Medication Adherence Patterns Among Patients with Multiple Serious Mental and Physical Illnesses. *Adv Ther*. May 2018;35(5):671-685. doi:10.1007/s12325-018-0700-6
73. Feldman CH, Collins J, Zhang Z, et al. Azathioprine and Mycophenolate Mofetil Adherence Patterns and Predictors among Medicaid Beneficiaries with Systemic Lupus Erythematosus. *Arthritis Care Res (Hoboken)*. Oct 24 2018;doi:10.1002/acr.23792
74. Franklin JM, Gopalakrishnan C, Krumme AA, et al. The relative benefits of claims and electronic health record data for predicting medication adherence trajectory. *Am Heart J*. Mar 2018;197:153-162. doi:10.1016/j.ahj.2017.09.019
75. MacEwan JP, Forma FM, Shafrin J, Hatch A, Lakdawalla DN, Lindenmayer JP. Patterns of Adherence to Oral Atypical Antipsychotics Among Patients Diagnosed with Schizophrenia. *J Manag Care Spec Pharm*. Nov 2016;22(11):1349-1361. doi:10.18553/jmcp.2016.22.11.1349
76. Winn AN, Dusetzina SB. The association between trajectories of endocrine therapy adherence and mortality among women with breast cancer. *Pharmacoepidemiol Drug Saf*. Aug 2016;25(8):953-9. doi:10.1002/pds.4012
77. Newman-Casey PA, Blachley T, Lee PP, Heisler M, Farris KB, Stein JD. Patterns of Glaucoma Medication Adherence over Four Years of Follow-Up. *Ophthalmology*. Oct 2015;122(10):2010-21. doi:10.1016/j.ophtha.2015.06.039
78. Juarez DT, Williams AE, Chen C, et al. Factors affecting medication adherence trajectories for patients with heart failure. *Am J Manag Care*. Mar 1 2015;21(3):e197-205.
79. Franklin JM, Krumme AA, Tong AY, et al. Association between trajectories of statin adherence and subsequent cardiovascular events. *Pharmacoepidemiol Drug Saf*. Oct 2015;24(10):1105-13. doi:10.1002/pds.3787

80. Wang CH, Huang LC, Yang CC, et al. Short- and long-term use of medication for psychological distress after the diagnosis of cancer. *Support Care Cancer*. Mar 2017;25(3):757-768. doi:10.1007/s00520-016-3456-z
81. Hickson RP, Annis IE, Killeya-Jones LA, Fang G. Opening the black box of the group-based trajectory modeling process to analyze medication adherence patterns: An example using real-world statin adherence data. *Pharmacoepidemiol Drug Saf*. Mar 2020;29(3):357-362. doi:10.1002/pds.4917
82. Vadhariya A, Fleming ML, Johnson ML, et al. Group-Based Trajectory Models to Identify Sociodemographic and Clinical Predictors of Adherence Patterns to Statin Therapy Among Older Adults. *Am Health Drug Benefits*. Jun-Jul 2019;12(4):202-211.
83. Hernandez I, He M, Chen N, Brooks MM, Saba S, Gellad WF. Trajectories of Oral Anticoagulation Adherence Among Medicare Beneficiaries Newly Diagnosed With Atrial Fibrillation. *J Am Heart Assoc*. Jun 18 2019;8(12):e011427. doi:10.1161/jaha.118.011427
84. Dillon P, Stewart D, Smith SM, Gallagher P, Cousins G. Group-Based Trajectory Models: Assessing Adherence to Antihypertensive Medication in Older Adults in a Community Pharmacy Setting. *Clin Pharmacol Ther*. Jun 2018;103(6):1052-1060. doi:10.1002/cpt.865
85. Dillon P, Smith SM, Gallagher P, Cousins G. The association between pharmacy refill-adherence metrics and healthcare utilisation: a prospective cohort study of older hypertensive adults. *Int J Pharm Pract*. Apr 10 2019;doi:10.1111/ijpp.12539
86. Aarnio E, Martikainen J, Winn AN, Huupponen R, Vahtera J, Korhonen MJ. Socioeconomic Inequalities in Statin Adherence Under Universal Coverage: Does Sex Matter? *Circ Cardiovasc Qual Outcomes*. Nov 2016;9(6):704-713. doi:10.1161/circoutcomes.116.002728
87. Mardby AC, Schioler L, Sundell KA, Bjerkeli P, Lesen E, Jonsson AK. Adherence to antidepressants among women and men described with trajectory models: a Swedish longitudinal study. *Eur J Clin Pharmacol*. Nov 2016;72(11):1381-1389. doi:10.1007/s00228-016-2106-1
88. Lambert-Cote L, Bouhnik AD, Bendiane MK, et al. Adherence trajectories of adjuvant endocrine therapy in the five years after its initiation among women with non-metastatic breast cancer: a cohort study using administrative databases. *Breast Cancer Res Treat*. Apr 2020;180(3):777-790. doi:10.1007/s10549-020-05549-x
89. Ajrouche A, Estellat C, De Rycke Y, Tubach F. Trajectories of Adherence to Low-Dose Aspirin Treatment Among the French Population. *J Cardiovasc Pharmacol Ther*. Jan 2020;25(1):37-46. doi:10.1177/1074248419865287
90. Chen CC, Cheng SH. Continuity of care and changes in medication adherence among patients with newly diagnosed diabetes. *Am J Manag Care*. Feb 2016;22(2):136-42.
91. Schaffer AL, Buckley NA, Pearson SA. Who benefits from fixed-dose combinations? Two-year statin adherence trajectories in initiators of combined amlodipine/atorvastatin therapy. *Pharmacoepidemiol Drug Saf*. Dec 2017;26(12):1465-1473. doi:10.1002/pds.4342
92. Tervonen HE, Daniels B, Tang M, Preen DB, Pearson SA. Patterns of endocrine therapy in a national cohort of early stage HER2-positive breast cancer patients. *Pharmacoepidemiol Drug Saf*. Jun 2019;28(6):812-820. doi:10.1002/pds.4751
93. Zongo A, Simpson S, Johnson JA, Eurich DT. Change in Trajectories of Adherence to Lipid-Lowering Drugs Following Non-Fatal Acute Coronary Syndrome or Stroke. *J Am Heart Assoc*. Dec 3 2019;8(23):e013857. doi:10.1161/jaha.119.013857

94. Feldman CH, Collins J, Zhang Z, et al. Dynamic patterns and predictors of hydroxychloroquine nonadherence among Medicaid beneficiaries with systemic lupus erythematosus. *Semin Arthritis Rheum*. Oct 2018;48(2):205-213. doi:10.1016/j.semarthrit.2018.01.002
95. Alhazami M, Pontinha VM, Patterson JA, Holdford DA. Medication Adherence Trajectories: A Systematic Literature Review. *J Manag Care Spec Pharm*. Sep 2020;26(9):1138-1152. doi:10.18553/jmcp.2020.26.9.1138
96. U.S. Centers for Medicare and Medicaid Services. Medicare 2020 Part C & D Star Ratings Technical Notes. 18.01.20, Accessed January 5, 2020, https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/Downloads/Star-Ratings-Technical-Notes-Oct-10-2019.pdf
97. Abughosh SM, Vadhariya A, Johnson ML, et al. Enhancing Statin Adherence Using a Motivational Interviewing Intervention and Past Adherence Trajectories in Patients with Suboptimal Adherence. *J Manag Care Spec Pharm*. Oct 2019;25(10):1053-1062. doi:10.18553/jmcp.2019.25.10.1053
98. Nagin DS, Jones BL, Passos VL, Tremblay RE. Group-based multi-trajectory modeling. *Stat Methods Med Res*. Jul 2018;27(7):2015-2023. doi:10.1177/0962280216673085
99. Juge PA, Granger B, Debray MP, et al. A Risk Score to Detect Subclinical Rheumatoid Arthritis-Associated Interstitial Lung Disease. *Arthritis Rheumatol*. May 18 2022;doi:10.1002/art.42162
100. Bredt LC, Peres LAB, Risso M, Barros L. Risk factors and prediction of acute kidney injury after liver transplantation: Logistic regression and artificial neural network approaches. *World J Hepatol*. Mar 27 2022;14(3):570-582. doi:10.4254/wjh.v14.i3.570
101. Campbell HM, Murata AE, Conner TA, Fotieo G. Chronic use of non-steroidal anti-inflammatory drugs (NSAIDs) or acetaminophen and relationship with mortality among United States Veterans after testing positive for COVID-19. *PLoS One*. 2022;17(5):e0267462. doi:10.1371/journal.pone.0267462
102. Li R, Zhan W, Huang X, et al. Association of Dietary Inflammatory Index (DII) and depression in the elderly over 55 years in Northern China: analysis of data from a multicentre, cohort study. *BMJ Open*. Apr 21 2022;12(4):e056019. doi:10.1136/bmjopen-2021-056019
103. Helle C, Hillesund ER, Overby NC. Associations between infant and maternal characteristics measured at child age 5 months and maternal feeding styles and practices up to child age two years. *PLoS One*. 2022;17(1):e0261222. doi:10.1371/journal.pone.0261222
104. Lee K, Ahn H, Moon H, Kodell RL, Chen JJ. Multinomial logistic regression ensembles. *J Biopharm Stat*. May 2013;23(3):681-94. doi:10.1080/10543406.2012.756500
105. Edrees H, Song W, Syrowatka A, Simona A, Amato MG, Bates DW. Intelligent Telehealth in Pharmacovigilance: A Future Perspective. *Drug Saf*. May 2022;45(5):449-458. doi:10.1007/s40264-022-01172-5
106. Moffat LM, Xu D. Accuracy of Machine Learning Models to Predict In-hospital Cardiac Arrest: A Systematic Review. *Clin Nurse Spec*. Jan-Feb 01 2022;36(1):29-44. doi:10.1097/NUR.0000000000000644

107. Ishwaran H, Lu M. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. *Stat Med*. Feb 20 2019;38(4):558-582. doi:10.1002/sim.7803
108. Deist TM, Dankers F, Valdes G, et al. Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers. *Med Phys*. Jul 2018;45(7):3449-3459. doi:10.1002/mp.12967
109. Couronne R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics*. Jul 17 2018;19(1):270. doi:10.1186/s12859-018-2264-5
110. Breiman L. Random forests. *Machine Learning*. Kluwer Academic Publishers; 2001:5-32. vol. 1.
111. Breiman L. Random Forests. *Machine Learning*. 2001/10/01 2001;45(1):5-32. doi:10.1023/A:1010933404324
112. Dietterich TG. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization. *Machine Learning*. 2000/08/01 2000;40(2):139-157. doi:10.1023/A:1007607513941
113. Freund Y. An Adaptive Version of the Boost by Majority Algorithm. *Machine Learning*. 2001:293-318. vol. 3.
114. Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. 2006:161-168.
115. Neumann A, Holstein J, Le Gall JR, Lepage E. Measuring performance in health care: case-mix adjustment by boosted decision trees. *Artif Intell Med*. Oct 2004;32(2):97-113. doi:10.1016/j.artmed.2004.06.001
116. Kuhn M. The caret package. Accessed 10/10/2020, https://topepo.github.io/caret/index.html
117. Kuhn M, Wing J, Weston S, et al. caret: Classification and Regression Training. Accessed 10/26/2020, https://cran.r-project.org/web/packages/caret/index.html
118. Fryar CD, Ostchega Y, Hales CM, Zhang G, Kruszon-Moran D. Hypertension Prevalence and Control Among Adults: United States, 2015-2016. *NCHS Data Brief*. Oct 2017;(289):1-8.
119. Carroll MD, Fryar CD, Nguyen DT. Total and High-density Lipoprotein Cholesterol in Adults: United States, 2015-2016. *NCHS Data Brief*. Oct 2017;(290):1-8.
120. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of Type 2 Diabetes - Global Burden of Disease and Forecasted Trends. *J Epidemiol Glob Health*. Mar 2020;10(1):107-111. doi:10.2991/jegh.k.191028.001
121. *SAS 9.4*. SAS Institute Inc.;
122. Jones BL, Nagin DS. A Note on a Stata Plugin for Estimating Group-based Trajectory Models. *Sociological Methods & Research*. 2013/11/01 2013;42(4):608-613. doi:10.1177/0049124113503141
123. *Stata Statistical Software: Release 17*. StataCorp LLC; 2021.
124. Kuhn M. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*. 11/10 2008;28(5):1 - 26. doi:10.18637/jss.v028.i05
125. Peterson AM, Nau DP, Cramer JA, Benner J, Gwadry-Sridhar F, Nichol M. A checklist for medication compliance and persistence studies using retrospective databases. *Value Health*. Jan-Feb 2007;10(1):3-12. doi:10.1111/j.1524-4733.2006.00139.x

126. Pednekar PP, Agh T, Malmenas M, et al. Methods for Measuring Multiple Medication Adherence: A Systematic Review-Report of the ISPOR Medication Adherence and Persistence Special Interest Group. *Value Health*. Feb 2019;22(2):139-156. doi:10.1016/j.jval.2018.08.006
127. Choudhry NK, Shrank WH, Levin RL, et al. Measuring concurrent adherence to multiple related medications. *Am J Manag Care*. Jul 2009;15(7):457-64.
128. Chu L, Kawatkar A, Gu A. A SAS® macro program to calculate medication adherence rate for single and multiple medication use. Lex Jansen SAS Conference Proceedings website. Accessed 01/04/2022, https://www.lexjansen.com/wuss/2011/hoc/Papers_Chu_L_74886.pdf
129. Bishop CM. *Pattern recognition and machine learning*. Information Science and Statistics. Springer; 2006.
130. Erosheva EA, Matsueda RL, Telesca D. Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond. *Annual Review of Statistics and Its Application*. 2014;1(1):301-332. doi:10.1146/annurev-statistics-022513-115701
131. Maddala GS. Qualitative and limited dependent variable models in econometrics. *Cambridge: Cambridge*. 1983;498
132. Hsiao C. Logit and probit models. *The Econometrics of Panel Data*. Springer; 1996:410-428.
133. Montgomery DC. *Introduction to linear regression analysis*. Fifth edition. ed. Wiley; 2013.
134. Kutner MH, Nachtsheim C, Neter J. *Applied linear regression models*. 4th / ed. vol Boston ; New York :. McGraw-Hill/Irwin; 2004.
135. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. vol 3. Springer; 2015.
136. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. *BMC Med Res Methodol*. Jun 20 2012;12:82. doi:10.1186/1471-2288-12-82
137. Loughran T, Nagin DS. Finite Sample Effects in Group-Based Trajectory Models. *Sociological Methods & Research*. 2006;35(2):250-278. doi:10.1177/0049124106292292
138. Herzog AR, Wallace RB. Measures of cognitive functioning in the AHEAD Study. *J Gerontol B Psychol Sci Soc Sci*. May 1997;52 Spec No:37-48. doi:10.1093/geronb/52b.special_issue.37
139. Crimmins EM, Kim JK, Langa KM, Weir DR. Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study. *J Gerontol B Psychol Sci Soc Sci*. Jul 2011;66 Suppl 1:i162-71. doi:10.1093/geronb/gbr048
140. Steffick DE. *Documentation of affective functioning measures in the Health and Retirement Study*. 2000. https://hrs.isr.umich.edu/sites/default/files/biblio/dr-005.pdf
141. Kaiser Family Foundation. A Snapshot of Sources of Coverage Among Medicare Beneficiaries in 2018. Accessed 05/28/2022, https://www.kff.org/medicare/issue-brief/a-snapshot-of-sources-of-coverage-among-medicare-beneficiaries-in-2018/
142. Stamler J, Vaccaro O, Neaton JD, Wentworth D. Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care*. Feb 1993;16(2):434-44. doi:10.2337/diacare.16.2.434

