Economic Burden of Low-value Healthcare on Patients with Localized Prostate Cancer: Statistical & Machine Learning Approaches

Ryan Fiano
rfiano@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Health Services Research Commons

Recommended Citation
https://researchrepository.wvu.edu/etd/8118

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Economic Burden of Low-value Healthcare on Patients with Localized Prostate Cancer: Statistical & Machine Learning Approaches

Ryan Fiano

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Health Services Research Commons
Economic Burden of Low-value Healthcare on Patients with Localized Prostate Cancer: Statistical & Machine Learning Approaches

Ryan Fiano, MPH

Dissertation submitted
to the School of Medicine
at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in
Clinical & Translational Science

Usha Sambamoorthi, PhD, Co-chair
Traci LeMasters, PhD, Co-chair
Kim Innes, PhD
Malcolm D. Mattes, MD
Chan Shen, PhD

West Virginia Clinical & Translational Science Institute

Morgantown, West Virginia
2021

Keyword: Prostate Cancer, Patient Centered Care, Conservative Management, Healthcare Costs, Machine Learning, Patient Reported Experience Measures, Care Fragmentation

Copyright 2021 Ryan Fiano
ABSTRACT

Economic Burden of Low-value Healthcare on Patients with Localized Prostate Cancer: Statistical & Machine Learning Approaches

Ryan M. Fiano, MPH

Adults with incident localized prostate cancer represent a large, medically complex population at risk for low-value care. Evidence-based guidelines recommend conservative management (CM) for localized prostate cancer patients with multimorbidity and limited life expectancy, however, 2 in 3 still choose treatment. This dissertation pursued three Aims to address research gaps related to healthcare practices associated with significant morbidity and economic burden on older men with incident localized prostate cancer: 1) examine the leading predictors of low-value healthcare practice of prostate cancer treatment for low-risk prostate cancer; 2) assess the role of patient-reported experience with care on high-value prostate cancer management; and 3) estimate the association of high-value care on non-cancer related healthcare expenditures using machine learning and statistical approaches. In this study, 2 in 3 adults received low-value prostate cancer treatment. Multimorbidity and care fragmentation were among the leading predictors of low-value prostate cancer treatment and contrary to expectations, life expectancy was a weak predictor of treatment receipt. Social determinants of health were highly ranked predictors of treatment. Higher “timeliness of care” patient reported experience scores were associated with high-value CM use. Other forms of low-value care before incident prostate cancer diagnosis were associated with higher non-cancer related healthcare expenditures while high-value CM was associated with lower costs. In summary, this dissertation highlights the negative effect of multimorbidity and care fragmentation on overtreatment, high-value care, and cost outcomes. Perceptions of timely care with healthcare providers and systems have significant impact on high-value CM use among older men with localized prostate cancer. This dissertation reports strong independent predictive associations of incremental low-value healthcare use before incident prostate cancer diagnosis to have significant increases on long-term non-cancer related costs.
DEDICATION

For Haddie and Anderson. May you always have the courage, grit, and patience to pursue your dreams.
ACKNOWLEDGEMENTS

I would like to first thank my mentor, Dr. Usha Sambamoorthi, for her commitment and dedication to me and this work. Dr. Usha is the most humble and intelligent I have encountered in my life. All her students recognize her incredible commitment as a mentor, and I am forever grateful for her patience and kindness.

Dr. Gregory Merrick has also made this dissertation possible. Dr. Merrick's unqualified support, encouragement, and constructive criticisms serve as a keystone in this work. Thank you, Doc!

To my colleagues Dr. JonDavid Pollock, Dr. Robert Galbreath, and Dr. Wayne Butler for both personal and professional guidance in this work and my life. You all have influenced my life in innumerable ways, and I thank you for your example.

I would like to thank my committee members Dr. Traci Lemasters, Dr. Kim Innes, Dr. Malcolm Mattes, and Dr. Chan Shen for their helpful input and patience throughout the completion of this dissertation.

To my CTSI and PSP classmates for being an overwhelming positive influence on my experience as a graduate student. I am grateful for the collegial atmosphere only possible among such selfless individuals. I wish you all the best in your promising futures.

I would like to thank Dr. Mark Olfert, Dr. Paul Chandler, Dr. Julie Lockman, and Dr. Joan Lakoski for facilitating a great academic experience.

I would like to thank my friends and family for their support. I would like to thank my wife, Amanda, and two children, Haddie and Anderson, who also endured the stress of completing a graduate degree. To my parents, Jacqueline and Victor Jr., my brothers Jason, Victor III, and Nicholas for their love, support, and patience. To my grandparents, Victor Fiano Sr. and Kathryn Fiano, for their unconditional love and jovial character that have formed the person I am today; you are missed but never forgotten. Together, we all share this accomplishment.
Table of Contents

List of Tables ............................................................................................................ vii
List of Figures ........................................................................................................ viii
List of Appendices .................................................................................................... ix
List of Abbreviations ................................................................................................ x

1 Introduction ............................................................................................................ 1
  1.1 Background and Significance ........................................................................... 1
  1.2 Innovation ......................................................................................................... 4
  1.3 Specific Aims ..................................................................................................... 5
  1.4 Approach .......................................................................................................... 5

2 Prediction of Low-value Cancer Care among Older Men with Low-Risk Prostate Cancer: A Machine Learning Approach ......................................................... 8
  2.1 Abstract ............................................................................................................ 8
  2.2 Introduction ....................................................................................................... 9
  2.3 Methods .......................................................................................................... 10
  2.4 Results ............................................................................................................ 14
  2.5 Discussion ....................................................................................................... 15
  2.6 Strengths and limitations ................................................................................ 17
  2.7 Conclusions ................................................................................................... 18

3 Associations of Multimorbidity and Patient-reported Experiences of Care with Conservative Management among Elderly Patients with Localized Prostate Cancer ................................................................. 25
  3.1 Abstract ............................................................................................................ 25
  3.2 Introduction ....................................................................................................... 26
  3.3 Methods .......................................................................................................... 27
  3.4 Results ............................................................................................................ 30
  3.5 Discussion ....................................................................................................... 32
  3.6 Strengths and Limitations ................................................................................ 34
  3.7 Conclusions ................................................................................................... 35

4 Healthcare Expenditures Associated with Low-value Care among Older Men with Incident Localized Prostate Cancer: Statistical and Machine Learning Approaches ......................................................... 42
  4.1 Abstract ............................................................................................................ 42
  4.2 Introduction ....................................................................................................... 44
  4.3 Methods .......................................................................................................... 44
  4.4 Results ............................................................................................................ 49
  4.5 Discussion ....................................................................................................... 50
  4.6 Strengths and Limitations ................................................................................ 53
  4.7 Conclusion ..................................................................................................... 53

5 Summary and Conclusion ..................................................................................... 62
5.1 Summary of Findings and Discussion ................................................................. 62
5.2 Implications and Suggestions for Future Research .............................................. 63
5.3 Strengths and Limitations ...................................................................................... 66
5.5 Conclusion .............................................................................................................. 67
6 References ................................................................................................................. 68
7 Appendices ............................................................................................................... 74
List of Tables

Table 2.1 Patient Characteristics by Treatment Use among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using Linked SEER Cancer Registry, 2009-2014 (n=13,870) ..................................................19

Continued Table 2.1 Patient Characteristics by Treatment Use among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using Linked SEER Cancer Registry, 2009-2014 (n=13,870) ............20

Table 3.1 Patient Characteristics by Conservative Management among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (n=496)...............................................36

Table 3.1 “continued” .........................................................................................................................................................37

Table 3.2 Multimorbidity and Patient Experiences by Conservative Management among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry with MCAHPS, 2002-2013 .................................................................38

Table 3.3 Unadjusted (UOR), Adjusted Odds Ratios (AOR) and 95% Confidence Intervals (CI) of Multimorbidity, Mental Health, Timeliness of Care, Low-risk Prostate Cancer, and Education on Likelihood of Conservative Management Use among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (N = 496)..........................................................39

Table 4.1 Patient Characteristics by Low-value Care Use among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671) ......................55

continued Table 4.1 ..........................................................................................................................................................56

Table 4.2 Healthcare Expenditures by Low-value Care Use among Elderly Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer SEER-Medicare, 2005-2014 (n=75671) .........................57

Table 4.3. Healthcare Expenditures by Conservative Management Use among Elderly Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer SEER-Medicare, 2005-2014 (n=75671) .........................58

Table 4.4 Parameter estimates of PCCI categories from unadjusted and adjusted generalized linear models on 12-24 months healthcare expenditures among elderly Medicare fee-for-service beneficiaries with incident localized prostate cancer diagnosis using SEER-Medicare data, 2005-2015 (n=75671). .........................59
List of Figures

Figure 2.1 Feature Importance of Treatment Prediction among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870) ................................................................. 21

Figure 2.2 Positive and Negative Feature Relationships of Treatment Prediction among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870) .................................................................................................................. 22

Figure 2.3 SHAP Partial Dependence Plot of Top 4 Predictive Features of Treatment among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870) .................................................................................................................. 23

Figure 2.4 SHAP Partial Density Plot of Life Expectancy and Multimorbidity Predictive Features on Treatment among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870) .................................................................................................................. 24

Figure 3.1 Cohort Selection and Exclusion ......................................................................................... 40

Figure 3.2 Adapted Competing Demands Framework ........................................................................ 41

Figure 4.1 SHAP Feature Importance and Summary Plot of Low-value Care and Conservative Management Use on Non-Treatment Related Total Healthcare Expenditures among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671) ......................... 60

Figure 4.2 SHAP Partial Density Plots of Selected Features on Non-Treatment Related Total Healthcare Expenditures among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671) ....................................................................................... 61
List of Appendices

Supplemental Figure 7.0 Aim 1 Inclusion & Exclusion Criteria ................................................................. 74

Supplemental Figure 7.1 Feature Importance of Treatment Prediction among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870) ....75

Appendix 7.2 Unadjusted and Adjusted Odds Ratios and 95% Confidence Intervals (CI) of PCCI, Mental Health, Timeliness of Care, Low-risk Prostate Cancer, and Education on Conservative Management Use in among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (N = 496) .................................................................................................................. 76

Appendix 7.3 Unadjusted and Adjusted Odds Ratios and 95% Confidence Intervals of PCCI, Multimorbidity, Mental Health, Timeliness of Care, Low-risk Prostate Cancer, and Education on Conservative Management Use up to 24 Months after Incident Prostate Cancer among Fee-for-Service Medicare Beneficiaries using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (N = 496) .................................................................................................................. 77

Appendix 7.4 Prostate cancer treatment codes used within 12 months of prostate cancer diagnosis to Identify Conservative Management Use ........................................................................................................ 78

Appendix 7.5 Patient Characteristics by Conservative Management among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671) ........................................................................................................ 79