143. Ho JE, Paultre F, Mosca L, Women's Pooling P. Is diabetes mellitus a cardiovascular disease risk equivalent for fatal stroke in women? Data from the Women's Pooling Project. *Stroke*. Dec 2003;34(12):2812-6. doi:10.1161/01.STR.0000102901.41780.5C
144. Valensi P, Lorgis L, Cottin Y. Prevalence, incidence, predictive factors and prognosis of silent myocardial infarction: a review of the literature. *Arch Cardiovasc Dis*. Mar 2011;104(3):178-88. doi:10.1016/j.acvd.2010.11.013
145. Tinoco MS, Groia-Veloso RCS, Santos J, Cruzeiro MGM, Dias BM, Reis AMM. Medication regimen complexity of coronary artery disease patients. *Einstein (Sao Paulo)*. 2021;19:eAO5565. doi:10.31744/einstein_journal/2021AO5565
146. Rettig SM, Wood Y, Hirsch JD. Medication regimen complexity in patients with uncontrolled hypertension and/or diabetes. *J Am Pharm Assoc (2003)*. Jul-Aug 2013;53(4):427-31. doi:10.1331/JAPhA.2013.13003
147. Ingersoll KS, Cohen J. The impact of medication regimen factors on adherence to chronic treatment: a review of literature. *J Behav Med*. Jun 2008;31(3):213-24. doi:10.1007/s10865-007-9147-y
148. Rapsomaniki E, Timmis A, George J, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet*. May 31 2014;383(9932):1899-911. doi:10.1016/S0140-6736(14)60685-1
149. Brown MT, Bussell J, Dutta S, Davis K, Strong S, Mathew S. Medication Adherence: Truth and Consequences. *Am J Med Sci*. Apr 2016;351(4):387-99. doi:10.1016/j.amjms.2016.01.010
150. Ho PM, Bryson CL, Rumsfeld JS. Medication Adherence. *Circulation*. 2009;119(23):3028-3035. doi:10.1161/circulationaha.108.768986
151. Hickson RP, Annis IE, Killeya-Jones LA, Fang G. Comparing Continuous and Binary Group-based Trajectory Modeling Using Statin Medication Adherence Data. *Med Care*. Nov 1 2021;59(11):997-1005. doi:10.1097/MLR.0000000000001625
152. Oates GR, Juarez LD, Hansen B, Kiefe CI, Shikany JM. Social Risk Factors for Medication Nonadherence: Findings from the CARDIA Study. *Am J Health Behav*. Mar 1 2020;44(2):232-243. doi:10.5993/AJHB.44.2.10
153. Rolnick SJ, Pawloski PA, Hedblom BD, Asche SE, Bruzek RJ. Patient characteristics associated with medication adherence. *Clin Med Res*. Jun 2013;11(2):54-65. doi:10.3121/cmr.2013.1113
154. Sherbourne CD, Hays RD, Ordway L, DiMatteo MR, Kravitz RL. Antecedents of adherence to medical recommendations: results from the Medical Outcomes Study. *J Behav Med*. Oct 1992;15(5):447-68. doi:10.1007/BF00844941
155. Gerlach LB, Kavanagh J, Watkins D, Chiang C, Kim HM, Kales HC. With a little help from my friends?: racial and gender differences in the role of social support in later-life depression medication adherence. *Int Psychogeriatr*. Sep 2017;29(9):1485-1493. doi:10.1017/S104161021700076X
156. Trivedi RB, Bryson CL, Udris E, Au DH. The influence of informal caregivers on adherence in COPD patients. *Ann Behav Med*. Aug 2012;44(1):66-72. doi:10.1007/s12160-012-9355-8
157. Mizokami F, Mase H, Kinoshita T, Kumagai T, Furuta K, Ito K. Adherence to Medication Regimens is an Effective Indicator of Cognitive Dysfunction in Elderly

Individuals. *Am J Alzheimers Dis Other Demen*. Mar 2016;31(2):132-6. doi:10.1177/1533317515598859
158. Kuzuya M, Hirakawa Y, Suzuki Y, et al. Association between unmet needs for medication support and all-cause hospitalization in community-dwelling disabled elderly people. *J Am Geriatr Soc*. May 2008;56(5):881-6. doi:10.1111/j.1532-5415.2008.01676.x
159. Wong AKC, Wong FKY, Chang K. Effectiveness of a community-based self-care promoting program for community-dwelling older adults: a randomized controlled trial. *Age Ageing*. Nov 1 2019;48(6):852-858. doi:10.1093/ageing/afz095
160. Gu D, Shen C. Assessing the Importance of Factors Associated with Cost-Related Nonadherence to Medication for Older US Medicare Beneficiaries. *Drugs Aging*. Dec 2019;36(12):1111-1121. doi:10.1007/s40266-019-00715-3
161. Health and Retirement Study Survey Research Center. HRS 2016 Final Release - Codebook. Accessed 01/15/2022, https://hrs.isr.umich.edu/sites/default/files/meta/2016/core/codebook/h16_00.html?_ga=2.158585854.539140789.1654381851-904238842.1638394310
162. Health and Retirement Study Survey Research Center. Data Collection Path Diagram. Institute for Social Research University of Michigan,. Accessed 01/15/2022, https://hrs.isr.umich.edu/data-products/collection-path
163. Krousel-Wood M, Craig LS, Peacock E, et al. Medication Adherence: Expanding the Conceptual Framework. *Am J Hypertens*. Sep 22 2021;34(9):895-909. doi:10.1093/ajh/hpab046
164. Chen C, Liaw A, Breiman L. *Using Random Forest to Learn Imbalanced Data*. 2004. 666. Accessed 06/04/2022. https://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf
165. Liu XY, Wu J, Zhou ZH. Exploratory Undersampling for Class-Imbalance Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*. 2009;39(2):539-550. doi:10.1109/TSMCB.2008.2007853
166. Chawla NV, Lazarevic A, Hall LO, Bowyer KW. SMOTEBoost: Improving Prediction of the Minority Class in Boosting. Springer Berlin Heidelberg; 2003:107-119.
167. Pontinha VM. Machine Learning Methods Session Draw Large Crowds at ISPOR 2022. *Value & Outcomes Spotlight*. 2022;8(3)(1)
168. Parekh N, Munshi KD, Hernandez I, Gellad WF, Henderson R, Shrank WH. Impact of Star Rating Medication Adherence Measures on Adherence for Targeted and Nontargeted Medications. *Value Health*. Nov 2019;22(11):1266-1274. doi:10.1016/j.jval.2019.06.009
169. Axon DR, Vaffis S, Chinthammit C, et al. Assessing the association between medication adherence, as defined in quality measures, and disease-state control, health care utilization, and costs in a retrospective database analysis of Medicare supplemental beneficiaries using statin medications. *J Manag Care Spec Pharm*. Dec 2020;26(12):1529-1537. doi:10.18553/jmcp.2020.26.12.1529
170. Campbell PJ, Axon DR, Taylor AM, et al. Associations of Renin-Angiotensin System Antagonist Medication Adherence and Economic Outcomes Among Commercially Insured US Adults: A Retrospective Cohort Study. *J Am Heart Assoc*. Sep 2020;9(17):e016094. doi:10.1161/JAHA.119.016094
171. Chinthammit C, Axon DR, Anderson S, et al. A retrospective cohort study evaluating the relationship between statin medication adherence and economic outcomes in

commercial health plans. *J Clin Lipidol*. Nov - Dec 2020;14(6):791-798. doi:10.1016/j.jacl.2020.09.005
172. Chinthammit C, Axon DR, Mollon L, et al. Evaluating the relationship between quality measure adherence definitions and economic outcomes in commercial health plans: a retrospective diabetes cohort study. *J Manag Care Spec Pharm*. Jan 2021;27(1):64-72. doi:10.18553/jmcp.2021.27.1.064
173. Menditto E, Cahir C, Malo S, et al. Persistence as a Robust Indicator of Medication Adherence-Related Quality and Performance. *Int J Environ Res Public Health*. May 3 2021;18(9)doi:10.3390/ijerph18094872
174. Vrijens B, De Geest S, Hughes DA, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol*. May 2012;73(5):691-705. doi:10.1111/j.1365-2125.2012.04167.x
175. The Office of the National Coordinator for Health Information Technology. Strategy on Reduction Regulatory and Administrative Burden Relating to the Use of Health IT and EHRs. Accessed 06/01/2022, https://www.healthit.gov/sites/default/files/page/2018-11/Draft%20Strategy%20on%20Reducing%20Regulatory%20and%20Administrative%20Burden%20Relating.pdf
176. Farley JF, Kumar A, Urick BY. Measuring Adherence: A Proof of Concept Study for Multiple Medications for Chronic Conditions in Alternative Payment Models. *Pharmacy (Basel)*. Jul 2 2019;7(3)doi:10.3390/pharmacy7030081
177. Health and Retirement Study Survey Research Center. HRS Medicare Claims and Summary Data Cross-Reference. Accessed 10/05/2021, https://hrs.isr.umich.edu/data-products/restricted-data/available-products/9677
178. Haviland A, Nagin DS, Rosenbaum PR. Combining propensity score matching and group-based trajectory analysis in an observational study. *Psychol Methods*. Sep 2007;12(3):247-67. doi:10.1037/1082-989X.12.3.247

8. APPENDICES

APPENDIX 1

Table 26 - List of drug names and drug classes included in the MA-PD and PDP Star Ratings Quality Metrics

Indication	Drug Class	Drug Name
Diabetes Medications	BIGUANIDES	METFORMIN
	DPP-4 INHIBITORS	SAXAGLIPTIN
	DPP-4 INHIBITORS	SITAGLIPTIN
	DPP-4 INHIBITORS	LINAGLIPTIN
	DPP-4 INHIBITORS	ALOGLIPTIN
	GLP-1 RECEPTOR AGONISTS	DULAGLUTIDE
	GLP-1 RECEPTOR AGONISTS	LIXISENATIDE
	GLP-1 RECEPTOR AGONISTS	LIRAGLUTIDE
	GLP-1 RECEPTOR AGONISTS	SEMAGLUTIDE
	GLP-1 RECEPTOR AGONISTS	ALBIGLUTIDE
	GLP-1 RECEPTOR AGONISTS	EXENATIDE
	MEGLITINIDES	NATEGLINIDE
	MEGLITINIDES	REPAGLINIDE
	SODIUM GLUCOSE CO-TRANSPORTER2 (SGLT2) INHIBITORS	ERTUGLIFLOZIN
	SODIUM GLUCOSE CO-TRANSPORTER2 (SGLT2) INHIBITORS	DAPAGLIFLOZIN
	SODIUM GLUCOSE CO-TRANSPORTER2 (SGLT2) INHIBITORS	EMPAGLIFLOZIN
	SODIUM GLUCOSE CO-TRANSPORTER2 (SGLT2) INHIBITORS	CANAGLIFLOZIN
	SULFONYLUREAS	GLIMEPIRIDE
	SULFONYLUREAS	TOLAZAMIDE
	SULFONYLUREAS	GLYBURIDE
	SULFONYLUREAS	GLIPIZIDE
	SULFONYLUREAS	CHLORPROPAMIDE
	SULFONYLUREAS	TOLBUTAMIDE
	THIAZOLIDINEDIONES	ROSIGLITAZONE
	THIAZOLIDINEDIONES	PIOGLITAZONE
Hypertension Medications	ACE INHIBITOR MEDICATIONS	LISINOPRIL
	ACE INHIBITOR MEDICATIONS	RAMIPRIL
	ACE INHIBITOR MEDICATIONS	PERINDOPRIL
	ACE INHIBITOR MEDICATIONS	QUINAPRIL
	ACE INHIBITOR MEDICATIONS	TRANDOLAPRIL
	ACE INHIBITOR MEDICATIONS	BENAZEPRIL
	ACE INHIBITOR MEDICATIONS	MOEXIPRIL
	ACE INHIBITOR MEDICATIONS	ENALAPRIL
	ACE INHIBITOR MEDICATIONS	FOSINOPRIL
	ACE INHIBITOR MEDICATIONS	CAPTOPRIL

Indication	Drug Class	Drug Name
	ARB MEDICATIONS	LOSARTAN
	ARB MEDICATIONS	VALSARTAN
	ARB MEDICATIONS	IRBESARTAN
	ARB MEDICATIONS	TELMISARTAN
	ARB MEDICATIONS	OLMESARTAN
	ARB MEDICATIONS	CANDESARTAN
	ARB MEDICATIONS	EPROSARTAN
	ARB MEDICATIONS	AZILSARTAN
	DIRECT RENIN INHIBITOR MEDICATIONS	ALISKIREN
High Blood Cholesterol Medications	STATINS	ATORVASTATIN
	STATINS	FLUVASTATIN
	STATINS	LOVASTATIN
	STATINS	PITAVASTATIN
	STATINS	PRAVASTATIN
	STATINS	ROSUVASTATIN
	STATINS	SIMVASTATIN

APPENDIX 2

Hypertension

Scenario	Foreign born	Non-white	Hispanic	Unmarried	Not college educated	Below poverty index	Medicaid beneficiary	CHAMPUS VA Beneficiary	Mobile home	Not home owner	Living nursing home	Excessive weight	Cognitive impairment	Cancer survivor	Stroke survivor	Heart condition present	Clinical depression
No risks	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Foreign born	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Non-white	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hispanic	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Unmarried	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Not college educated	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Below poverty index	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Medicaid beneficiary	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
CHAMPUSVA Beneficiary	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Mobile home	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Not home owner	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Living nursing home	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Excessive weight	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Cognitive impairment	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Cancer survivor	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Stroke survivor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Heart condition present	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Clinical depression	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
All risks	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 37 – Scenarios for imputing probabilities of trajectory membership in the hypertension medications trajectory model

Statins

Scenario	Female	Foreign born	Non-white	Hispanic	Unmarried	Not college educated	Below poverty index	Medicaid beneficiary	CHAMPUS VA Beneficiary	Mobile home	Not home owner	Living nursing home	Excessive weight	Cognitive impairment	Cancer survivor	Stroke survivor	Heart condition present	Clinical depression
No risks	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Female	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Foreign born	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Non-white	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hispanic	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Unmarried	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Not college educated	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Below poverty index	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Medicaid beneficiary	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
CHAMPUSVA Beneficiary	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Mobile home	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Not home owner	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Living nursing home	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Excessive weight	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Cognitive impairment	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Cancer survivor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Stroke survivor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Heart condition present	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Clinical depression	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
All risks	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 38 - Scenarios for imputing probabilities of trajectory membership in the statins trajectory model

Diabetes Medications																
Scenario	Female	Foreign born	Non-white	Hispanic	Unmarried	Not college educated	Below poverty index	Medicaid beneficiary	Not home owner	Living nursing home	Excessive weight	Cognitive impairment	Cancer survivor	Stroke survivor	Heart problems survivor	Clinical depression
No risks	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Female	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Foreign born	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Non-white	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Hispanic	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Unmarried	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Not college educated	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Below poverty index	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Medicaid beneficiary	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Not home owner	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Living nursing home	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Excessive weight	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Cognitive impairment	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Cancer survivor	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Stroke survivor	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Heart problems survivor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Clinical depression	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
All risks	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 39 - Scenarios for imputing probabilities of trajectory membership in the diabetes medications trajectory model

APPENDIX 3

RANDOM FOREST ALGORITHM

Model Accuracies

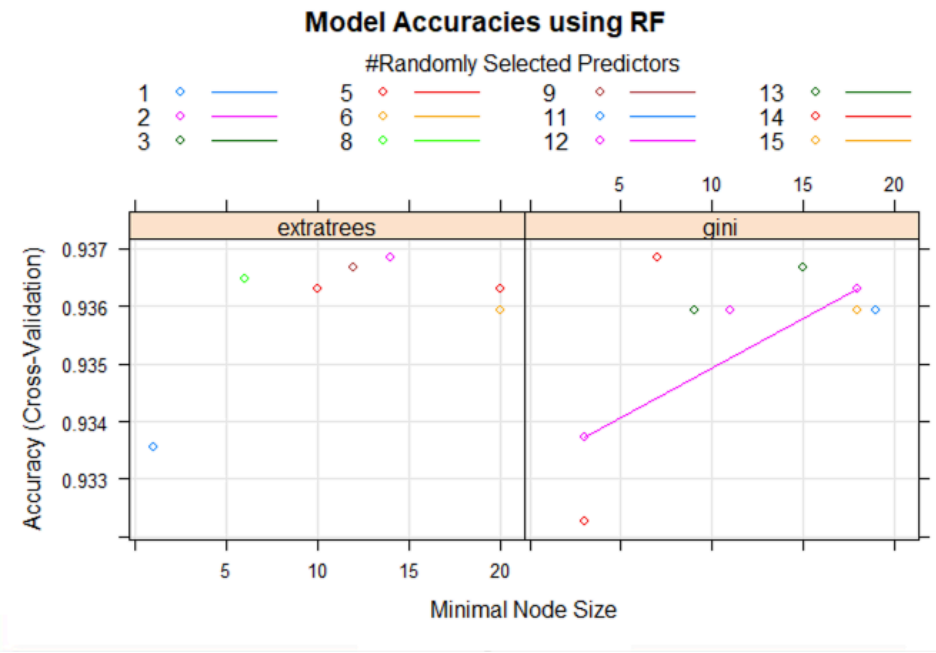

Figure 40 - Model accuracies for hypertension trajectory model and MI outcome

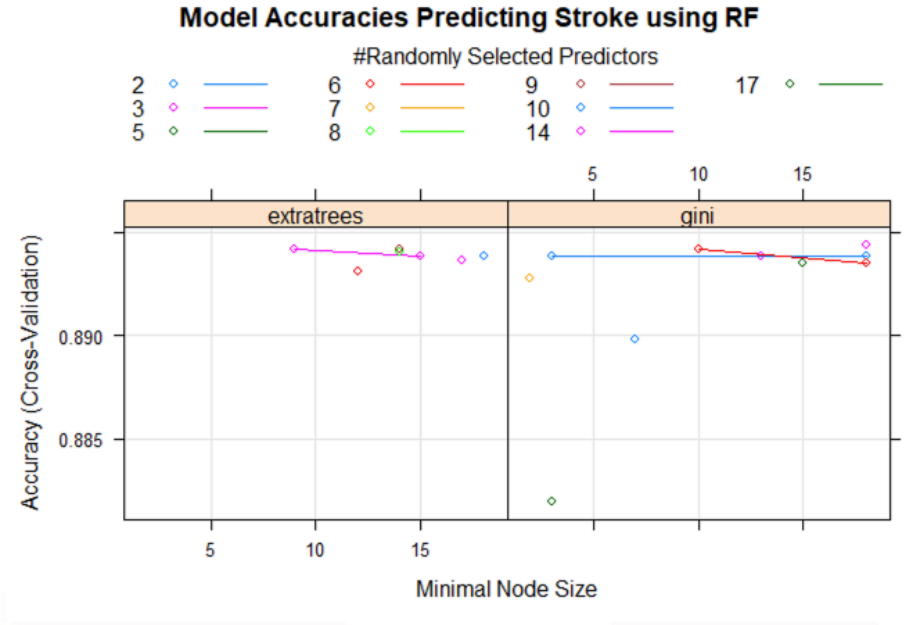

Figure 41 - Model accuracies for hypertension trajectory model and stroke outcome

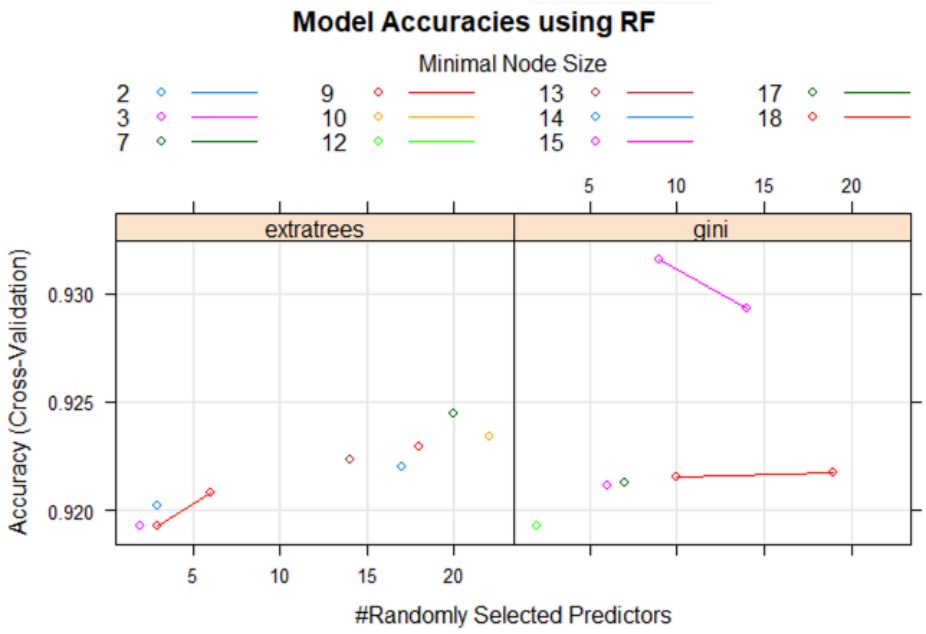

Figure 42 - Model accuracies for statins trajectory model and MI outcome

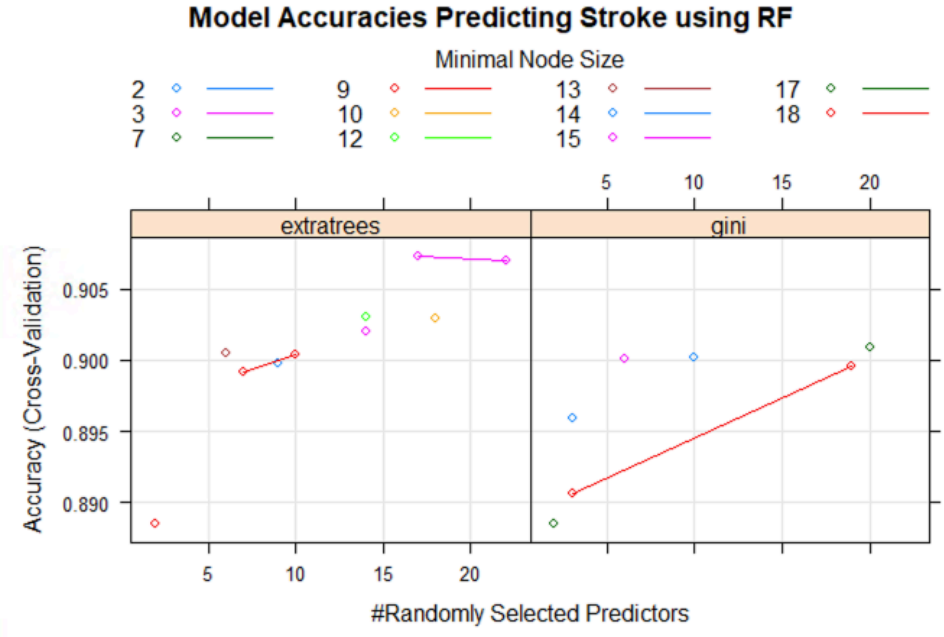

Figure 43 - Model accuracies for statins trajectory model and stroke outcome

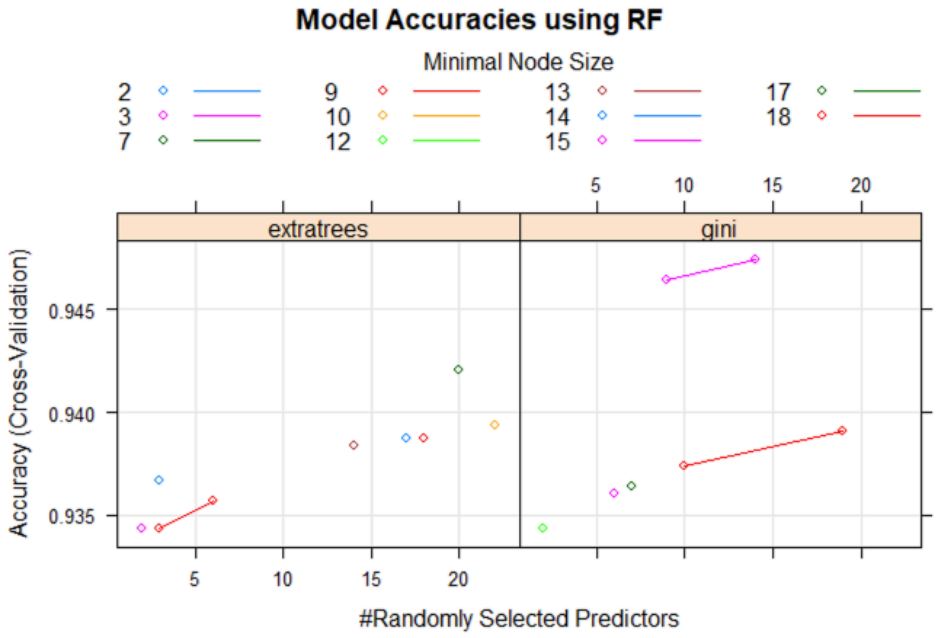

Figure 44 - Model accuracies for diabetes trajectory model and MI outcome

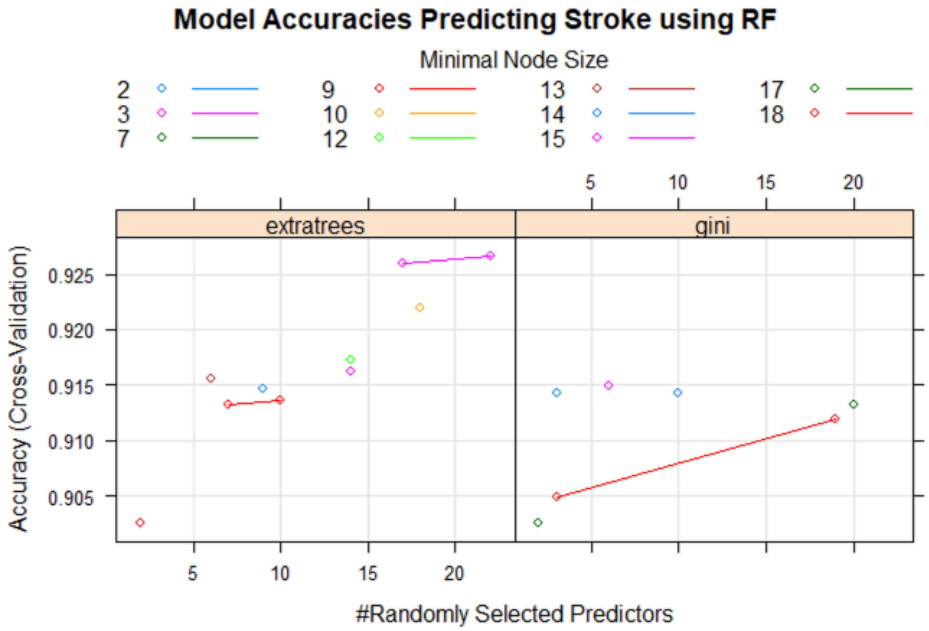

Figure 45 - Model accuracies for diabetes trajectory model and stroke outcome

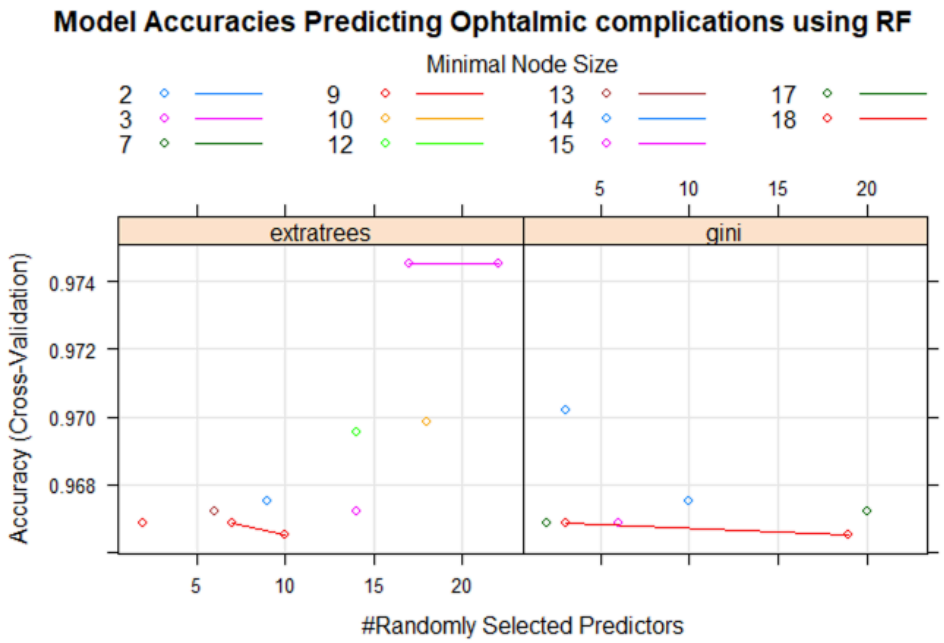

Figure 46 - Model accuracies for diabetes trajectory model and ophthalmic complications outcome