Appendix 7.6 Low-value Care by Conservative Management Use among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671) ...........81

Supplemental Figure 7.7 .............................................................................................................................. 83

Appendix 7.8 Aim 1 python codes ........................................................................................................... 84

Appendix 7.8 Aim 3 python codes ........................................................................................................... 105
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM</td>
<td>Conservative Management</td>
</tr>
<tr>
<td>PSA</td>
<td>Prostate Specific Antigen</td>
</tr>
<tr>
<td>PREM</td>
<td>Patient Reported Experience Measure(s)</td>
</tr>
<tr>
<td>LVC</td>
<td>Low Value Care</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>GLM</td>
<td>Generalized Linear Model</td>
</tr>
<tr>
<td>SDM</td>
<td>Shared Decision Making</td>
</tr>
<tr>
<td>SHAP</td>
<td>SHapley Additive exPlanations</td>
</tr>
<tr>
<td>PDP</td>
<td>Partial Density Plot</td>
</tr>
<tr>
<td>PCCI</td>
<td>Prostate Cancer Comorbidity Index</td>
</tr>
<tr>
<td>BBI</td>
<td>Bice-Boxerman Continuity of care Index</td>
</tr>
<tr>
<td>BETOS</td>
<td>Berenson-Eggers Type of Service codes</td>
</tr>
<tr>
<td>HCPCS</td>
<td>Healthcare Common Procedure Coding System</td>
</tr>
<tr>
<td>SEER</td>
<td>Surveillance Epidemiology, and End Results</td>
</tr>
<tr>
<td>MCAHPS</td>
<td>Medicare Consumer Assessment of Healthcare Providers and Systems surveys</td>
</tr>
<tr>
<td>CCI</td>
<td>Charlson Comorbidity Index</td>
</tr>
<tr>
<td>AOR</td>
<td>Adjusted Odds Ratio</td>
</tr>
<tr>
<td>UOR</td>
<td>Unadjusted Odds Ratio</td>
</tr>
<tr>
<td>FFS</td>
<td>Fee-for-Service</td>
</tr>
</tbody>
</table>
CHAPTER 1

1 Introduction

1.1 Background and Significance

Epidemiology of localized prostate cancer

Prostate cancer is the most frequently diagnosed malignancy and the 2nd leading cause of cancer death among men in the United States. In the last several decades, prostate cancer-specific mortality has decreased significantly and future estimates project robust survivorship growth from 3.3 to 4.5 million by 2026. Wide-spread use of prostate-specific antigen testing has increased detection of malignant and indolent cancer thus presenting evidence-based management challenges to address prostate cancer specific mortality of higher-, and avoid overtreatment of lower-, risk disease.

Localized, low-risk prostate cancer

American Joint Commission on Cancer (AJCC) establishes a standard staging system using Tumor, Node, Metastasis (TNM) classifications for incident prostate cancer diagnoses; primary tumor (T), regional lymph node involvement (N), and distant metastasis (M) are assessed. Localized prostate cancer, or cancer that is confined to the prostate, is classified as clinical stage T1 to T2a. Risk stratification, broadly, includes clinical staging, Prostate Specific Antigen testing, and Gleason Score with 3 main risk groups: low-, intermediate-, and high-risk prostate cancer. Low-risk prostate cancer is defined as T1 to T2a, Gleason 6, PSA <10ng/mL, intermediate risk defined as T2b-T2c or Gleason 7 or higher or PSA 10-20 ng/mL, and high risk defined as T3a or Gleason 8 or higher or PSA > 20 ng/mL.

Treatment selection for localized prostate cancer

Older men with indolent localized prostate cancer and/or limited life expectancy are vulnerable to adverse effects of overtreatment. Overtreatment of localized prostate cancer is associated with negative health outcomes such as sexual, urinary, and rectal dysfunction including impotence, incontinence, and bowel irritation. To avoid overtreatment morbidities, non-curative conservative management (CM) strategies have been established as an evidence-based option for adults with localized, low-risk cancer and/or limited life expectancy. Clinical practice guidelines have recommended CM approaches for a decade. More recently, providing curative treatment of low-risk prostate cancer or higher-risk disease among patients with
less than 5 years of life expectancy without discussing CM is considered an indicator of sub-optimal healthcare quality.\(^{(12,14)}\) Thus, the premise of CM is to avoid overtreatment-related adverse effects as localized malignancies pose little risk of prostate cancer-specific mortality.

*The Burden of Low-value Healthcare*

For decades, wasteful healthcare spending in the United States has not improved the cost, quality, and outcomes in comparison to other developed countries. Efforts to identify and address waste, as outlined in the Institute of Medicines (IOM) seminal work *Crossing the Quality Chasm*, have promoted a national effort to develop healthcare quality measurements to promote patient centered care; timely, well-coordinated care to inform healthcare decisions that reflect the patient’s values.\(^{(15)}\) To build on the IOM seminal work, The National Quality Task Force has redefined effective and patient-centered care as “appropriate, person-centered care”, thereby recognizing the adverse effects associated with inappropriate healthcare and affirming the need to understand chronic conditions management within the context of overall well-being.\(^{(16)}\)

Low-value care, defined as inappropriate healthcare lacking a net clinical benefit to the patient, contributes to avoidable morbidities and excessive healthcare costs.\(^{(17)}\) Low-value care is prevalent among elderly patients affecting 1 in 4 Medicare beneficiaries\(^{(18)}\) and accounts for $75.7 - $101.2 billion in annual healthcare costs.\(^{(19,20)}\) Use of low-value healthcare services among medically complex patients can lead to cascades of unnecessary down-stream care, cumulating costs as high as 10 times the original low-value healthcare service.\(^{(21)}\)

To combat and eliminate low-value care, initiatives supported by the American Board of Internal Medicine Foundation have established the Choosing Wisely campaign to work with other professional societies to develop lists of low-value screenings, tests, and procedures.\(^{(22)}\) Since the establishment of Choosing Wisely in 2012 study of low-value care has increased substantially. The National Quality Forum outlined opportunities for increasing requirements to “educate and engage” specific patient populations to reduce low-value care, while increasing patient-centered, high-value care.\(^{(16)}\) However, few studies are oncology focused, with many examining the overuse of screening (i.e., breast and prostate cancer screening) versus more harmful impacts associated with overtreatment.\(^{(23)}\) Significant research gaps exist in the understanding how low-value care contributes to “down-stream” overtreatment and cost within specific clinical scenarios.
**Patient-centered care**

Patients with incident localized prostate cancer represent a large, medically complex population at risk of low-value care in the form of overtreatment. Older patients with localized prostate cancer have high rates of multimorbidity, defined as more than 2 chronic conditions, and often experience an increase in care fragmentation at diagnosis.(24) Care fragmentation is known to present barriers to shared decision making and evidence-based recommendations among adults with multimorbidity,(25,26) therefore adults with multimorbidity may be at increased risk for receiving overtreatment, low-value care, and/or increased downstream survivorship costs.

Shared decision making facilitates patient choice amid the risks and benefits of treatment and non-treatment options. Incident prostate cancer patients are challenged by multimorbidity and care fragmentation that influence perception of the healthcare system, potentially pushing them to choose an immediate cure(27) versus managing an additional chronic condition in a fragmented system. Fear of treatment regret, anxiety, and misunderstanding the risks and benefits of curative treatment within the context of cancer risk grouping and life expectancy are likely influenced by a patient’s perception of care.(27)

The use of patient-reported experience measures can amplify the patient’s perspective of the continuity of care among primary care and specialty clinicians. Patient reported experiences of care include domains such as physician communication, timely care, and perceptions of getting needed care and are increasingly used as quality measures by health plans, medical groups, and physician practices.(28) Positive patient-reported experience scores are associated with adherence to medical advice, improved clinical outcomes, and lower utilization of unnecessary health-care services.(15,29) Identifying specific measures of patient-reported experiences that facilitate CM use among patients with incident localized prostate cancer and multimorbidity is needed to promote evidence-based cancer care.(30) For example, in colorectal cancer populations, patient-reported experiences with perceived timely care are associated with evidenced-based follow-up.(31) Understanding the relationship between patient-reported experiences of care on CM use can inform patient-centered care approaches by identifying gaps in timely care from the patient’s perspective to improve adoption of CM use thereby reducing the adverse effects of overtreatment among older patients with multimorbidity and localized prostate cancer.
Economic burden of localized prostate cancer

Many studies have examined comparative costs between different prostate cancer treatment modalities including CM. CM use is known to reduce short-term treatment-related costs, however, long term costs associated with overtreatment or low-value care is not known. Use of low-value services prior to incident prostate cancer may be associated with increased costs throughout survivorship regardless of curative treatment or CM modalities.

1.2 Innovation

Using recent literature reviews(27,33–37) and healthcare initiative recommendations(16,22), the research presented in this dissertation addresses significant research gaps as related to healthcare practices that provide limited clinical benefit, potentially harmful effects, and a significant economic burden on older men with incident localized prostate cancer. Specifically, we used novel machine learning techniques with nationally representative samples of older men with localized prostate cancer to answer research questions with both clinical and population relevance. We used interpretative model-agnostic approaches to understand the complex machine learning outputs with a focus on understanding predictors of localized prostate cancer treatment. We also used traditional statistical approaches to understand the impact of patient-reported outcomes on CM use and non-treatment related costs associated with the use or low-value care prior to incident localized prostate cancer diagnosis. We identified misalignment between prostate cancer specific, validated measures of life expectancy and use of guideline recommended care among localized prostate cancer patients. We also identified the predictive qualities of care fragmentation and use of low-value care on non-cancer related healthcare expenditures during survivorship. The innovation of this work addresses the strategic objectives, research recommendations, and general research gaps as we identify novel factors, such as care fragmentation, low-value care use, and patient-reported outcomes, on high- and low-value treatment and costs outcomes among older adults with localized prostate cancer.

To highlight healthcare practices that provide limited clinical benefit, potentially harmful effects, and a significant economic burden among older men with incident localized prostate cancer, we pursue the following specific aims:
1.3 Specific Aims

Aim 1: Use machine learning (ML) to identify leading predictors of cancer treatment within 12 months of diagnosis among older men with localized, low-risk prostate cancer.

- Quasi-hypothesis H1: Leading predictors of low-risk prostate cancer treatment will be age, multimorbidity, care fragmentation, and life expectancy.

Aim 2: Examine the associations of patient-reported experiences with providers and healthcare systems and multimorbidity on high-value care (CM use) among older men with localized prostate cancer.

- H2: Older men with multimorbidity and higher patient experience scores will be more likely to use CM compared to those without multimorbidity and lower patient experience scores, after adjusting for patient-, clinician-, and practice ecosystem factors.

Aim 3: Identify leading predictors and estimate the association of high- and low-value care use on non-cancer related healthcare expenditures among older men with localized prostate cancer.