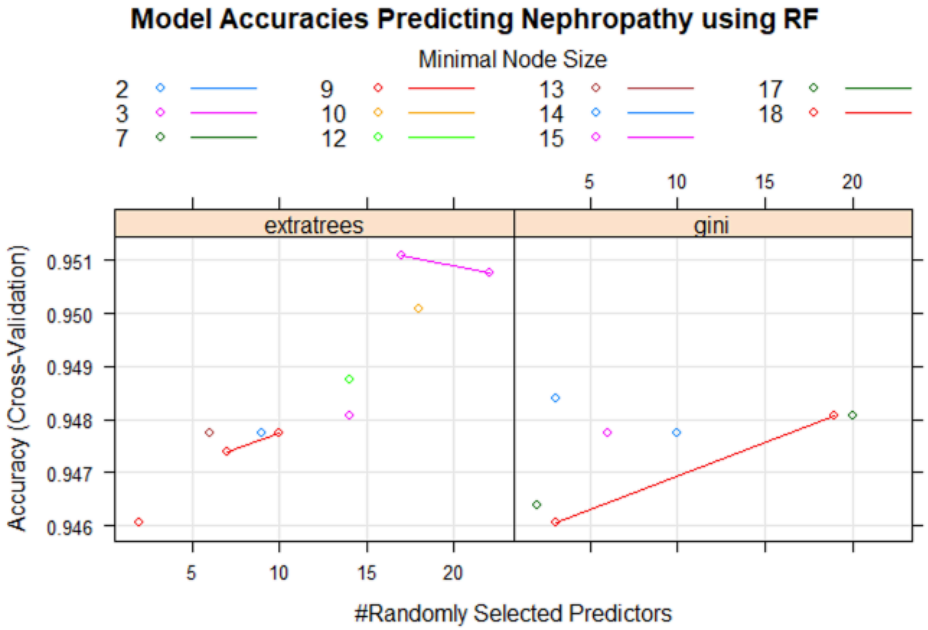

Figure 47 - Model accuracies for diabetes trajectory model and nephropathy complications outcome

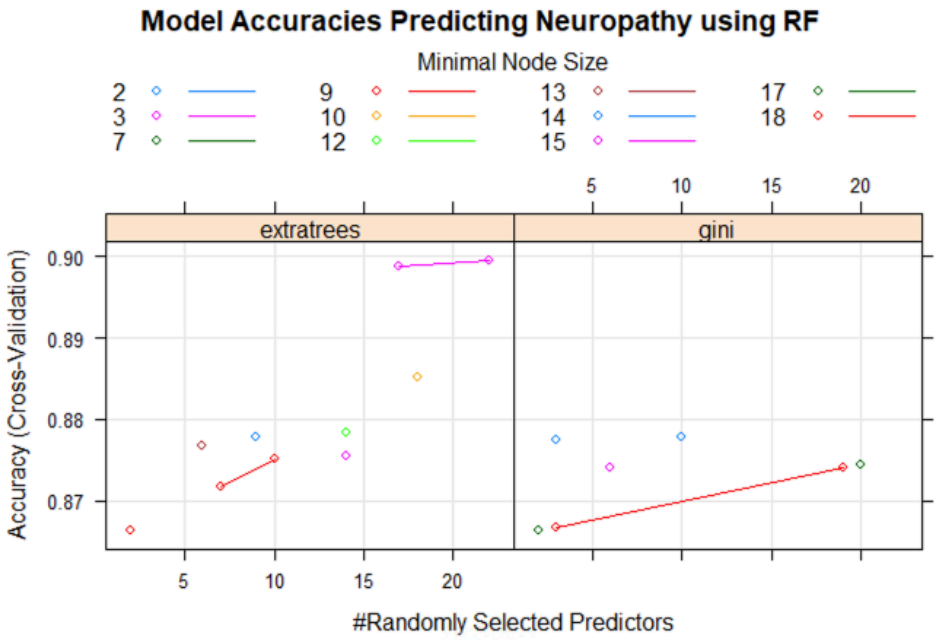

Figure 48 - Model accuracies for diabetes trajectory model and neuropathy complications outcome

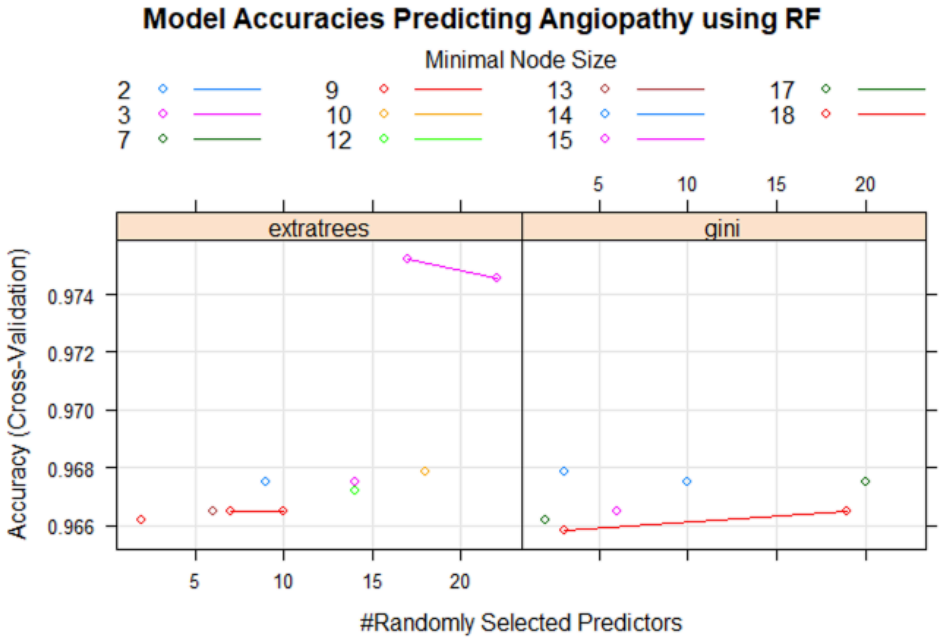

Figure 49 - Model accuracies for diabetes trajectory model and peripheral angiopathy complications outcome

Variable Importance Plots

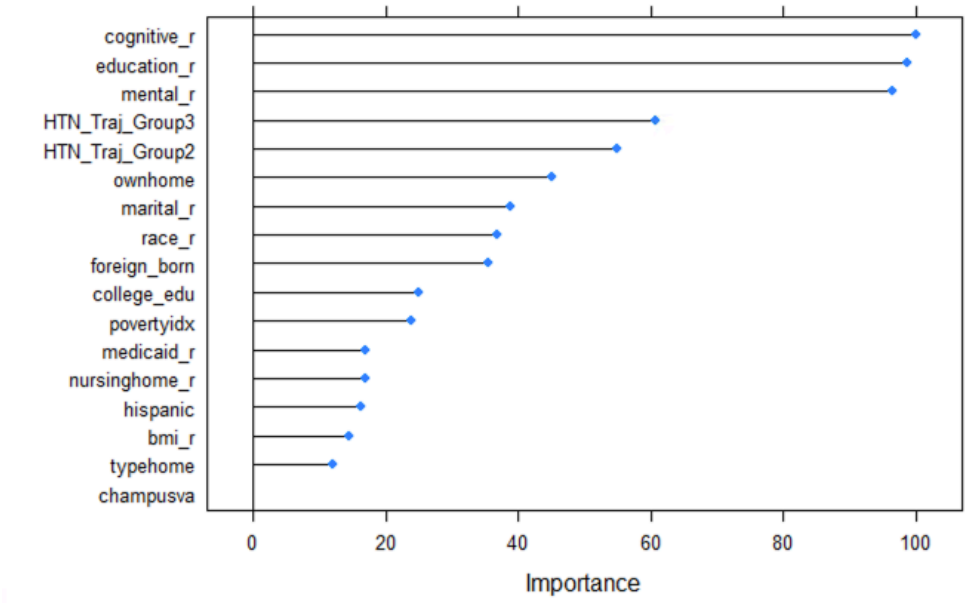

Figure 50 - Variable importance plot for the hypertension model and MI outcome

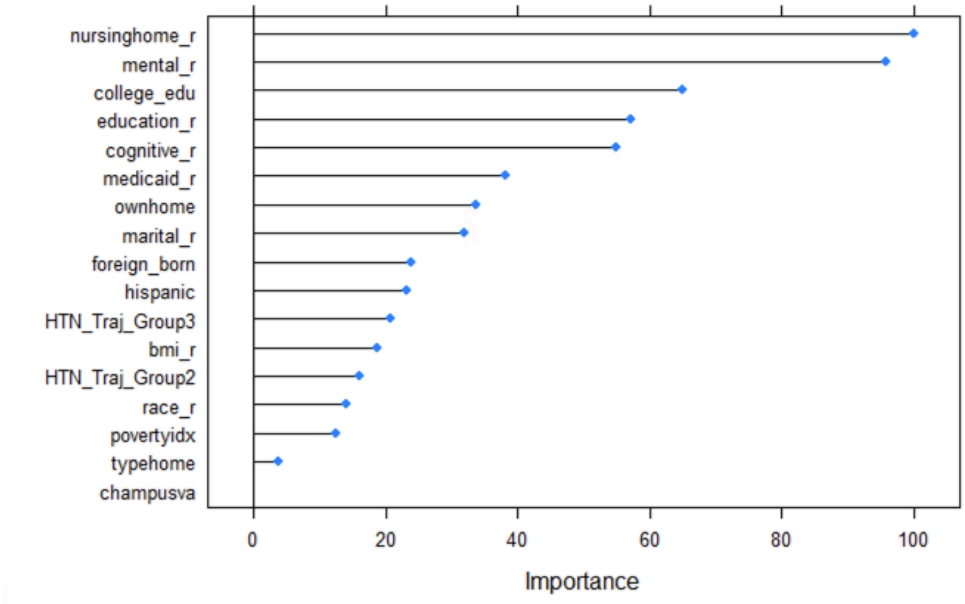

Figure 51 - Variable importance plot for the hypertension model and stroke outcome