- Quasi-hypothesis H3.1: Leading predictors of non-cancer related healthcare expenditures will be age, life expectancy, low- and high-value care.

- Quasi-hypothesis H3.2: Independent of treatment or conservative management use, adults using low-value care, versus no low-value care, will have higher non-treatment related healthcare costs 12-24 months after a localized prostate cancer diagnosis.

1.4 Approach

Conceptual Framework

Management of preexisting multimorbidity, care fragmentation, and the shared decision making process associated with choosing treatment or conservative management for incident prostate cancer requires the use of limited resources (i.e., time to manage chronic conditions and availability of health-care professionals and resources). Therefore, in this dissertation we used a competing demands model to conceptualize factors known to affect localized prostate cancer treatment selection within clinician, patient, and practice ecosystem domains.(33,38,39)

Data Sources

For aims 1 and 3 we used the Surveillance, Epidemiology, and End Results (SEER)-Medicare linked
data files. SEER-Medicare data includes data on all incident prostate cancer cases occurring in 18 regions throughout the United States. Incident prostate cancer, prostate cancer specific clinical information (i.e., Gleason Score and Prostate-specific antigen values), and cancer staging information were derived from the SEER database. Medicare claims were derived from inpatient, outpatient, and other files linked to SEER.

For aim 2 we used a separate SEER with Medicare Consumer Assessment of Healthcare Providers & Systems (MCAHPS®) survey linkages. MCAHPS, administered by the CMS, use standardized and validated questionnaires to collect information on patient-reported experiences with health-care providers.(40) MCAHPS collection methodologies use a weighted probability sampling procedure covering all the 50 US states, DC, and Puerto Rico, which are then linked to SEER patients.(40,41)

Area Health Resource File (AHRF) files were linked via MEDPAR FIPS state and county codes and used to calculate urologist densities.(42) Census files were linked to calculate county-level median income and college education.

We used both statistical and machine learning approaches in this dissertation. All model adjustments were made with factors known to influence treatment or conservative management selection. We used frequency and percentages (i.e., Chi-square tests and t-tests) to identify significant group differences by categorical variables and standard deviations for continuous variables. Generalized linear models (GLM) with log-link and gamma distribution were used to estimate key independent variable associations with healthcare cost. In aim 2, multivariable logistic regression was used to estimate the effect of patient-reported experience measures of care and healthcare providers on conservative management use.

ML methods can determine the most effective and parsimonious model/algorithm through classification of a binary target (i.e., yes or no dependent variable) via iterative learning. We used the XGBoost classifier, a decision-tree ensemble machine learning algorithm, to predict prostate cancer treatment or CM use among patients with low-risk prostate cancer diagnosis.(43)

Traditional statistical approaches use the entire data set and goodness of fit statistics for model specification.(44) Machine learning techniques utilize a training and test data set to estimate the predictive accuracy of a model. In ML models, original datasets were split (70%/30%) into training and test data. Next, predicted labels and probabilities, or a regression form such as log expenditures, were optimized on the
training data via hyperparameter tuning using stratified 10-fold cross validation. Hyperparameters can adjust regularization through iterative tuning runs. Stratified cross validation describes the process of splitting the training data into identical “folds” that can be used as versions of “unseen” data during testing. Once the XGBoost algorithm is optimized on training data for accuracy, it is “tested” on the original 30% split. We assessed predictive performance on the test set using precision, recall, accuracy, and Precision-Recall Receiver Operating Curve scores for machine learning models in aim 1. In aim 3, XGBoost regressors we used to assess prediction of log-healthcare costs and performance was assessed using $R^2$ and negative mean squared error.

**Shapley Additive exPlanations**

To improve machine learning model interpretability, Shapley Additive exPlanations (SHAP) were used to visually summarize the contributions and associations of the features to the target variable. SHAP values represent the magnitude of effect from each feature on the model output by conditional expectation when the feature is “hidden”. Attributed changes induced on the model output are then averaged overall possible feature orderings in the data set (i.e. coalition) to provide the SHAP value. Feature importance can be calculated by sorting absolute average SHAP values whereas summary plots sort average SHAP outputs with positive and negative variations on the x-axis. Partial dependence plots (PDP) average SHAP values for a feature by a feature’s unit value (for example, average age SHAP values by the feature age in years). Unlike traditional PDP, SHAP PDP shows dispersion around feature units reflecting the impact of all the other features in the model. Lastly, PDP interactions display the attribution effect of two features simultaneously. For aim 1 and 3, model SHAP outputs are log-odds and log-cost, respectively.
CHAPTER 2

2 Prediction of Low-value Cancer Care among Older Men with Low-Risk Prostate Cancer: A Machine Learning Approach

2.1 Abstract

Background: Older men with incident prostate cancer are vulnerable to low-value prostate cancer treatment. Despite evidence-based support for conservative management (i.e., non-treatment), approximately 2 in 3 Medicare beneficiaries receive treatment for low-risk prostate cancer. Adults with multimorbidity who experience care fragmentation are vulnerable to departures from evidence-based medicine. A comprehensive analysis of clinical and non-clinical factors, such as life expectancy and care fragmentation, that may drive low-value prostate cancer treatment is lacking.

Objective: Use machine learning (ML) to identify leading predictors of cancer treatment within 12 months of diagnosis among older men with low-risk prostate cancer. Novel predictors included validated prostate-cancer specific life expectancy and care fragmentation.

Methods: In this retrospective cohort study we linked Surveillance, Epidemiology, and End Results cancer Registry (SEER), Medicare Claims, Census, and Area Health Resource files and included older men with incident low-risk prostate cancer from 2009 to 2014 (n=13,870). We used claims data to identify treatment (Yes/No) in the first 12 months after diagnosis. We used the XGboost algorithm and SHapley Additive exPlanations (SHAP) to rank feature importance in treatment prediction.

Results: In our study cohort (n=13,870), 66.9% of older adults received cancer treatment. Age, multimorbidity, care fragmentation, social support, and social determinants were leading predictors of cancer treatment (Accuracy=0.70, Precision=0.71, Recall=.92, Precision-Recall Area Under the Curve = 0.78). Relationships of college education, income, and care fragmentation on low-value cancer treatment were nonlinear and complex. Life expectancy was a weak predictor of prostate cancer treatment.

Conclusion: Our results suggest that non-clinical factors such as social determinants, care fragmentation, and social support are the most important predictors of treatment among older men diagnosed with low-risk prostate cancer. Despite a critical role in evidence-based treatment recommendations, life expectancy had limited impact on treatment selection.
2.2 Introduction

Localized prostate cancer diagnoses are often indolent, representing a large potential source of low-value cancer care. Older men with low-risk prostate cancer and limited life expectancy risk significant urinary, erectile, and rectal morbidities without the benefit of significant improvement in prostate cancer specific mortality outcomes.\(^{(7,8,10)}\) High-level evidence discourages treatment \(^{(3)}\) of individuals with low-risk prostate cancer with less than 10 years of life expectancy or higher-risk disease with a life expectancy of 5 years or less.\(^{(6)}\) Despite evidence-based support for CM in recent years, treatment for low-risk prostate cancer is highly prevalent, for example, among Medicare beneficiaries as a majority (58%) still receive treatment as compared to integrated healthcare network cohorts (20%).\(^{(46)}\)

Decision to treat low-risk prostate cancer with radiation, surgery, cryotherapy, chemotherapy, or hormone therapy requires life expectancy estimation within the context of prostate cancer specific mortality and death from competing risks. Evidence-based guidelines recommend conservative management for low-risk prostate cancer and life expectancy of 10 or fewer years.\(^{(6)}\) Examining the impact of life expectancy as a predictor is critical in assessing low-value treatment of low-risk prostate cancer, however, many methods used in population studies vary considerably and many lack validity in prostate cancer populations.\(^{(34)}\)

Multimorbidity is another factor that may influence treatment of low-risk cancer. Over 60% of older adults with incident prostate have multimorbidity defined 2 or more chronic conditions.\(^{(47),(48)}\) Prior studies with SEER-Medicare data have shown that older men with 3 or more chronic conditions were significantly more likely to get treated for localized prostate cancer in adjusted multivariable models.\(^{(49)}\)

Adults with multimorbidity who experience care fragmentation are vulnerable to departures from evidence-based medicine and poorer health outcomes.\(^{(26)}\) Approximately 50% of adults with multimorbidity already see 3 or more specialists \(^{(50)}\) and increased care fragmentation is associated with diagnosis of incident prostate cancer.\(^{(24)}\) Medically complex adults experiencing care fragmentation prior to diagnosis are likely to encounter additional barriers to shared-decision making process critical for understanding the risks and benefits of treatment. Shared decision-making engagement fosters trusting relationships that facilitate communication of evidence-based recommendations and is associated with reductions in both cancer and non-cancer low-value care use.\(^{(27,51)}\)
Supervised ML models are increasingly used to predict healthcare outcomes, quality, and cost outcomes. However, many studies investigating low-value prostate cancer care use only conventional statistical approaches. A comprehensive analysis of clinical and non-clinical factors, namely, valid life expectancy measures, multimorbidity, and care fragmentation, that drive low-value cancer treatment among adults with incident low-risk prostate cancer is lacking. Supervised ML can be used to identify the leading predictors and non-linear relationships among medically complex adults with higher probability of overtreatment.

Therefore, the objective of this study was to identify leading predictors of low-value cancer care among older men with low-risk prostate cancer using machine learning methods (ML). We used supervised ML models to identify leading predictors of low-value prostate cancer care by using nationally representative linked cancer registry-Medicare claims data incorporating a comprehensive list of patient, clinical, and practice ecosystem features.

2.3 Methods

Study Design

In this retrospective cohort study, the date of incident low-risk prostate cancer diagnosis was used as an index date and 12 months before diagnosis was used as the baseline period. We defined the treatment for low-risk prostate cancer (radiation, surgery, chemo- and hormone- therapies) during a period of 12 months after diagnosis based on validated methods for claims data.

Study population

Older adults (age ≥ 66 years) diagnosed with low-risk prostate cancer during the study period (2009-2014) were included in the study population. Cancer diagnosis was identified using International Classification of Diseases for Oncology (ICD-9). Low-risk prostate cancer was defined as American Joint Committee on Cancer stage T2a or less, Gleason Score ≤ 6, and a PSA test ≤ 10 ng/mL.

Data sources

This study used the Surveillance, Epidemiology, and End Results (SEER)-Medicare linked data files. SEER-Medicare data includes data on all incident prostate cancer cases occurring in 18 regions throughout the United States. Incident prostate cancer, prostate cancer specific clinical information (i.e., Gleason Score and...
Prostate-specific antigen values), and cancer staging information were derived from the SEER database. Patient Medicare claims were derived from inpatient, outpatient, and other files linked to SEER.

Area Health Resource File (AHRF) files were linked via MEDPAR FIPS state and county codes and were used to calculate urologist densities. Census files were linked to calculate county-level median income and college education (percentage with college degree).

**Study inclusion/exclusion criteria**

Adults with low-risk prostate cancer were included in the study if alive and continuously enrolled in Medicare Parts A and B throughout the baseline and follow-up periods (Appendix Supplemental Figure 7.7). Adults with missing cancer stage, PSA values, Gleason Scores, and/or those diagnosed with prostate cancer at autopsy were excluded.

**Target variable: Prostate cancer treatment (yes/no)**

Prostate cancer treatment was estimated within the first 12 months after diagnosis using a previously validated, claims-based algorithm and included radiation, surgery, and chemo-, cyro-, and/or hormone-therapy. (53)

**Key features**

*Life expectancy*

Prostate cancer comorbidity index (PCCI) is a weighted comorbidity index validated in prostate cancer patient populations to predict life expectancy. (54) Medical colleges, patient advocacy groups and evidence-based recommend informed decision-making in using conservative management or treatment for men with low-risk prostate cancer and a life expectancy of greater than 10 years. (6) We calculated PCCI during the baseline period to create predictive models based on validated life expectancy (10 year or greater life expectancy; PCCI<8) to reflect evidence-based recommendations. (6,54) Categorical variables at clinically meaningful cut-offs are presented in Table 2.1 and were also included in sensitivity analyses for interpretative purposes.

*Multimorbidity*

The multimorbidity framework developed by the United States Department of Health and Human Services for guiding programs, practice, and policy guided the selection of chronic conditions as follows:
arthritis, asthma, coronary artery disease, congestive heart failure, chronic kidney disease, chronic obstructive pulmonary disease, cardiac arrhythmias, acute myocardial infarction, dementia, diabetes, depression, hepatitis, hyperlipidemia, hypertension, human immunodeficiency virus, osteoporosis, substance abuse, schizophrenia, stroke, and anemia. (55) We defined multimorbidity presence of 3 or more chronic conditions.

**Care fragmentation**

We used a modified version of the Bice-Boxerman continuity of care index to calculate care fragmentation during the 12 month baseline period. (56,57) The continuity of care index represents the concentration of visits per patient among health care providers based on visit number, proportion of encounters to each provider, and total number of visits. The BBI has been used to compare care fragmentation among cancer survivors and is highly correlated with other measures of care fragmentation. (58) We used physician specialty codes representing primary care, oncology, and various specialist visit encounters; where \( n \) is the total number of visits, \( nk \) is the total number of visits associated with a physician specialty, and \( k \) the total number of physician specialty codes. (59) For example, multiple encounters with a single provider would result in a score of zero, however, multiple encounters among several health care providers would result in a score approaching 100. We included primary care, oncology, and many other specialty codes.

**Other features**

Patient factors included age, race, marital status, county-level median income and education (college graduation percent). Practice ecosystem included urologist/radiation oncologist density, diagnosis year, SEER region and Rural groups (Metro, Urban, Rural). All features were calculated in the 12 months before cancer diagnosis.

Propensity to seek care was also measured by two domains: the number of low-value care services as well as high-value care. Low-value healthcare services were operationalized using a claims-based algorithm representing Choosing Wisely campaign recommendations via procedure, diagnosis, hospitalization, and Berenson-Eggers Type of Service (BETOS) codes using previously published methods. (60–62) We estimated the feature importance of individual low-value care procedures and summed those with an importance score greater than zero. We operationalized preventative services based on the National Commission on Prevention Priorities by the American Academy of Family Physicians as “high-value” care which included lipid and A1c
screenings, influenza vaccinations, and primary care well visits. We used Healthcare Common Procedure Coding System (HCPCS) codes to identify preventative care.

All predictors were measured during the 12 month baseline period before incident low-risk prostate cancer diagnosis.

**Statistical Analysis**

Chi-square tests and t-tests were used to identify significant group differences in prostate cancer treat or CM use by categorical variables. All statistical tests were 2-sided with a 5% Type I error rate and completed in SAS (SAS Institute Inc., Cary, NC, USA. Version 9.4, using Windows 10).

**Machine learning**

ML methods can determine the most effective and parsimonious model/algorithm through classification of a binary target (i.e., yes or no dependent variable) via iterative learning. We used the XGBoost classifier, a decision-tree ensemble machine learning algorithm, to predict prostate cancer treatment or CM use among patients with low-risk prostate cancer diagnosis.(43)

First, our original dataset was split (70%/30%) into training and test data. Next, predicted labels and probabilities were optimized on the training data via hyperparameter tuning using repeated stratified 10-fold cross validation. We used hyperparameter tuning on the training data to optimize performance and avoid overfitting. Hyperparameter tuning parameters codes are available in Appendix Table 7.1.

Final model predictions were evaluated using the original “hold-out” test data. We assessed predictive performance on the test set using precision, recall, accuracy, and Precision-Recall Receiver Operating Curve scores. Machine learning analyses were conducted using Python 3.7 with open-source scikit-learn 0.21.3 wrapper interface for XGBoost classifier.

**Machine learning interpretation**

In the ML model, Shapley Additive exPlanations (SHAP) were used to visually summarize the contributions and associations of the features to the target variable. SHAP values represent the magnitude of effect from each feature on the model output by conditional expectation when the feature is “hidden”. Attributed changes induced on the model output are then averaged overall possible feature orderings in the data set (i.e. coalition) to provide the SHAP value.(45) Feature importance can be calculated by sorting
absolute average SHAP values whereas summary plots sort average SHAP outputs with positive and negative variations on the x-axis. Partial dependence plots (PDP) average SHAP values for a feature by a feature’s unit value (for example, average age SHAP values by the feature age in years). Unlike traditional PDP, SHAP PDP shows dispersion around feature units reflecting the impact of all the other features in the model. Lastly, PDP interactions display the attribution effect of two features simultaneously. All model SHAP outputs are log-odds and are ranked in decending order.

TreeSHAP was used to generate SHAP values, feature importance, partial dependence plots, and PDP interactions with Python 3.7 and the SHAP package (0.29.2).

2.4 Results

Descriptive cohort statistics

The study cohort was predominantly non-Hispanic, white (83.3%), marital status of married (69.1%) with a median age of 72.0 (Mdn=71.1, SD=4.46).

Overall, marital status of married, younger men in the age group 66-74 years, higher income, higher education level, lower physician quartiles, and adults residing in Northeastern SEER regions more frequently received treatment for low-risk prostate cancer (Table 2.1). In recent years, (2013 and 2014) mean age was lower among adults using treatment (M=71.3, SD=4.05) versus no treatment (M=71.7, SD=4.29; p<0.001). Life expectancy did not significantly differ between adults using treatment versus no treatment, however, multimorbidity was significantly more frequent among treatment groups (Table 2.1). Preventative care, flu vaccinations, and primary care visits were significantly more frequent among treatment groups.

Highest ranking predictors

Marital status of married, men in the age group 66-74 years, care fragmentation, median income, and college education were the top 5 predictive features of low-risk prostate cancer treatment (Figure 2.1). Men who were married, aged 66-74 years, and experienced care fragmentation were more likely, and those with higher median income and college education were less likely to receive treatment (Figure 2.1). Northeast SEER region, diagnostic year 2010, multimorbidity, and west SEER region, more recent diagnostic year (2014 and 2013), were associated with higher and lower SHAP values (log-odds) of treatment prediction, respectively (Figure 2.1).
Figure 2.3 shows the complex individual variation among the top 4 predictive features. Overall, SHAP log-odds decreased with higher median income, however, substantial individual-level variation in log-odds by distinct median income values was observed (Figure 2.3 (d)).

**Lowest ranking predictors**

Life expectancy of 10 or more years (PCCI) and non-cancer low-value care use ranked 22nd and 27th, respectively in feature importance out of 29 total features (Figure 2.1). SHAP partial density plots illustrate variation among adults with life expectancies less than 10 years with SHAP log-odds of treatment (Figure 2.4). Hispanic, black, and white race were 23rd, 24th, and 25th ranked predictors, respectively (Appendix 7.1)

2.5 Discussion

In this first study using ML to predict treatment among older men with low-risk prostate cancer, we observed that 2 in 3 adults with low-risk prostate cancer received treatment. Other recent SEER-Medicare data studies also report high (57.9%)(63), while integrated care network studies report low (21%), rates of low-risk prostate cancer treatment among older men (65+ years). High rates of low-value treatment are concerning as many men are likely to experience avoidable, negative health related quality of life and higher healthcare costs. For example, in the Prostate Cancer Intervention vs Observation Trial (PIVOT), a survival benefit to treatment versus conservative management was not significant after 10 years of follow-up. Treatment of low-risk prostate cancer confers significant risk of avoidable morbidities(10) and imposes a heavy economic burden of 1.2 billion dollars in annual cost(64). We speculate that treatment for low-risk prostate cancer may be high as older men experience adverse mental health effects associated with notification of higher PSA levels, pending biopsy results, and a prostate cancer diagnosis.(27) However, we observed that diagnosis year was a leading predictor of treatment, with adults diagnosed in later years (i.e. 2013 and 2014) less likely to receive treatment for low-risk prostate cancer. Our findings revealing diagnostic year as a highly ranked predictive feature are consistent with published studies that show treatment of Medicare FFS adults with low-risk prostate cancer have decreased substantially in the last decade from 86% to 58%(63).

Care fragmentation was among the leading predictors of treatment for low-risk prostate cancer in this study. Adults experiencing care fragmentation and incident prostate cancer may be inclined to “take care of it” (27), as managing non-cancer conditions becomes more complex after a prostate cancer diagnosis.(65) In our
adjusted machine learning models, care fragmentation prior to incident prostate cancer diagnosis was the 3rd highest in SHAP feature importance, suggesting a strong associative prediction of low-value prostate cancer treatment. We also observed that older men with multimorbidity were more likely to receive treatment, consistent with findings of our recently published study indicating that multimorbidity was significantly and positively associated with likelihood of treatment among older men with localized prostate cancer.(49) Multimorbidity adds to the complexity to prostate cancer management decisions given that limited physician and patient resources often compete for management. Specialty healthcare (i.e., urologists and oncologists) need after an incident prostate cancer diagnosis among older men with multimorbidity is known to increase care fragmentation.(24) However, although multimorbidity was associated with higher odds of treatment use in the current study, interactions with care fragmentation were not informative.