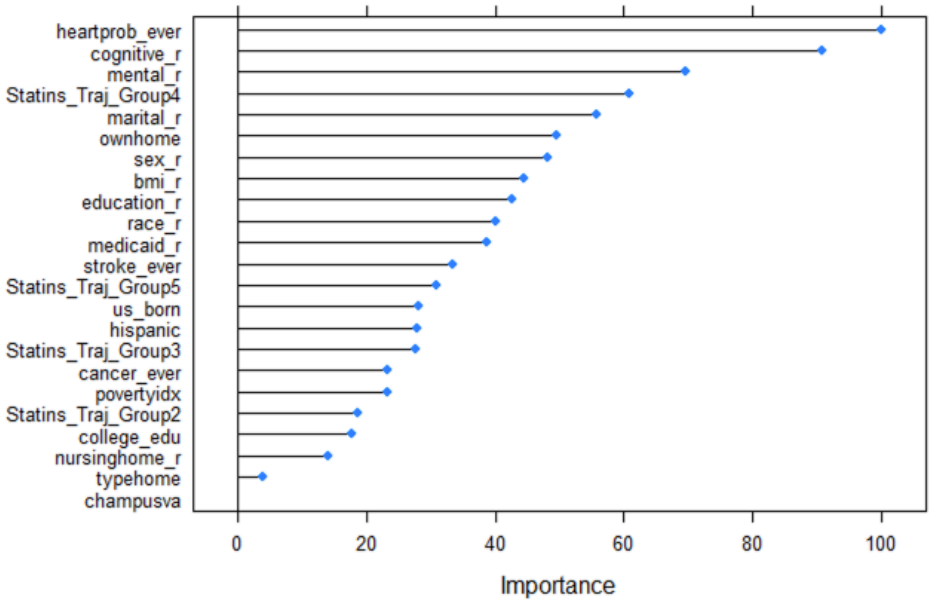

Figure 52 - Variable importance plot for the statins model and MI outcome

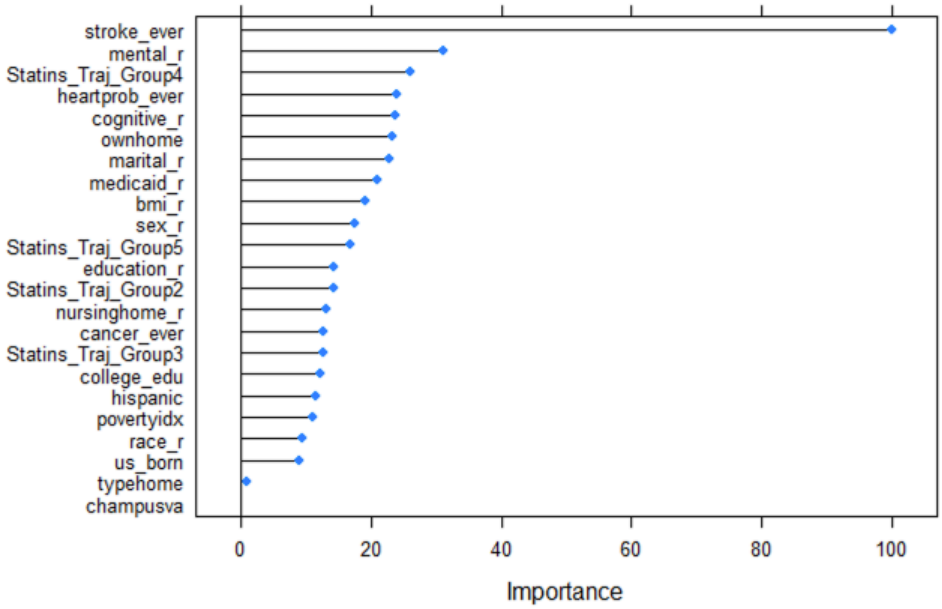

Figure 53 - Variable importance plot for the statins model and stroke outcome

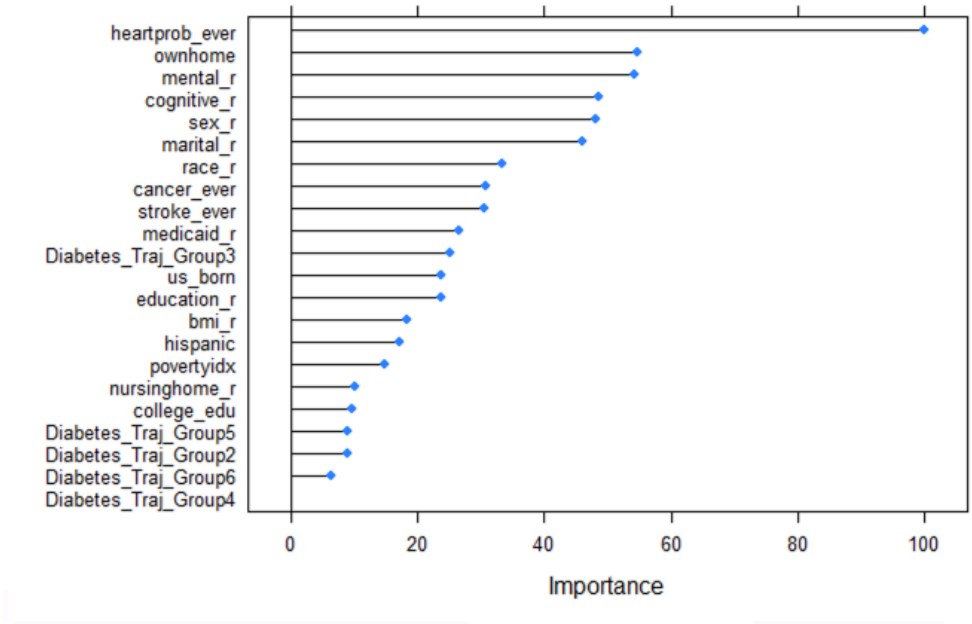

Figure 54 - Variable importance plot for the diabetes model and MI outcome

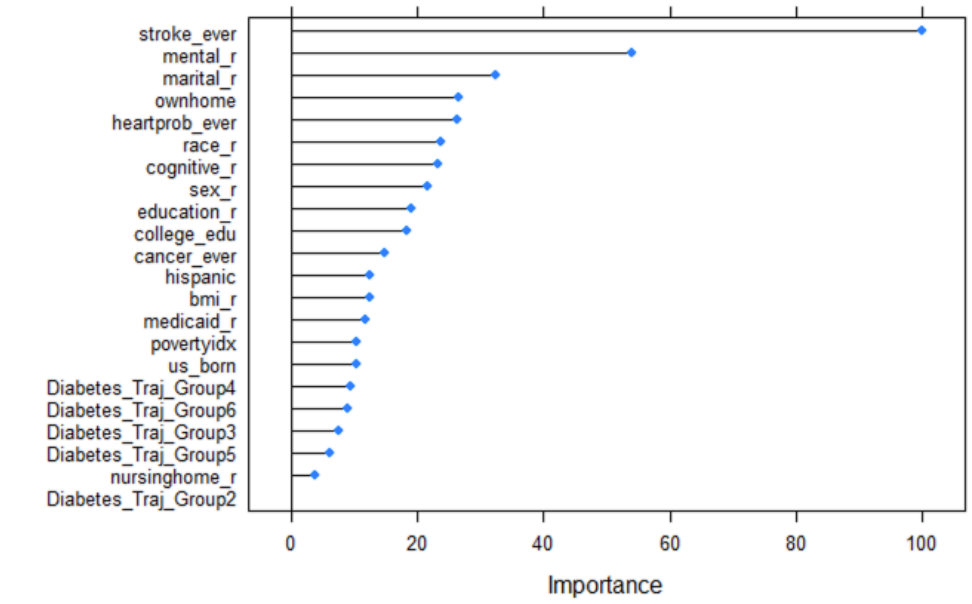

Figure 55 - Variable importance plot for the diabetes model and stroke outcome

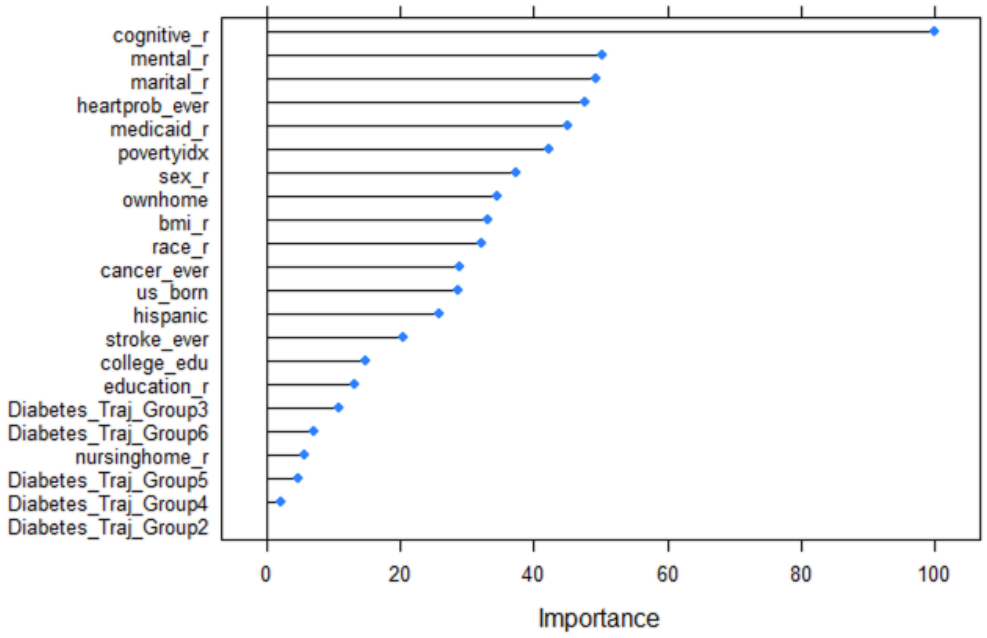

Figure 56 - Variable importance plot for the diabetes model and ophthalmic complications outcome

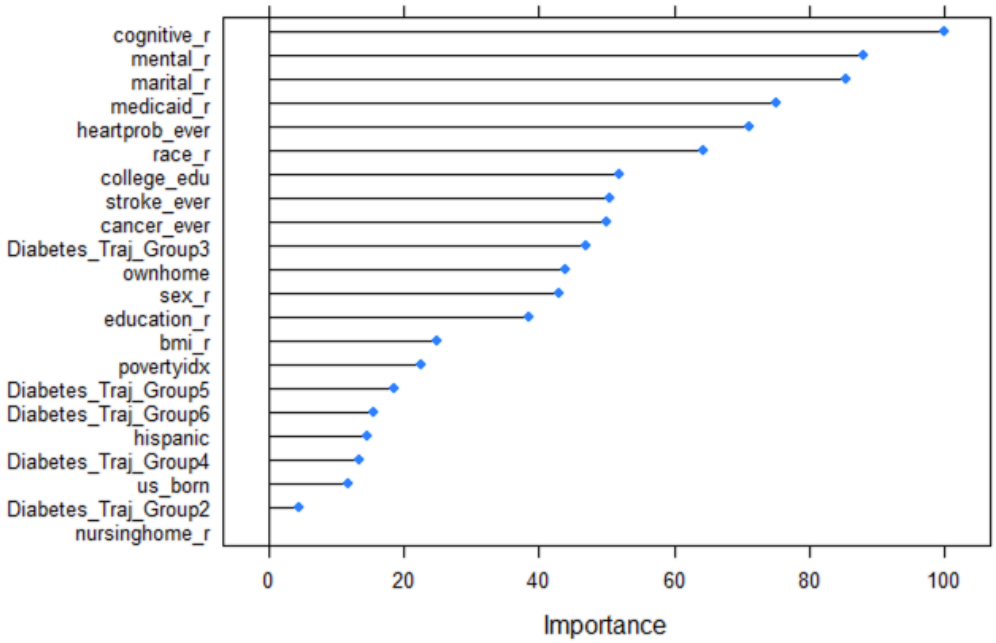

Figure 57 - Variable importance plot for the diabetes model and nephropathy complications outcome

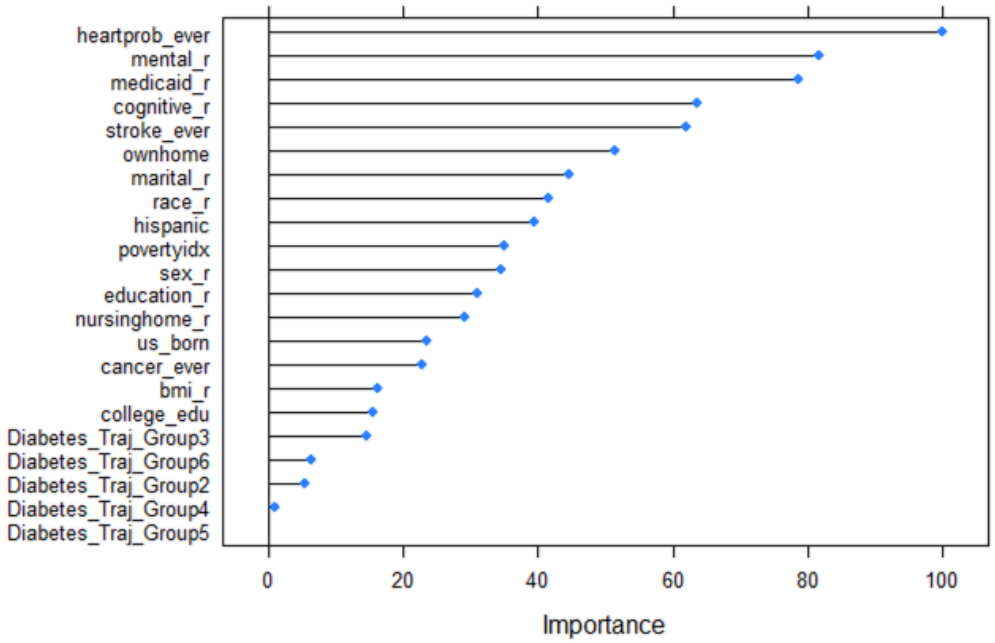

Figure 58 - Variable importance plot for the diabetes model and neuropathy complications outcome

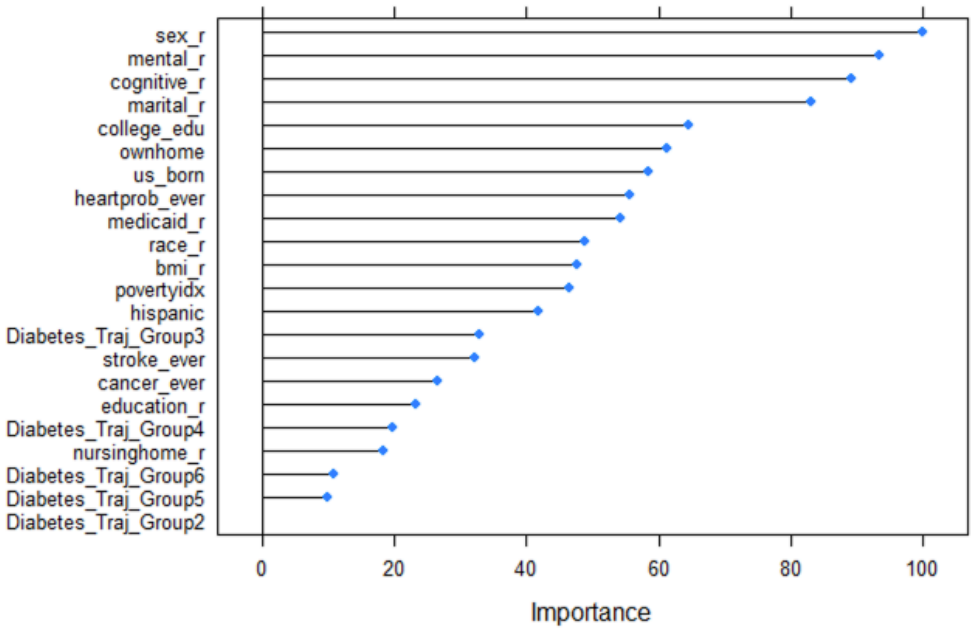

Figure 59 - Variable importance plot for the diabetes model and peripheral angiopathy complications outcome