Surprisingly, life expectancy was not associated with treatment receipt in this study. Evidence-based guidelines are based on risk group stratification (i.e., PSA, Gleason Score) and life expectancy. Although clinical assessment of life expectancy is a critical element to informed decision making for older adults with low-risk prostate cancer, our study findings suggest that treatment selection might not sufficiently take into consideration life expectancy.(6) We speculate that greater variation in low-value care use between organizations, versus between physicians, may reflect practice patterns associated with physician affiliation and organization management (i.e., compensation, practice guidelines).(66) Findings of other published studies using alternative methods of life expectancy estimation have been inconsistent, with investigations showing both positive (67) and negative relationships (68) between comorbidity burden and CM use in Medicare FFS populations.(49) Taken together, clinical and population differences in comorbidity life expectancy definitions are likely to account for these mixed findings. We speculate that discrepancies may be explained by the use of common use of comorbidity indexes that imprecisely measure life expectancy in prostate cancer populations and likely account for mixed findings.(34,69)

A noteworthy finding of the current study is the apparent role of social determinants (i.e., education and income) in predicting treatment of low-risk prostate cancer. The relationships were complex among individuals living in counties with varying levels of income and education. Socio-demographic factors, such as SEER region, income, college education, and marital status were highly ranked predictors of treatment in this
and many other (27) investigations. Social determinants of health have also been used to accurately predict inpatient and emergency room utilization, demonstrating how social determinants can improve predictive ability of ML models. (70) We found race to be a low-ranking predictor of treatment among low-risk prostate cancer. Unequal access to care in racial minorities may result in poor management and/or overtreatment of low-risk prostate cancer as rates of low-risk prostate cancer treatment among racial groups vary between integrated and fee-for-service cohorts. (71,72) However, we found race to be a low-ranking predictor of treatment among low-risk prostate cancer in this study.

To our knowledge, this study is the first application of machine learning to estimate predictors of prostate cancer treatment among adults with low-risk prostate cancer. Unlike traditional statistical approaches, robust resampling and cross-validation techniques can be used to “train” machine learning algorithms to accurately predict healthcare utilization. More recently, model agnostic interpretations, such as SHAP used in this study, enable feature importance ranking and visualizations of individual-level contributions of patient, clinician, and practice ecosystem factors on target variables. Trade-offs between high performance and interpretability of complex “black-box” algorithms can be mitigated with SHAP feature importance ranking and data visualization. Our approach identified the top predictors of low-value treatment of low-risk prostate cancer and complex non-linear associations among features of high importance. Our results inform clinicians, payers, and policy makers regarding the predictive associations of features articulated in guideline recommendations, namely the lack of predictive importance of life expectancy, and the importance of social determinants, multimorbidity, and care fragmentation.

2.6 Strengths and limitations

Our study has several notable strengths. We used novel machine learning and model agnostic interpretations (SHAP) to predict and understand treatment use among older men with low-risk prostate cancer. We included validated measures of prostate cancer-specific life expectancy to identify older men with low-risk prostate cancer that could benefit from conservative management. Commonly used proxy measures of comorbidity burden are limited in estimating death risk from prostate cancer. We distinguished between comorbidity and life expectancy while previous studies have used comorbidity burden as a surrogate of life expectancy. (34) Using SHAP, we noted both average and individual level variations in CM use which suggests
many older men could benefit from policies and/or interventions designed to reduce low-value prostate cancer treatment.

Our study must be interpreted within the context of several limitations. Foremost, the study is a claims-based, retrospective analysis; therefore, we cannot establish causality and our results are subject to unobservable variable bias and selection bias of paid Medicare claims. To minimize the proportion of missed claims we limited analyses to continuously enrolled in Medicare (A&B, without HMO) throughout the study period. Second, our analysis may not be generalizable to commercial insurance beneficiaries as we only observed Medicare FFS beneficiaries. Third, social determinants of health (i.e., income, education) were not available at the individual level therefore county level measures were used. We were unable to estimate patient preference with SEER-Medicare data, and were thus unable to account for adults with longer life expectancies who prefer curative treatment as conservative management or active surveillance are evidence-based treatment options.

2.7 Conclusions

Using interpretable machine learning approaches provided evidence for several strong, modifiable predictors of low-value treatment of low-risk prostate cancer, including care fragmentation and social determinants of health. Healthcare policies could reduce low-value treatment by addressing fragmentation of care and management of multimorbidity. “Health in all policies” (73) that address upstream features such as income and education should be considered to reduce low-value prostate cancer care associated with significant adverse health related quality of life and cost burdens. Despite a decade’s worth of evidence-based recommendations of observation for adults with low-risk prostate cancer and a life expectancy of more than 10 years, we observed life expectancy to be a low-ranked predictor of treatment. Physician- and patient- focused education may be needed to utilize more precise life expectancy estimates in shared decision making of treatment modalities of low-risk prostate cancer.
Table 2.1 Patient Characteristics by Treatment Use among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer (Linked SEER Cancer Registry and Medicare claims files, 2009-2014 (n= 13,870)

<table>
<thead>
<tr>
<th></th>
<th>CM</th>
<th>Treatment</th>
<th>X²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=4596</td>
<td>33.1</td>
<td>9274</td>
<td>66.9</td>
</tr>
<tr>
<td><strong>Multimorbidity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2 MM</td>
<td>2157</td>
<td>37.0</td>
<td>3678</td>
<td>63.0</td>
</tr>
<tr>
<td>&gt;2 MM</td>
<td>2439</td>
<td>30.4</td>
<td>5596</td>
<td>69.6</td>
</tr>
<tr>
<td><strong>Life expectancy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10+ years</td>
<td>3555</td>
<td>33.5</td>
<td>7057</td>
<td>66.5</td>
</tr>
<tr>
<td>5-10 years</td>
<td>714</td>
<td>31.1</td>
<td>1580</td>
<td>68.9</td>
</tr>
<tr>
<td>&lt;5 LE</td>
<td>327</td>
<td>33.9</td>
<td>637</td>
<td>66.1</td>
</tr>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66-74</td>
<td>3287</td>
<td>30.8</td>
<td>7384</td>
<td>69.2</td>
</tr>
<tr>
<td>75 and above</td>
<td>1309</td>
<td>40.9</td>
<td>1890</td>
<td>59.1</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>3796</td>
<td>32.8</td>
<td>7762</td>
<td>67.2</td>
</tr>
<tr>
<td>Black</td>
<td>497</td>
<td>35.4</td>
<td>908</td>
<td>64.6</td>
</tr>
<tr>
<td>Hispanic</td>
<td>48</td>
<td>22.6</td>
<td>164</td>
<td>77.4</td>
</tr>
<tr>
<td>Others</td>
<td>231</td>
<td>36.7</td>
<td>398</td>
<td>63.3</td>
</tr>
<tr>
<td>Unknown</td>
<td>24</td>
<td>36.4</td>
<td>42</td>
<td>63.6</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmarried</td>
<td>318</td>
<td>35.5</td>
<td>577</td>
<td>64.5</td>
</tr>
<tr>
<td>Married</td>
<td>2843</td>
<td>29.7</td>
<td>6742</td>
<td>70.3</td>
</tr>
<tr>
<td>Sep/div/wid</td>
<td>545</td>
<td>37.6</td>
<td>903</td>
<td>62.4</td>
</tr>
<tr>
<td>Unknown</td>
<td>890</td>
<td>45.8</td>
<td>1052</td>
<td>54.2</td>
</tr>
<tr>
<td><strong>Income quartile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$2,512-43,709</td>
<td>997</td>
<td>30.4</td>
<td>2286</td>
<td>69.6</td>
</tr>
<tr>
<td>$43,711-57,34</td>
<td>1058</td>
<td>32.6</td>
<td>2185</td>
<td>67.4</td>
</tr>
<tr>
<td>$57,350-76,87</td>
<td>1142</td>
<td>34.6</td>
<td>2156</td>
<td>65.4</td>
</tr>
<tr>
<td>$76,888-250,0</td>
<td>1283</td>
<td>34.4</td>
<td>2442</td>
<td>65.6</td>
</tr>
<tr>
<td><strong>Education quartile</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.00-24.0</td>
<td>1192</td>
<td>33.5</td>
<td>2370</td>
<td>66.5</td>
</tr>
<tr>
<td>24-29.0</td>
<td>1066</td>
<td>31.6</td>
<td>2306</td>
<td>68.4</td>
</tr>
<tr>
<td>29.0-34.1</td>
<td>1088</td>
<td>32.4</td>
<td>2269</td>
<td>67.6</td>
</tr>
<tr>
<td>34.1-100</td>
<td>1135</td>
<td>34.8</td>
<td>2130</td>
<td>65.2</td>
</tr>
<tr>
<td><strong>Rural group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metro</td>
<td>3925</td>
<td>33.4</td>
<td>7824</td>
<td>66.6</td>
</tr>
<tr>
<td>Urban</td>
<td>587</td>
<td>32.0</td>
<td>1248</td>
<td>68.0</td>
</tr>
<tr>
<td>Rural</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Unknown</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

Based on 13870 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident low-risk prostate cancer (Gleason 6, PSA <10, Stage ≤T2a) between 2009 and 2014. Age group included 66-74 and 75 and over categories. SEER= Surveillance, Epidemiology, and End Results cancer Registry. **= Cells suppressed in accordance with Data Usage Agreement due to <11 cell count.Bold values denote statistical significance at the P-value < .05 level.
Continued Table 2.1 Patient Characteristics by Treatment Use among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using Linked SEER Cancer Registry, 2009-2014 (n=13,870)

| Metro group       | | 5.253 | 0.072 |
|-------------------|--|--|--|--|
| Metro county      | 3940 | 33.5 | 7827 | 66.5 |
| Non-metro county  | 648  | 31.1 | 1436 | 68.9 |
| **SEER Region**   | | | | |
| Northeast         | 838  | 27.8 | 2171 | 72.2 |
| South             | 1171 | 30.5 | 2666 | 69.5 |
| North Central     | 433  | 32.6 | 896  | 67.4 |
| West              | 2154 | 37.8 | 3541 | 62.2 |
| **Radiation oncology quartile** | | | | |
| 0-0.44            | 1128 | 31.2 | 2485 | 68.8 |
| 0.44-0.99         | 1207 | 32.3 | 2527 | 67.7 |
| 0.99-1.47         | 1085 | 35.2 | 1997 | 64.8 |
| 1.48-22.0         | 1168 | 34.1 | 2254 | 65.9 |
| **Urologist quartile** | | | | |
| 0-1.25            | 1117 | 30.4 | 2562 | 69.6 |
| 1.25-2.38         | 1255 | 34.2 | 2411 | 65.8 |
| 2.38-3.22         | 1039 | 34.3 | 1990 | 65.7 |
| 3.23-28.4         | 1177 | 33.9 | 2300 | 66.1 |
| **Preventative A1c screening** | | | | |
| Yes               | 650  | 34.6 | 1226 | 65.4 |
| No                | 3946 | 32.9 | 8048 | 67.1 |
| **Preventative Flu vaccination** | | | | |
| Yes               | 2113 | 30.1 | 4898 | 69.9 |
| No                | 2483 | 36.2 | 4376 | 63.8 |
| **Preventative lipid screening** | | | | |
| Yes               | 1283 | 33.1 | 2592 | 66.9 |
| No                | 3313 | 33.1 | 6682 | 66.9 |
| **Primary care physician visit** | | | | |
| Yes               | 3618 | 30.6 | 8215 | 69.4 |
| No                | 628  | 38.9 | 988  | 61.1 |

Based on 13870 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident low-risk prostate cancer (Gleason 6, PSA <10, Stage ≤T2a) between 2009 and 2014. Age group included 66-74 and 75 and over categories. SEER= Surveillance, Epidemiology, and End Results cancer Registry. Bold values denote statistical significance at the P-value < .05 level.
Figure 2.1 Feature Importance of Treatment Prediction using SHAP values among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870)

Top 15 mean SHAP values in descending order of log-odds. Based on 13870 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident low-risk prostate cancer (Gleason 6, PSA <10, Stage ≤T2a) between 2009 and 2014. Age group included 66-74 and 75 and over categories.
Care fragmentation = Bice-Boxerman continuity of care index to calculate care fragmentation during the 12 month baseline period (See Methods).
SEER= Surveillance, Epidemiology, and End Results cancer Registry.
**Figure 2.2** Positive and Negative Feature Relationships of Treatment Prediction among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870)

Mean SHAP values in descending order of log-odds. Based on 13870 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident low-risk prostate cancer (Gleason 6, PSA <10, Stage ≤T2a) between 2009 and 2014.

- Age group 66-74 vs. adults over age 74
- Multimorbidity group = 3 or more chronic conditions.
- Care fragmentation = Bice-Boxerman continuity of care index to calculate care fragmentation during the 12 month baseline period (See Methods).
- SEER= Surveillance, Epidemiology, and End Results cancer Registry.
**Figure 2.3** SHAP Partial Dependence Plot of Top 4 Predictive Features of Treatment among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870)

SHAP Partial Dependence plots (PDP) of log-odds SHAP values by (A.) Marital status “Married”, (B.) Age group 66-74, (C) Care fragmentation, and (D) Median Income. Based on 13870 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident low-risk prostate cancer (Gleason 6, PSA <10, Stage ≤T2a) between 2009 and 2014.

Age group 66-74 vs. adults over age 74
Care fragmentation = Bice-Boxerman continuity of care index to calculate care fragmentation during the 12 month baseline period (See Methods).
Median income = County-level median income from Area Health Resource File linkage.
SHAP = Shapley Additive ePlanations.
SEER= Surveillance, Epidemiology, and End Results cancer Registry.
**Figure 2.4** SHAP Partial Density Plot of Life Expectancy and Multimorbidity Predictive Features on Treatment among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870)

SHAP Partial Dependence plots (PDP) of log-odds SHAP values by (A.) PCCI “Life expectancy over 10 years”, (B.) “Multimorbidity”. Based on 13870 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident low-risk prostate cancer (Gleason 6, PSA <10, Stage ≤T2a) between 2009 and 2014.

PCCI = Prostate Cancer Comorbidity Index.
Multimorbidity defined as 1 = *two or fewer* non-cancer chronic conditions, 2 = *three or more* non-cancer chronic conditions.

SHAP = Shapley Additive ePlanations.
SEER = Surveillance, Epidemiology, and End Results cancer Registry.
CHAPTER 3

3 Associations of Multimorbidity and Patient-reported Experiences of Care with Conservative Management among Elderly Patients with Localized Prostate Cancer

3.1 Abstract

**Background:** Many elderly localized prostate cancer patients could benefit from conservative management (CM). This retrospective cohort study examined associations of patient-reported access to care and multimorbidity on CM use patterns among Medicare Fee-for-Service (FFS) beneficiaries with localized prostate cancer.

**Methods:** We used linked Surveillance, Epidemiology and End Results cancer Registry, Medicare Claims, and the Medicare Consumer Assessment of Healthcare Providers and Systems survey files. We identified FFS Medicare Beneficiaries (Age ≥ 66; continuous enrollment in Parts A & B) with incident localized prostate cancer from 2003–2013 and a completed MCAHPS surveys within 12-months of diagnosis (n=496). We used multivariable models to examine MCAHPS measures (getting needed care, timeliness of care, and doctor communication) and multimorbidity on CM use.

**Results:** Localized prostate cancer patients with multimorbidity were less likely to use CM (adjusted odds ratio (AOR)=0.42 (0.27-0.66), P<0.001); those with higher scores on timeliness of care (AOR=1.21 (1.09, 1.35), P<0.001), higher education attainment (3.21=AOR (1.50,6.89), P=0.003), and impaired mental health status (4.32=AOR (1.86, 10.1) P<0.001) were more likely to use CM.

**Conclusion(s):** Timeliness of care was significantly and positively associated, and multimorbidity, significantly and inversely associated with CM use. Addressing specific modifiable barriers to timely care along the cancer continuum for elderly localized prostate cancer patients with limited life expectancy, and the reasons for lower CM use among those with multimorbidity, could reduce the adverse effects of overtreatment on health outcomes and costs.

**Text pages:** 17, Tables: 4, and 2 Figures

**Keywords:** prostate cancer, conservative management, active surveillance, multimorbidity, patient-centered care.
3.2 Introduction

Conservative management (CM) has emerged as a common disease management approach for older adults with localized prostate cancer. (63) CM use is supported by high-level evidence for localized prostate cancer patients with low or favorable intermediate risk disease or higher risk disease with limited life expectancy. (6) CM includes protocols for low or intermediate risk prostate cancer, such as follow-up biopsies and PSA testing, or “watchful waiting” for patients with less than 5 years of life expectancy. Use of CM among patients with low-risk prostate cancer or limited life expectancy improves health-related quality of life (ie, urinary, bowel, and/or erectile dysfunction) and could reduce excessive annual health-care costs of overtreatment by $1.2 billion. (10, 64)

CM decisions are complex as 60% of older adults (age > 65 years) with localized prostate cancer have pre-existing multimorbidity. (47, 48) Multimorbidity affects life expectancy (47) and more than 50% of patients with multimorbidity seek care from three or more healthcare specialists. (50) Patients with multimorbidity and cancer may be vulnerable to poor quality of cancer care and have prompted greater attention in measuring, monitoring, and incentivizing patient-centered care. (15) Measures of patient-centered care, such as patient-reported experiences with care include domains of physician communication, timely care, and perceptions of getting needed care, are increasingly used as quality measures by health plans, medical groups, and physician practices. Positive patient-reported experience scores are associated with adherence to medical advice, improved clinical outcomes, and lower utilization of unnecessary health-care services (29, 74) such as overtreatment of low-risk localized prostate cancer.

Patient-reported experiences may differ by multimorbidity status, which may further complicate or facilitate treatment choices for low-risk prostate cancer. (75) Identifying specific measures of patient-reported experiences that facilitate CM use among patients with incident localized prostate cancer and multimorbidity is needed to promote evidence-based cancer care. (30) For example, in colorectal cancer populations, patient-reported experiences with perceived timely care are associated with evidenced-based follow-up. (31) Understanding the relationship between patient-reported experiences of care on CM use can inform patient-centered care approaches to improve adoption of CM use, thereby reducing the adverse effects of overtreatment among older patients with multimorbidity and localized prostate cancer.
Despite the importance of patient-reported experiences, CM studies primarily focus on disease characteristics, clinical, and sociodemographic factors. To date, no studies have investigated the impact of patient-reported experiences on CM use among medically complex patients with localized prostate cancer. Therefore, the primary objective of this study is to examine the associations of multimorbidity and patient-reported experiences on CM use among Fee-for-Service (FFS) Medicare beneficiaries with localized prostate cancer using Consumer Assessment of Healthcare Providers and Systems (MCAHPS®) patient surveys and Medicare claims linkages.

3.3 Methods

The study cohort included men diagnosed with localized prostate cancer defined as American Joint Committee on Cancer stage T2a or less, aged 66 or older, with continuous enrollment in FFS Medicare Parts A and B throughout the study period (Figure 3.1).

Date of incident localized prostate cancer diagnosis was used as an index date and 12 months before diagnosis was used as the baseline period. During the baseline period, we identified multimorbidity using Medicare claims and calculated life expectancy estimates.

We also defined the “CM measurement” period as 12 months after diagnosis. During this period, we identified CM based on validated methods for claims data.

As MCAHPS surveys can be administered at varying points during the post diagnosis period, we followed individuals for an additional period of 12 months. Thus, our follow-up period consisted of 24 months after incident localized prostate cancer diagnosis.

To account for varying months from diagnosis to survey administration, we included time from diagnosis to survey as one of the independent variables in the models. However, as this variable was not significant and did not affect our main results, we did not include time from diagnosis to survey administration variable in the final model. As a sensitivity analysis, we also estimated CM use during 24 months after diagnosis (Appendix 7.3).

Data sources

The Surveillance, Epidemiology and End Results (SEER) cancer registry contains tumor and demographic information for patients diagnosed with cancer while residing in a SEER region. We derived
Medicare eligibility from the SEER data (Figure 3.1). We extracted fee-for-service Medicare claims from Medicare Provider Analysis and Review (MEDPAR), Carrier Claims, Outpatient Claims, Home Health Agency, and Durable Medical Equipment files.

Medicare Consumer Assessment of Healthcare Providers & Systems (MCAHPS®) surveys, administered by the CMS, use standardized and validated questionnaires to collect information on patient-reported experiences with health-care providers. MCAHPS collection methodologies use a weighted probability sampling procedure covering all the 50 US states, DC, and Puerto Rico, which are then linked to SEER patients.

Area Health Resource File (AHRF) files were linked via MEDPAR FIPS state and county codes and were used to calculate radiation oncologist and urologist densities. Census files were linked to calculate county-level median income quartiles.

**Dependent Variable**

We operationalized CM use as the absence of curative treatment within 12 months after incident localized prostate cancer. Treatment was identified using International Classification of Diagnosis 9th edition (ICD9), ICD9 procedure codes, and Healthcare Common Procedure Coding System (HCPCS) codes from FFS Medicare claims (Appendix 7.4).