EXTREME GRADIENT BOOSTING

Learning rates

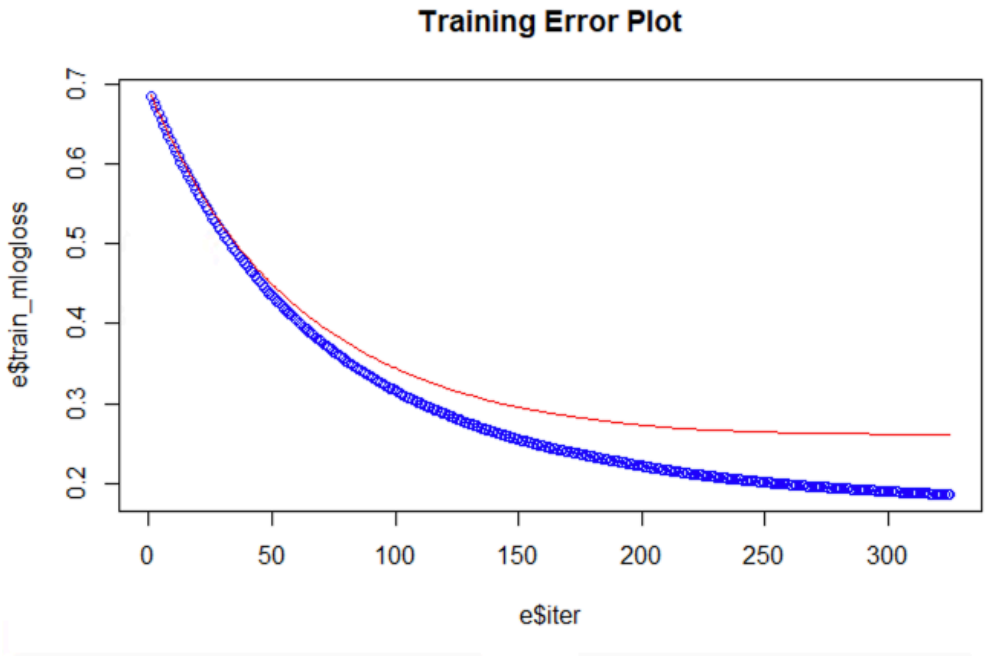

Figure 60 - Learning rate plot of hypertension model and MI outcome

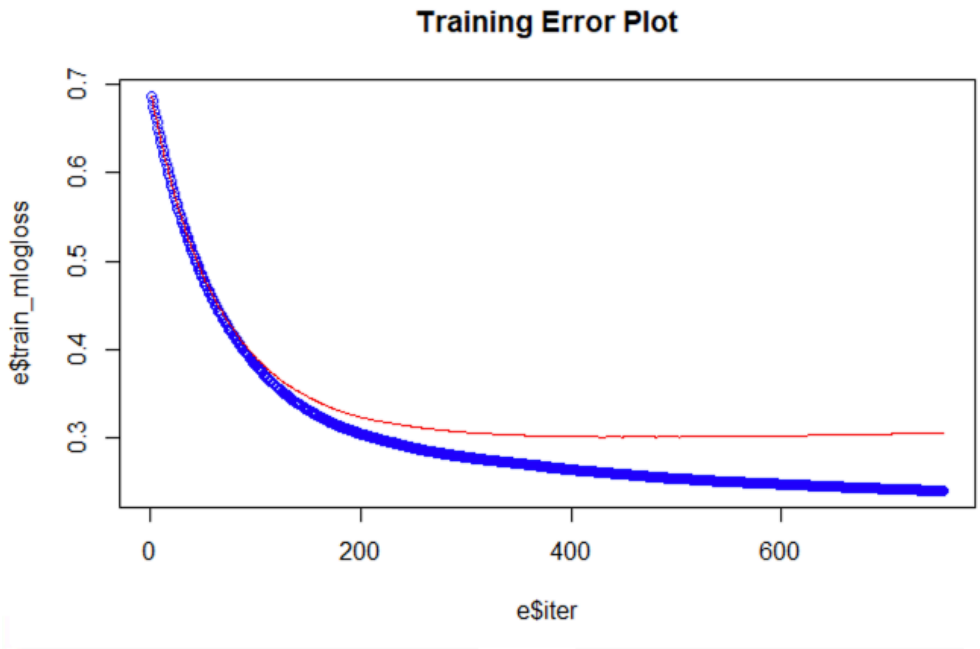

Figure 61 - Learning rate plot of hypertension model and stroke outcome

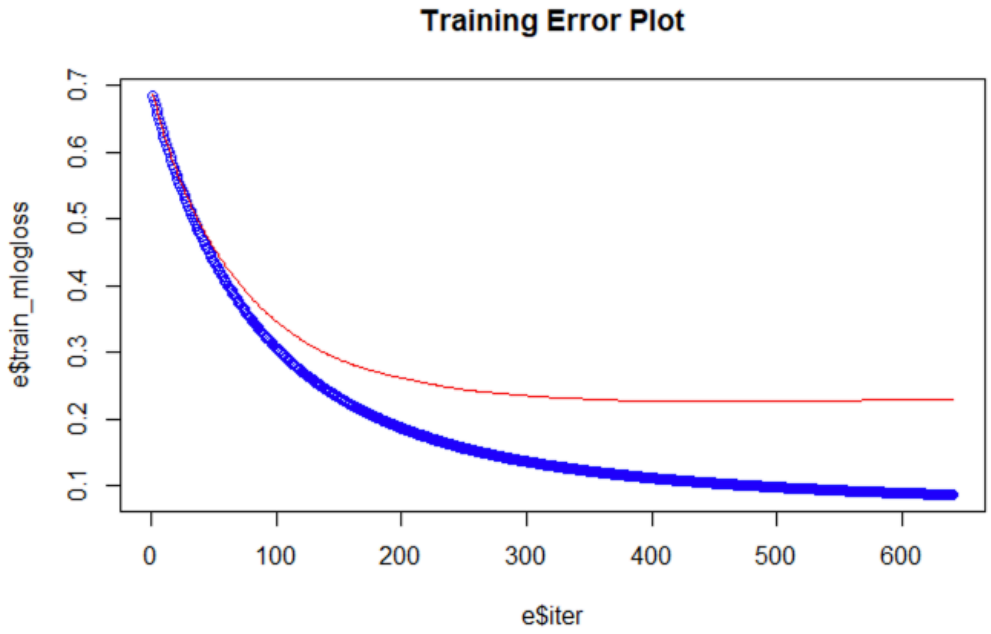

Figure 62 - Learning rate plot of statins model and MI outcome

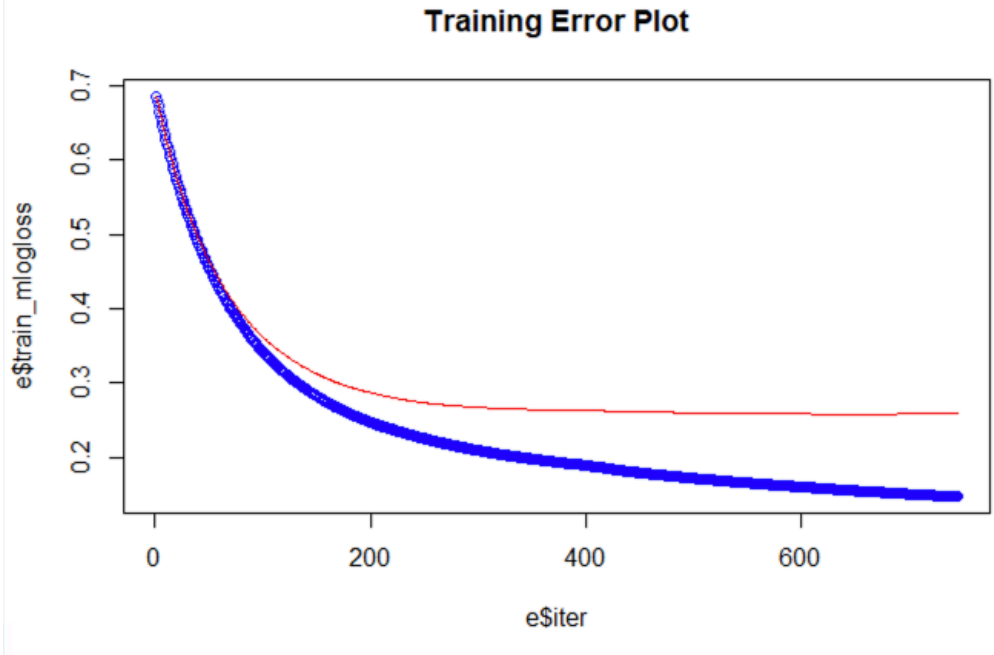

Figure 63 - Learning rate plot of statins model and stroke outcome

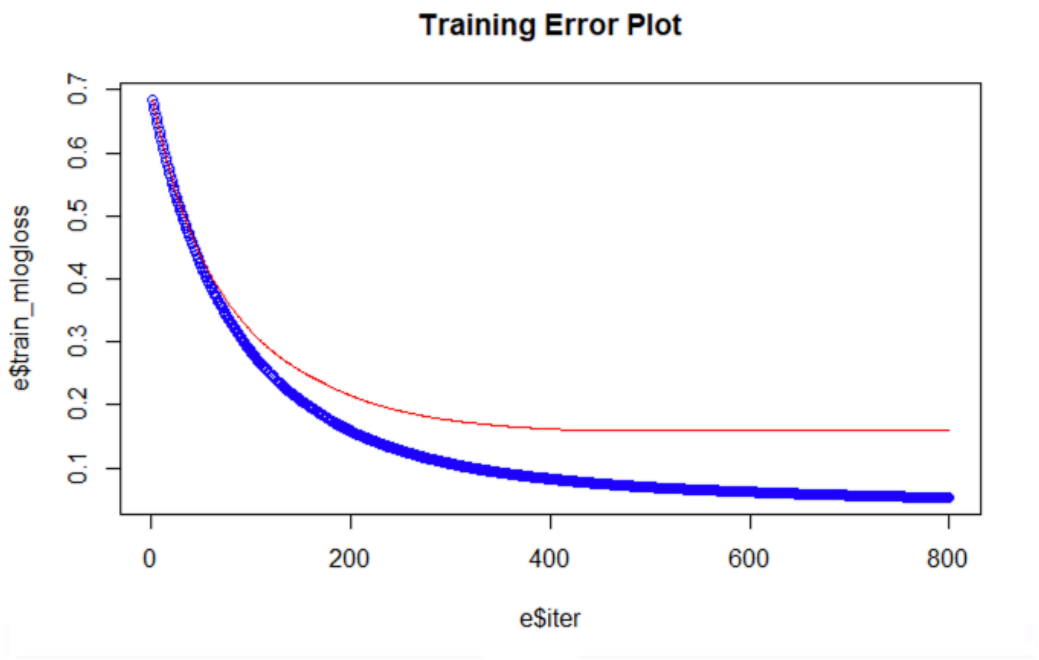

Figure 64 - Learning rate plot of diabetes model and MI outcome

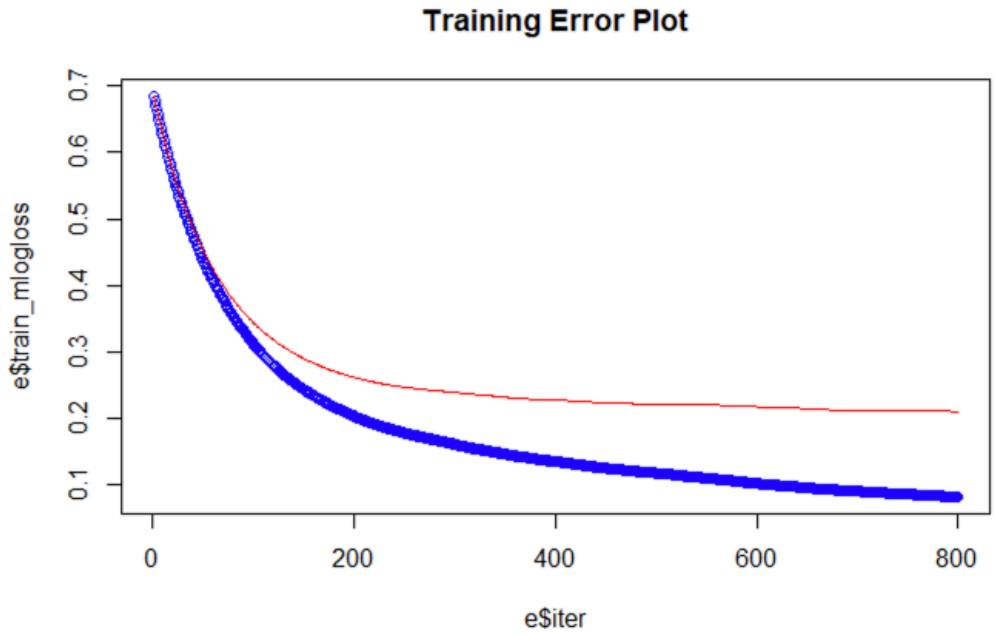

Figure 65 - Learning rate plot of diabetes model and stroke outcome

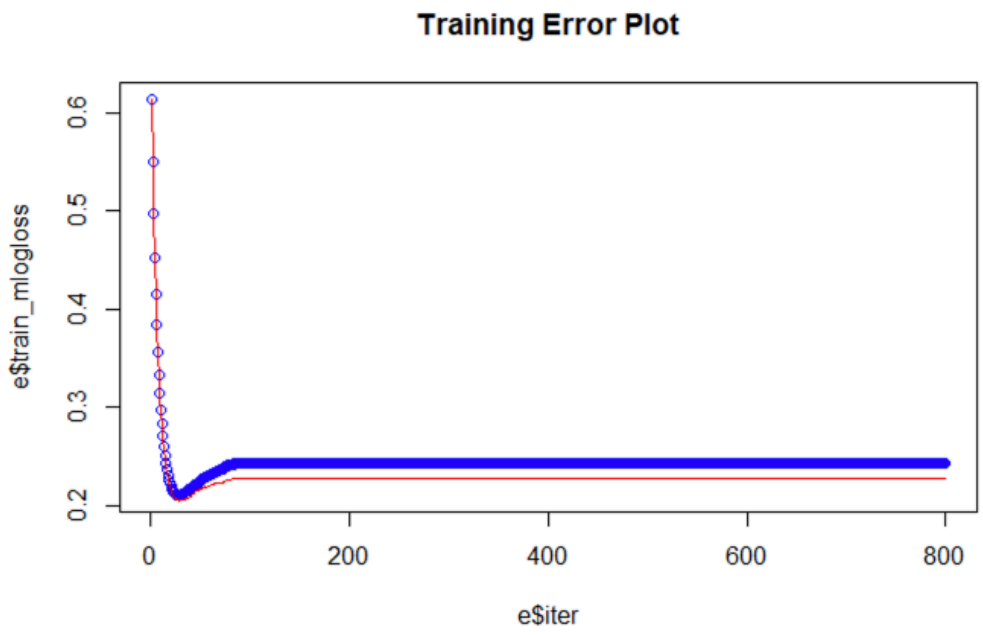

Figure 66 - Learning rate plot of diabetes model and ophthalmic complications outcome