**Key Independent Variables**

The multimorbidity framework developed by the United States Department of Health and Human Services for guiding programs, practice, and policy guided the selection of chronic conditions as follows: arthritis, asthma, coronary artery disease, congestive heart failure, chronic kidney disease, chronic obstructive pulmonary disease, cardiac arrhythmias, acute myocardial infarction, dementia, diabetes, depression, hepatitis, hyperlipidemia, hypertension, human immunodeficiency virus, osteoporosis, substance abuse, schizophrenia, stroke, anemia, and lower limb fracture. The most common definition of multimorbidity is the concurrent presence of two or more conditions in the same individual. We defined multimorbidity as the presence of three or more conditions in the same individuals as older men diagnosed with prostate cancer at age 65 or higher are at high risk for competing risk mortality. For example, among men with three or more
comorbid conditions, aged 61-74 and 75 years or older, 10-year other cause mortality is 40% and 71%, respectively.(77)

*Prostate cancer comorbidity index* (PCCI), a weighted comorbidity index validated prostate cancer patient populations, was used to predict 5- and 10-year life expectancy in prostate cancer patient populations.(77) PCCI was calculated during the baseline period to estimate 5- and 10-year life expectancy. PCCI was categorized into three groups: 0–8 (>10-year life expectancy); 9 to 13 (5- to 10-year life expectancy); and > 13 (<5-year life expectancy). In all models, PCCI total 0-8 was used as the reference group.

Published research in prostate cancer patients often uses Charlson Comorbidity Index (CCI); therefore, we conducted a supplemental analysis using CCI. In these analyses, CCI scores were dichotomized with “0-1” as the reference group.(78)

We included three MCAHPS composite measures—“getting needed care,” “getting care quickly,” and “doctor communication”—which rates the ability to get needed appointments, timeliness of care when care is needed, and how well the physician communicated. Patients report experiences with health-care access in the last 6 months. MCAHPS surveys have been extensively validated for measuring patient-reported access to care and are commonly used for quality improvement as well as value-based purchasing initiatives.(40) MCAHPS are based on a 0-100 scale with 0 representing the lowest and 100 representing the highest score; we examined the effect of 10 unit changes in composite items on the dependent variable.

Management of preexisting multimorbidity and shared prostate cancer treatment decision-making requires the use of limited resources (ie, time to manage chronic conditions and availability of health-care professionals and resources). Therefore, for other independent variables, the competing demands model was used to conceptualize factors known to affect localized prostate cancer treatment selection within clinician, patient, and practice ecosystem domains.(38,53,79) (Figure 3.2).

Multivariable models were adjusted with independent variables: diagnosis year group (2003-2009 and 2010-2013), low-risk prostate cancer (operationalized as Gleason Score ≤ 6 and PSA test ≤ 10 ng/mL or Gleason Score>=7 or PSA>= 10 ng/mL), self-reported general and mental health status, education-level, zip-code income quartiles, and county-level quartiles of urologists and radiation oncologists per 10,000 persons over age 65.
Our analyses include several case-mix adjustment variables such as age, education, general health status, mental health status, income level, and race. Secondary analyses using additional recommended case-mix adjustment variables, such as dual-eligible Medicaid respondents, “proxy” survey completion, and time from cancer diagnosis to survey completion, did not significantly improve model specification.(80)

To assess the potential influence of missing data, we examined missing data patterns using covariate-dependent missingness methods.(81) Mean values were imputed to independent variables of interest. For categorical variables, including general and mental health status, missing data indicators were created and included as a separate category in the regression models.

Chi-square tests and t tests were used to identify significant group differences in CM use by categorical variables. Multivariable models were fit using separate unadjusted and adjusted logistic regressions to identify independent and interactive associations of multimorbidity and patient care experiences on CM use. All statistical tests were two-sided with a 5% Type I error rate and were completed in STATA (StataCorp, College Station, TX).

3.4 Results

The study cohort was predominantly non-Hispanic, whites (84.5%). The median age at diagnosis was 72.8 years and did not differ by year of diagnosis (2003-2009:73.6 = M, 5.38; 2010-2013:73.6 = M, SD = 5.14). Average composite scores for doctor communication, getting needed care, and getting care quickly were 91.0 (SD = 12.2), 88.6 (SD = 15.6), and 70.8 (SD = 21.7), respectively.

Overall, 33.5% used CM, defined as no curative treatment within 12 months of incident localized prostate cancer diagnosis. Use of CM was only marginally higher in men with low-risk relative to those with higher-risk disease (≤ cT2a and PSA>= 10 ng/mL or Gleason Score > 6) (38.7% vs 30.9%, respectively, P = .08) (Table 3.1). High-school graduation, college education, low-risk prostate cancer diagnosis, and mental health status were significantly more frequent among patients using CM (Table 3.1). CM use by localized prostate cancer patients with higher-risk disease was 30.9%. Higher-risk disease was significantly more common among age groups 75+ (75.9%) vs 66-74 (62.4%) (P = .002) and significant differences by patients with multimorbidity 24.6% (n = 45) vs those without multimorbidity 38.7% (n = 58), (Χ2 = 7.65, P = .006) were observed.
In our study cohort, 57.2% had multimorbidity. Patients 75 years or older were significantly more likely to have multimorbidity than those aged 66-74 years (64.4% vs 53.4%). Blacks had a higher percentage of multimorbidity as compared to whites (76.1% vs 53.9%). Patients with multimorbidity using treatment (n = 207) did not differ significantly by patient, clinician, or practice ecosystem factors except for mental health status of excellent/very good (74.6%) and good (77.5%) vs patients using CM (P = .031). Patients with multimorbidity and higher-risk disease (n = 183) significantly more frequently used treatment if aged 66-74 (82.5%) (P = .011). Average composite scores for doctor communication, getting needed care, and getting care quickly did not differ by multimorbidity status. CM use was significantly lower in men with vs without multimorbidity (27.1% vs 72.9%, respectively, P < .001).

Average getting care quickly composite scores (ie, timely care) were higher for those with CM use as compared to those without CM use (Table 3.2). CM use significantly differed by PCCI categories, with lower percentages among groups with less than 10 (27.0%) and 5 (27.0%) vs more than 10 (38.9%) years of life expectancy ($X_2 = 7.82, P = .020$) (Table 3.2).

PCCI life expectancy groups did not statistically differ by CM use. Higher-risk patients reporting fair or poor mental health status (62.1%; P = .002) vs excellent mental health status, and college education (33%) or high-school graduates (37.5%) vs no high-school graduation (13.7%), significantly used CM more frequently. Getting care quickly composite scores were significantly higher among higher-risk patients (n = 333) using CM (M = 75.8) vs curative treatment (M = 69.8), ($t=−2.43, CI 95\% = 69.3, 73.9, P = .016$).

Multimorbidity was significantly and inversely related to CM use in unadjusted logistic regression analyses (odds ratios (OR) = 0.55; 95% CI = 0.35, 0.75). Adjustment for other factors, including timeliness of care, further strengthened this association (adjusted OR (AOR) = 0.42, CI 0.27- 0.66) (Table 3.3); additional models adjusting for other patient experience domains or CCI were not significant. Getting care quickly showed a significant, positive association with CM use in both the unadjusted analyses (OR = 1.15; 95% CI = 1.05, 1.27) and the fully adjusted models (AOR = 1.21; 95% CI = 1.09, 1.34). In models including PCCI life expectancy categories, less than 10- and 5-year life expectancy were inversely associated with CM use (Appendix 7.2).

CM use was also significantly and positively associated with fair/poor mental health status, low-risk prostate cancer diagnosis, college education or more, and high-school graduation in all adjusted models (Table
3.3; Appendix 7.1 and 7.2). We found no evidence for a modifying effect of patient-experience variables, multimorbidity, PCCI, or other independent variables on the observed associations.

3.5 Discussion

In this study, we assessed the independent associations of multimorbidity and patient-reported experiences with care on CM use among older men with localized prostate cancer. Despite proven benefits of CM, one in three (33.5%) of all men with localized prostate cancer, and only two in five (41%) men over the age of 75 years, used CM. Our estimates of CM use among patients with localized prostate cancer are lower than those reported in recent investigations using SEER-Medicare data (42.1% in 2015) (63) but higher than reported in an investigation of Michigan Medicare beneficiaries (22.3% in 2014). (53) We speculate that these differences could be due to variation in study population characteristics (ours included prostate cancer patients from many regions of the US) and geographic practice patterns. (82–84)

Multimorbidity and life expectancy are critical components of patient counseling after a localized prostate cancer diagnosis as many older men do not live long enough to benefit from treatment. Patients with low or favorable intermediate-risk disease or higher-risk disease with limited life expectancy could avoid significant urinary, erectile, and rectal treatment morbidities without increasing the risk of prostate cancer-specific mortality with CM. (85, 86) However, in our study, men with multimorbidity were significantly less likely to use CM compared to those without multimorbidity after controlling for age, low-risk prostate cancer, and other sociodemographic variables. We speculate that men with multimorbidity and low-risk cancer may not opt for treatment because they may have a preference for immediate cure (ie, “take care of it”) (87) and may not want to add one more condition that requires long-term management. Furthermore, men with multimorbidity may fear nontreatment regret, (88) emotional distress, (89) and anxiety. (90) Strong multidisciplinary management strategies, including significant psychological support from primary care physicians and specialists (ie, urologist and/or medical and radiation oncologists), are needed to mitigate decisional conflict (91) and therefore facilitate CM use. (27)

In adjusted models, including validated life expectancy measures for prostate cancer survivors, patients meeting evidence-based criteria2 for CM were 58% less likely to use CM based on life expectancy alone (ie, less than 5 years). Previous studies using CCI report both positive and negative relationships between comorbidity
burden and CM use in Medicare FFS populations.(67,68,92) In a supplemental analysis in this study, CCI was not significantly associated with CM use. Taken together, these findings suggest that clinical and population differences in comorbidity definitions are likely to account for mixed findings in several previous investigations.(69) By defining multimorbidity using a list of conditions prioritized by policy makers in the US,(55) our study made a unique contribution to this field. However, as pointed out by a systematic review that current life expectancy prediction tools lack both practical and theoretical utility,(34) comorbidity measures that can be easily operationalized in a clinical setting are needed. Recently, age-adjusted indexes, such as the PCCI used in our study, were developed to provide life expectancy estimates in patients with prostate cancer.(54) Certain types (cardiovascular disease) and combinations (cardiometabolic and respiratory)(65) of chronic conditions are associated with treatment regardless of patient, clinician, and health-care ecosystem factors. Additional research is needed to understand the relationship between more precise estimates of life expectancy and multimorbidity on CM use in FFS Medicare populations.

In our study, patient-reported experiences, specifically timeliness of care, were positively associated with CM use. Patients with higher timeliness of care scores were significantly more likely to use CM after adjusting for demographic, clinical, socioeconomic, and health-care system factors. Timely access to care for localized prostate cancer patients is not limited to initial diagnosis of prostate cancer, but the opportunity and ease by which a patient is able to utilize needed services along the continuum of care throughout survivorship.(93) Choices for elderly localized prostate cancer patients involve selecting curative and non-curative treatments with trade-offs in efficacy, potential adverse quality of life effects, and competing risk mortality. MCAHPS timeliness of care domains, such as perceived barriers to appointment scheduling, are fundamental to shared decision-making among multiple health-care providers that significantly influence localized prostate cancer treatment choice.(94–96) We speculate that patients with higher timely care ratings may choose CM because they may have a favorable perception of health-care system capacity to provide services once a need is detected.