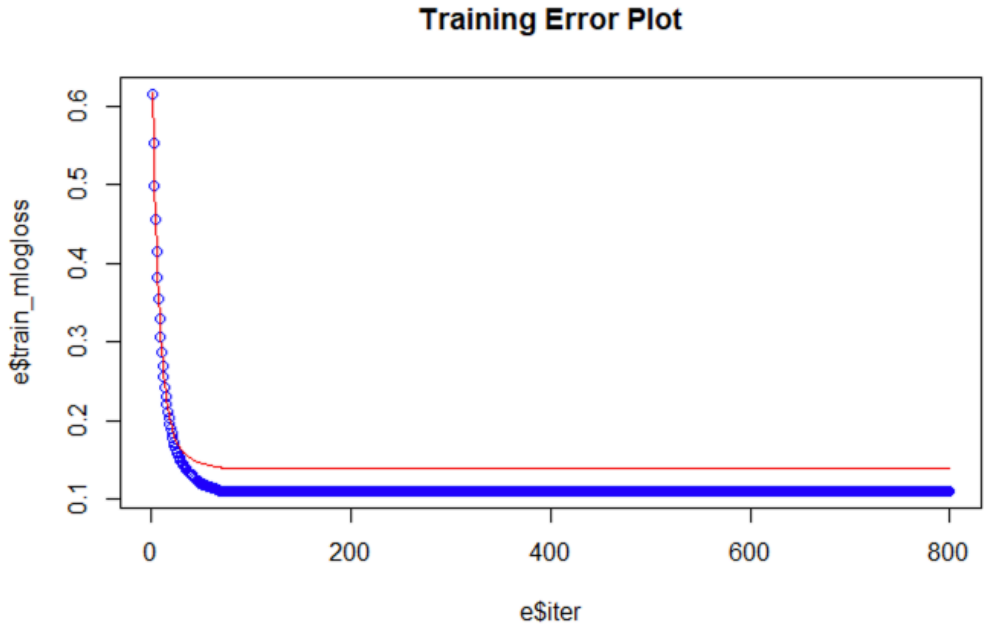

Figure 67 - Learning rate plot of diabetes model and nephropathy complications outcome

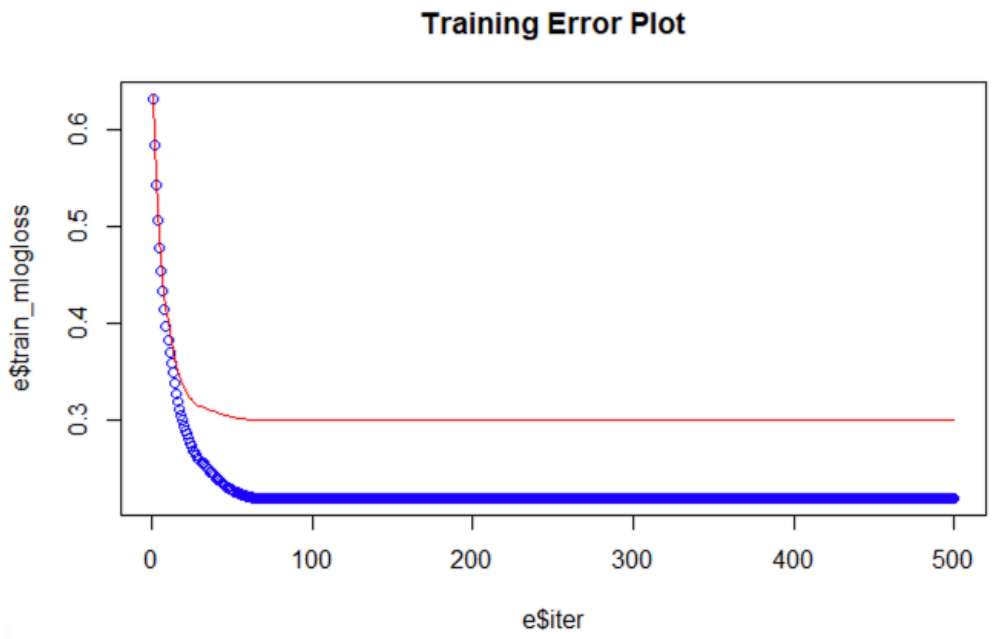

Figure 68 - Learning rate plot of diabetes model and neuropathy complications outcome

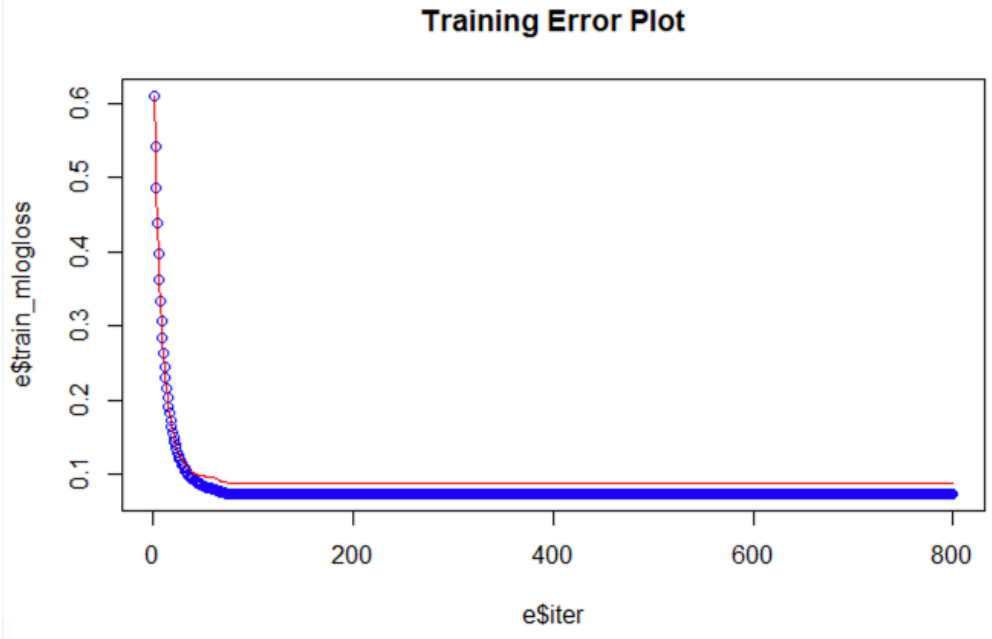

Figure 69 - Learning rate plot of diabetes model and peripheral angiopathy complications outcome

Variable importance plots

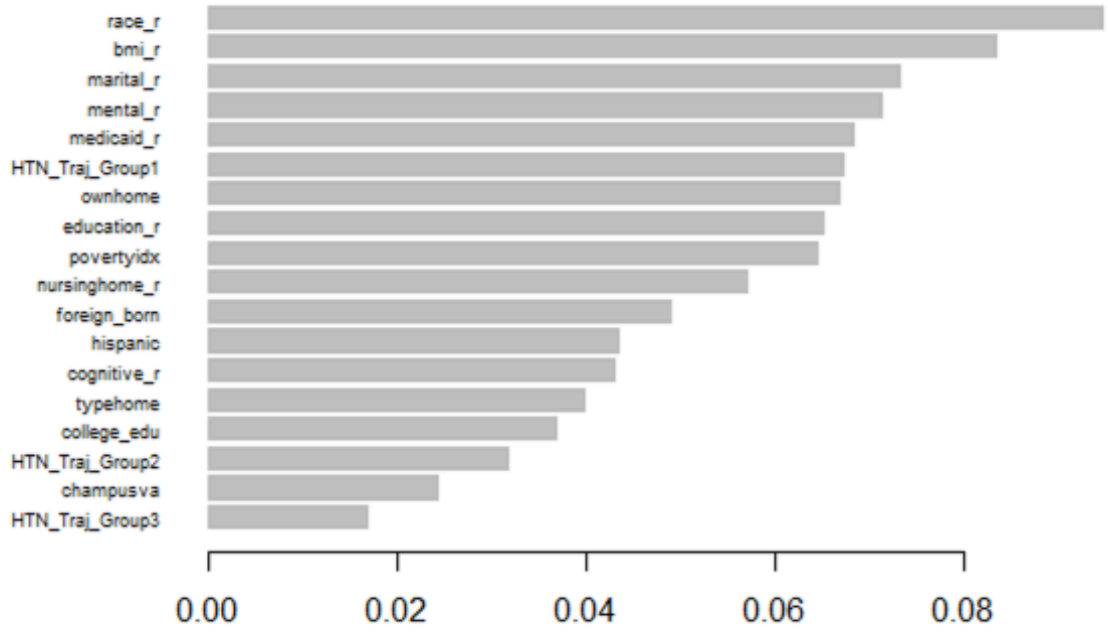

Figure 70 - Variable importance plot for hypertension model and MI outcome

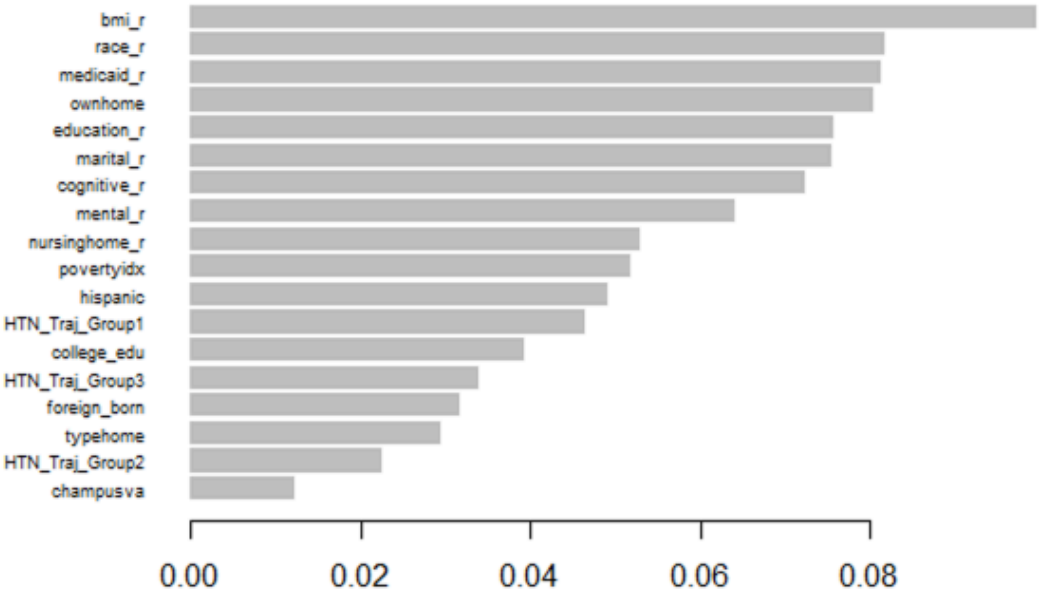

Figure 71 - Variable importance plot for hypertension model and stroke outcome

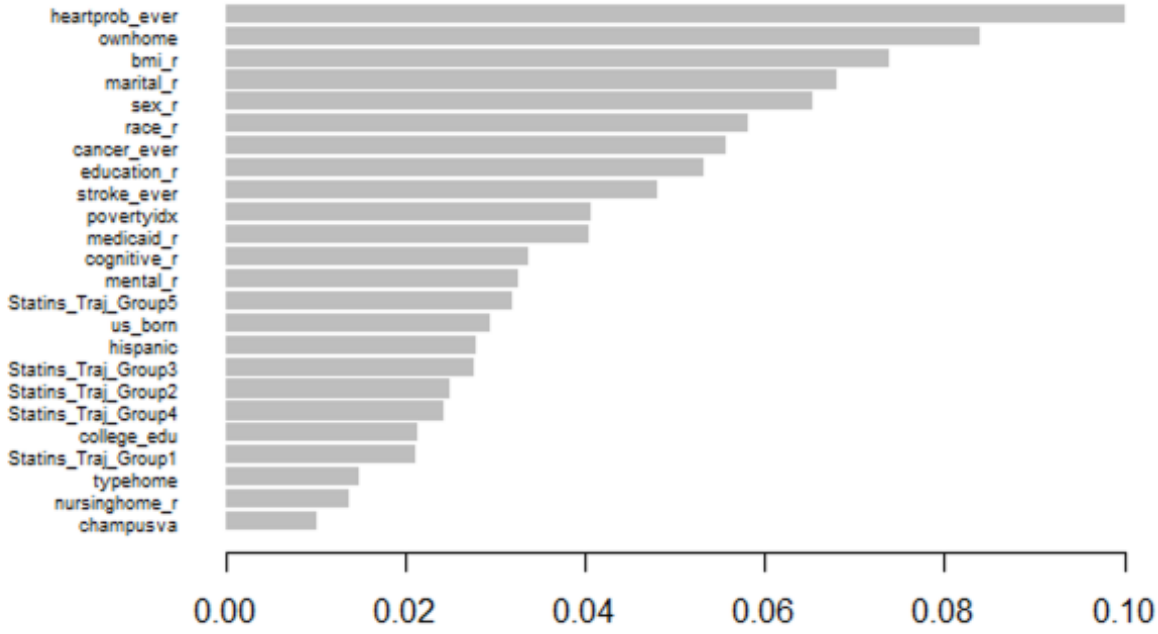

Figure 72 - Variable importance plot for statins model and MI outcome

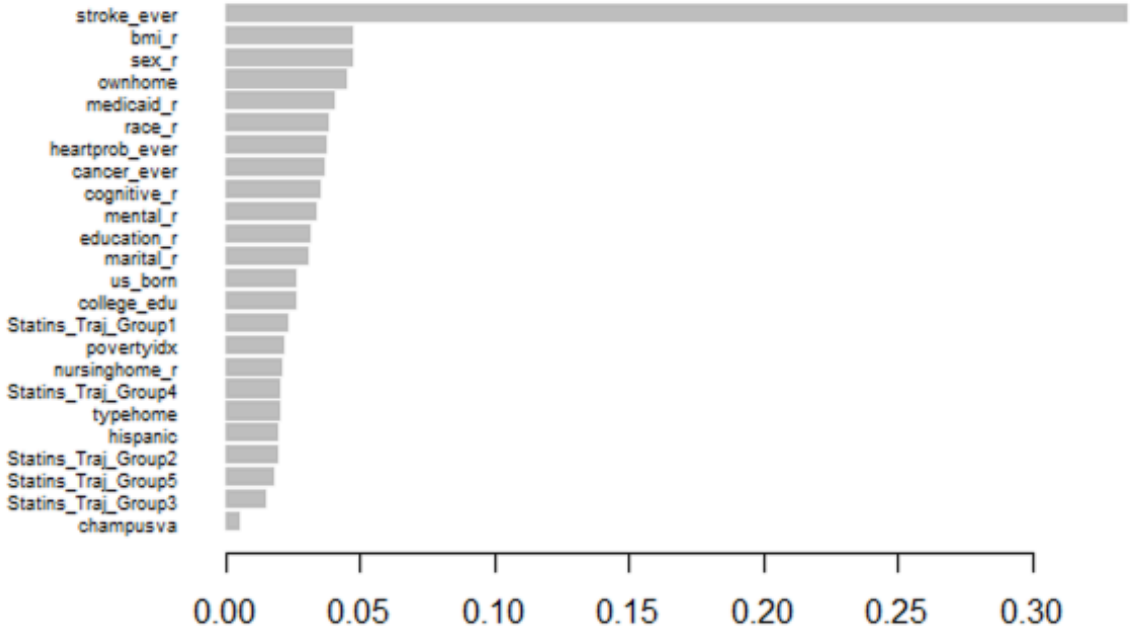

Figure 73 - Variable importance plot for statins model and stroke outcome

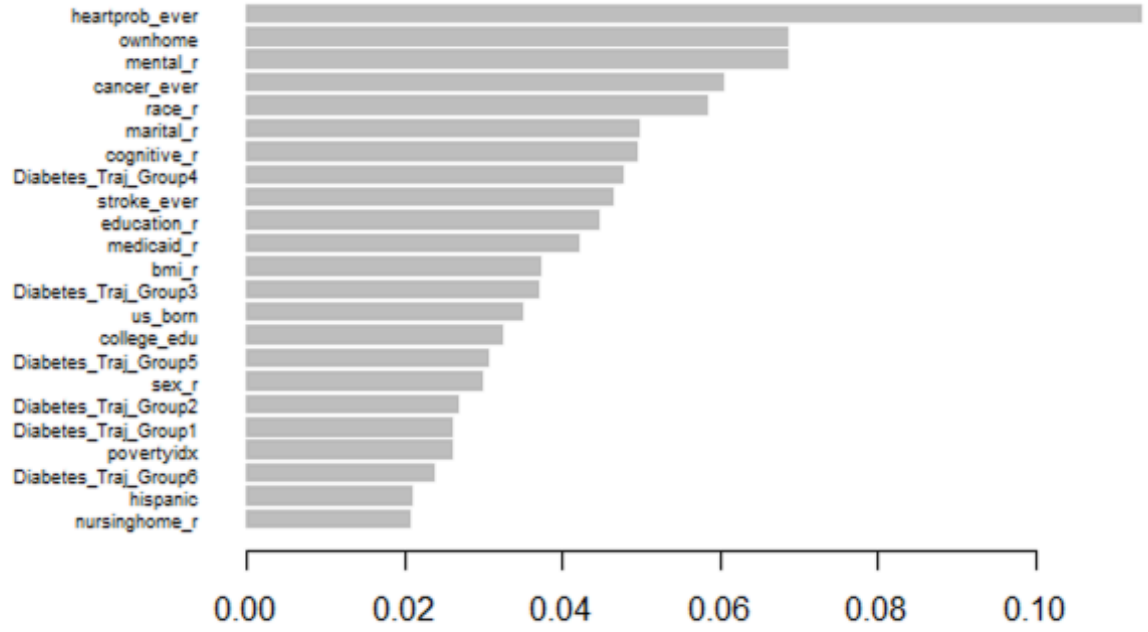

Figure 74 - Variable importance plot for diabetes model and MI outcome

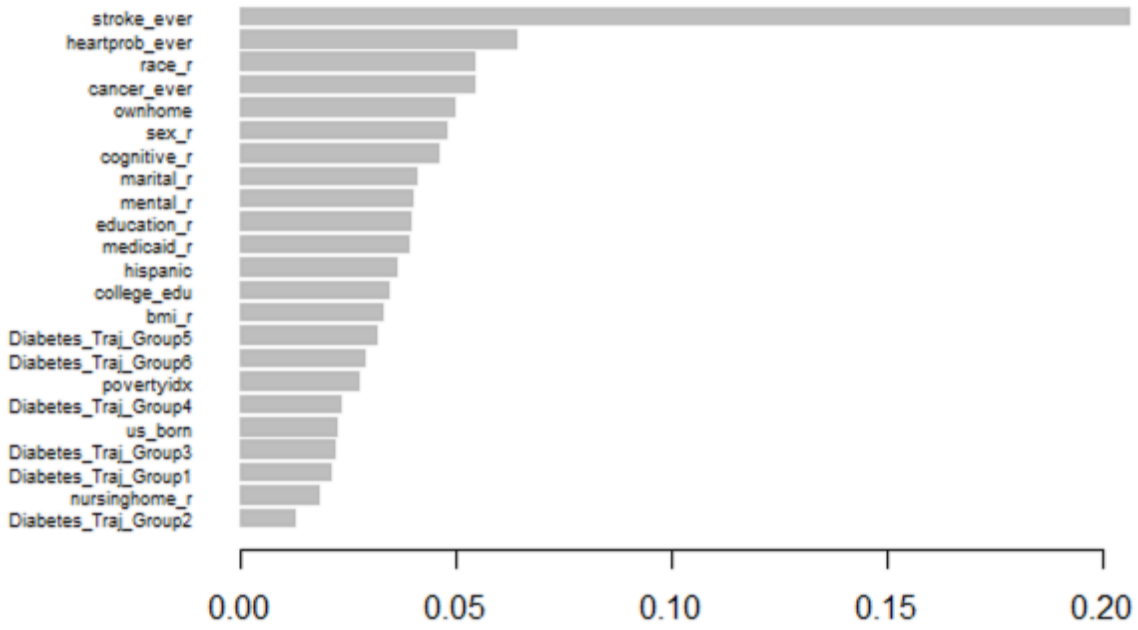

Figure 75 - Variable importance plot for diabetes model and stroke outcome

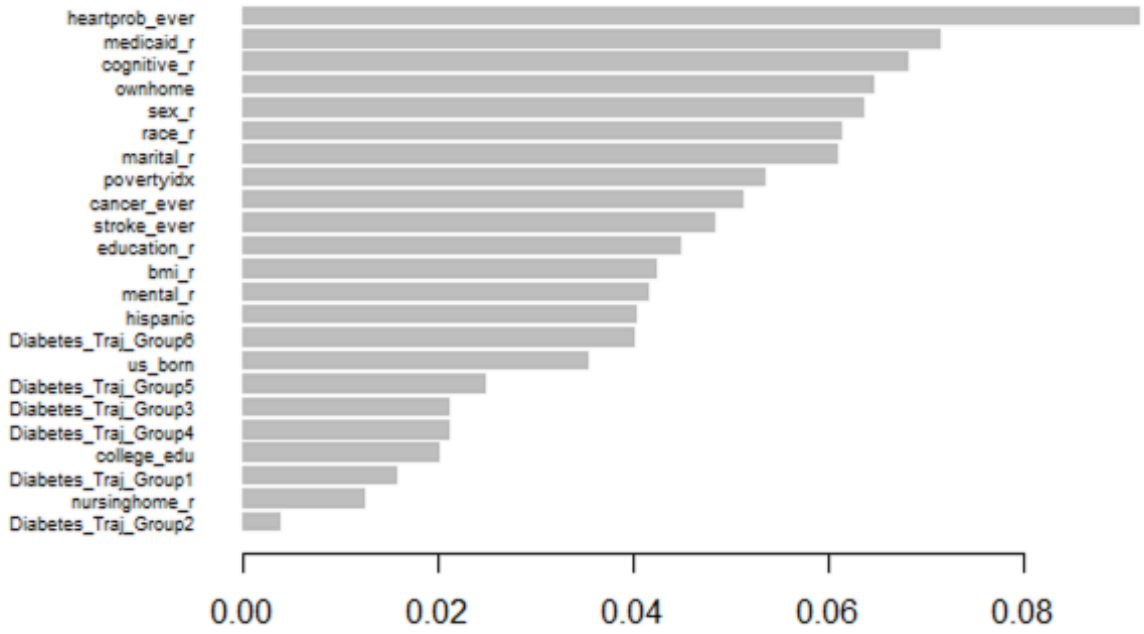

Figure 76 - Variable importance plot for diabetes model and ophthalmic complications outcome

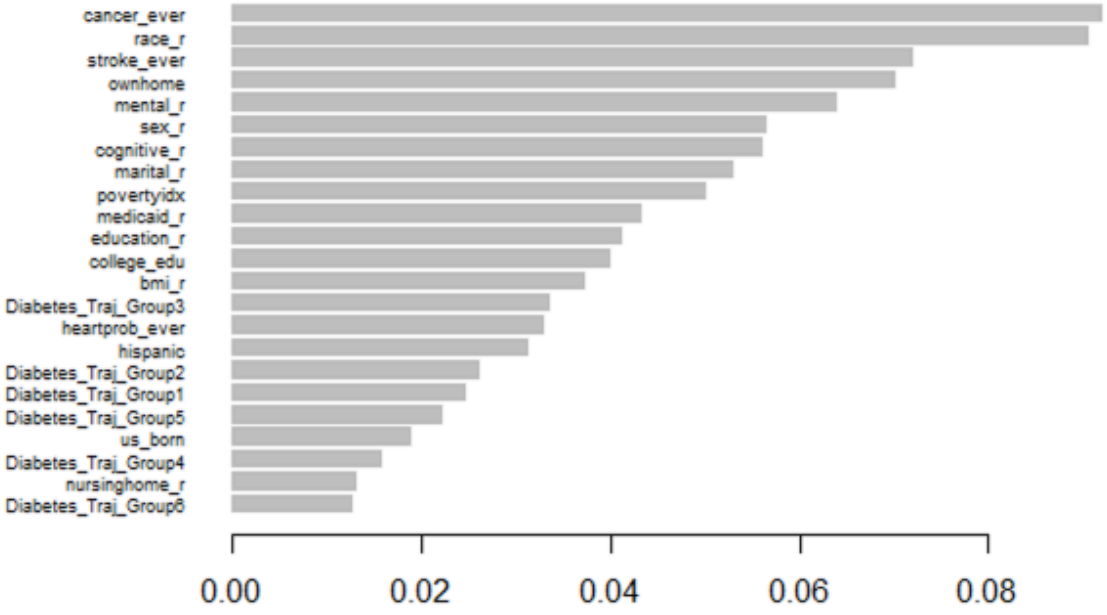

Figure 77 - Variable importance plot for diabetes model and nephropathy complications outcome

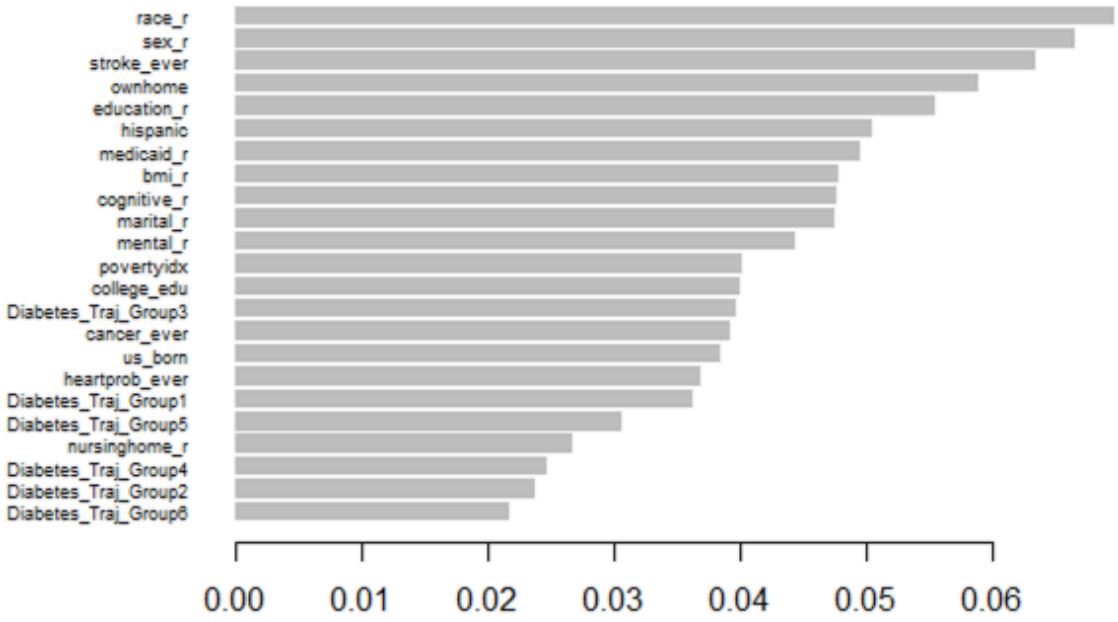

Figure 78 - Variable importance plot for diabetes model and neuropathy complications outcome

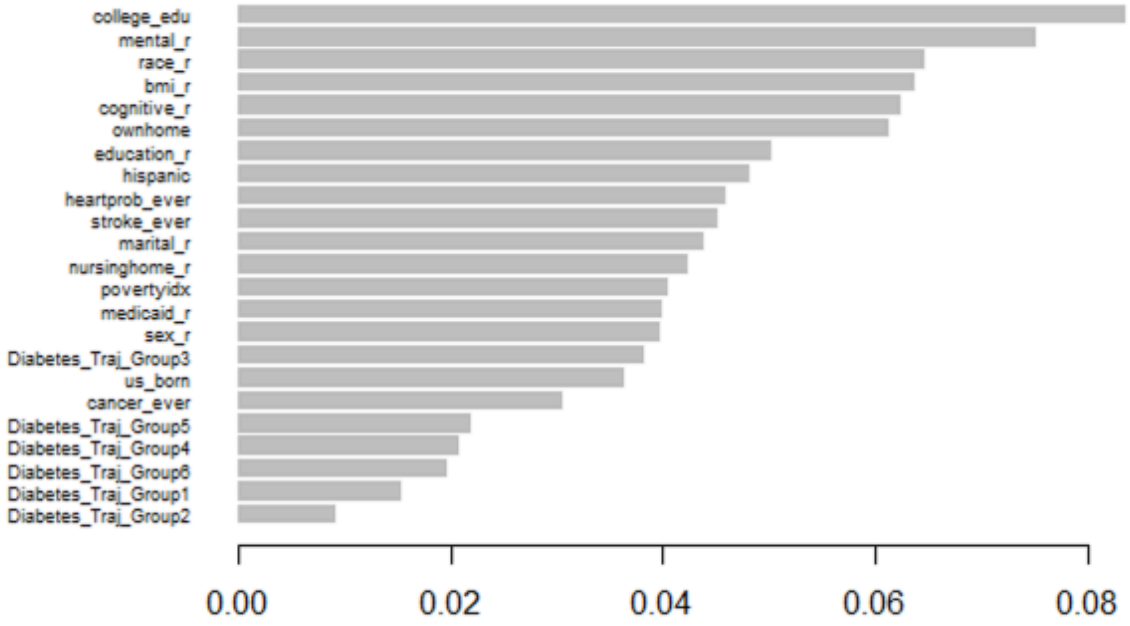

Figure 79 - Variable importance plot for diabetes model and peripheral angiopathy complications outcome