Our study findings have important policy implications. Currently, no value-based mechanisms exist to support the use of CM in Medicare FFS populations. Existing literature also suggest that CM use in FFS system varies among physician practices by 5.1%-71.2%.(84) Emerging oncology care models by CMS may need to
incorporate risk-adjusted CM use as a quality indicator along the cancer care continuum(97) to promote CM use among men with localized prostate cancer. Recently, the NCCN Quality and Outcomes committee identified significant gaps in evaluating high-quality cancer care with patient experience measures and evidence-based practice.(85) More research is needed to identify specific barriers to timely care and how validated patient-reported experience measures can be used to support evidence-based management of localized prostate cancer patients in oncology care models.

We also observed that elements of social determinants, such as education, were associated with CM use. Although educational attainment may not be modifiable among older adults, initiatives such as “health in all policies” by World Health Organization and the Centers for Disease and Prevention Control “integrate and articulated health considerations” into community health policy.(73) These experts concluded that social, economic, and physical environments have a significant impact on the health of an individual and these effects should be considered in the development of all public policies and programs.

3.6 Strengths and Limitations

The SEER-CAHPS data linkage is a robust and unique resource that provides an ideal opportunity to study patient-centered care delivery of contemporary treatment patterns among patients with localized prostate cancer and multimorbidity. We build on previous findings by including validated estimations of life expectancy and definitions of multimorbidity to access the impact of comorbid conditions on patterns of contemporary treatment options for older localized prostate cancer patients.

Our study results must be interpreted with important limitations. MCAHPS surveys request patient-reported experiences with care “in the last 6 months”.(40) Due to relatively small sample size, we included surveys completed within 6 months after localized prostate cancer diagnosis which may overlap with the baseline period. However, our results were robust to multivariable adjustments for time between cancer diagnosis and survey completion. Due to MCAHPS survey administration processes and collection, we cannot directly attribute MCAHPS composite ratings to physician specialty or the prostate cancer diagnosis; instead, our results are generalizable to the entire patient experience after diagnosis which may include multiple care providers for multiple conditions. The study sample comprised of predominantly non-Hispanic white, urban
adults, potentially limiting generalizability to ethnic minorities, rural, or other populations. Our study was restricted to Medicare FFS enrollees 65 years or older and may not be generalizable to younger adults or individuals on private insurance. Lastly, due to sample size limitations, we were unable to analyze the relationship of individual chronic conditions with CM use.

3.7 Conclusions

Our results highlight the effect of patient-reported experiences, multimorbidity, and life expectancy on CM use among older men with localized prostate cancer. While factors such as multimorbidity and life expectancy are critical clinical components that may affect the choice of CM vs over treatment, our study highlights the role of nonclinical factors, specifically patient-reported experiences with care on treatment of localized prostate cancer. Our findings support a “population health-based” oncology care model in which both clinical and nonclinical factors, such as patient-reported experiences, are integrated to promote CM use and avoid overtreatment among older men with localized prostate cancer.
### Table 3.1 Patient Characteristics by Conservative Management among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (n=496)

<table>
<thead>
<tr>
<th></th>
<th>CM</th>
<th>No CM</th>
<th>X²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td><strong>ALL</strong></td>
<td>166</td>
<td>33.5</td>
<td>330</td>
<td>66.5</td>
</tr>
<tr>
<td><strong>Age in Years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66-74</td>
<td>102</td>
<td>31.7</td>
<td>220</td>
<td>68.3</td>
</tr>
<tr>
<td>75+</td>
<td>64</td>
<td>36.8</td>
<td>110</td>
<td>63.2</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>140</td>
<td>33.4</td>
<td>279</td>
<td>66.6</td>
</tr>
<tr>
<td>Black</td>
<td>15</td>
<td>32.6</td>
<td>31</td>
<td>67.4</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>35.5</td>
<td>20</td>
<td>64.5</td>
</tr>
<tr>
<td><strong>Marital Status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>115</td>
<td>32.9</td>
<td>235</td>
<td>67.1</td>
</tr>
<tr>
<td>Unmarried</td>
<td>22</td>
<td>34.9</td>
<td>41</td>
<td>65.1</td>
</tr>
<tr>
<td>Unknown</td>
<td>29</td>
<td>34.9</td>
<td>54</td>
<td>65.1</td>
</tr>
<tr>
<td><strong>Income quartiles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>38</td>
<td>33.3</td>
<td>76</td>
<td>66.7</td>
</tr>
<tr>
<td>Second</td>
<td>39</td>
<td>33.6</td>
<td>77</td>
<td>66.4</td>
</tr>
<tr>
<td>Third</td>
<td>38</td>
<td>31.1</td>
<td>84</td>
<td>68.9</td>
</tr>
<tr>
<td>Four</td>
<td>51</td>
<td>35.4</td>
<td>93</td>
<td>64.6</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College or more</td>
<td>100</td>
<td>36.6</td>
<td>173</td>
<td>63.4</td>
</tr>
<tr>
<td>High School Grad.</td>
<td>42</td>
<td>35.9</td>
<td>75</td>
<td>64.1</td>
</tr>
<tr>
<td>No High School Grad.</td>
<td>13</td>
<td>18.8</td>
<td>56</td>
<td>81.2</td>
</tr>
<tr>
<td><strong>General health status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent/Very Good</td>
<td>59</td>
<td>35.5</td>
<td>107</td>
<td>64.5</td>
</tr>
<tr>
<td>Good</td>
<td>54</td>
<td>28.6</td>
<td>135</td>
<td>71.4</td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>47</td>
<td>37.6</td>
<td>78</td>
<td>62.4</td>
</tr>
<tr>
<td><strong>Mental health status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent/Very Good</td>
<td>104</td>
<td>31.6</td>
<td>225</td>
<td>68.4</td>
</tr>
<tr>
<td>Good</td>
<td>34</td>
<td>30.1</td>
<td>79</td>
<td>69.9</td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>22</td>
<td>57.9</td>
<td>16</td>
<td>42.1</td>
</tr>
</tbody>
</table>

Bold values denote statistical significance at the $P$-value < .05 level.
Based on 496 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident localized prostate cancer between 2003 and 2013.
Abbreviations: CM, Conservative management; MCAHPS, Medicare Claims and the Medicare Consumer Assessment of Healthcare Providers and System surveys; SEER, Surveillance, Epidemiology, and End Results cancer Registry.

36
Table 3.1 “continued”

<table>
<thead>
<tr>
<th>Urologist density</th>
<th>4.99</th>
<th>0.173</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 1.41</td>
<td>41</td>
<td>33.1</td>
</tr>
<tr>
<td>1.41 to 2.49</td>
<td>33</td>
<td>26.6</td>
</tr>
<tr>
<td>2.5 to 3.46</td>
<td>51</td>
<td>39.8</td>
</tr>
<tr>
<td>3.47 to 10.2</td>
<td>41</td>
<td>34.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Radiation oncologist density</th>
<th>5.43</th>
<th>0.143</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 0.44</td>
<td>37</td>
<td>29.8</td>
</tr>
<tr>
<td>0.45 to 1.07</td>
<td>41</td>
<td>33.1</td>
</tr>
<tr>
<td>1.07 to 1.49</td>
<td>52</td>
<td>41.6</td>
</tr>
<tr>
<td>1.51 to 5.35</td>
<td>36</td>
<td>29.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SEER region</th>
<th>5.09</th>
<th>0.166</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td>33</td>
<td>35.5</td>
</tr>
<tr>
<td>South</td>
<td>35</td>
<td>30.4</td>
</tr>
<tr>
<td>North-central</td>
<td>13</td>
<td>22.4</td>
</tr>
<tr>
<td>West</td>
<td>85</td>
<td>37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metro Status</th>
<th>0.71</th>
<th>0.401</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metro</td>
<td>138</td>
<td>34.3</td>
</tr>
<tr>
<td>Non-Metro</td>
<td>28</td>
<td>29.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis Year</th>
<th>0.44</th>
<th>0.509</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002-2007</td>
<td>96</td>
<td>32.3</td>
</tr>
<tr>
<td>2008-2013</td>
<td>70</td>
<td>35.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low-risk prostate cancer</th>
<th>2.93</th>
<th>0.087</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>63</td>
<td>38.7</td>
</tr>
<tr>
<td>No</td>
<td>103</td>
<td>30.9</td>
</tr>
</tbody>
</table>

Bold values denote statistical significance at the *P*-value < .05 level.

Based on 496 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident localized prostate cancer between 2003 and 2013.

Abbreviations: CM, Conservative management; MCAHPS, Medicare Claims and the Medicare Consumer Assessment of Healthcare Providers and System surveys; SEER, Surveillance, Epidemiology, and End Results cancer Registry.
<table>
<thead>
<tr>
<th></th>
<th>CM</th>
<th>No CM</th>
<th>X²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multimorbidity</strong></td>
<td>166</td>
<td>330</td>
<td>12.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Yes</td>
<td>77</td>
<td>207</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>89</td>
<td>123</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PCCI</strong></td>
<td></td>
<td></td>
<td>7.82</td>
<td>0.020</td>
</tr>
<tr>
<td>&lt; 5 years life expectancy</td>
<td>20</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 5 and &lt; 10 years life expectancy</td>
<td>41</td>
<td>111</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 10 years life expectancy</td>
<td>105</td>
<td>165</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Getting Needed Care</td>
<td>87.1</td>
<td>89.3</td>
<td>1.48</td>
<td>N.S.</td>
</tr>
<tr>
<td>Getting Care Quickly</td>
<td>75.0</td>
<td>68.7</td>
<td>-3.06</td>
<td>0.002</td>
</tr>
<tr>
<td>Doctor/Patient Communication</td>
<td>89.8</td>
<td>91.6</td>
<td>1.61</td>
<td>N.S.</td>
</tr>
</tbody>
</table>
Table 3.3 Unadjusted (UOR), Adjusted Odds Ratios (AOR) and 95% Confidence Intervals (CI) of Multimorbidity, Mental Health, Timeliness of Care, Low-risk Prostate Cancer, and Education on Likelihood of Conservative Management Use among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (N = 496).

<table>
<thead>
<tr>
<th></th>
<th>UOR [95% CI]</th>
<th>AOR [95% CI]</th>
<th>p-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient Experience: Getting Care Quickly</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimorbidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.51</td>
<td>0.42</td>
<td>0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Getting Care Quickly</td>
<td>1.15</td>
<td>1.21</td>
<td>0.003</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Low-risk prostate cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.41</td>
<td>1.76</td>
<td>0.088</td>
<td>0.01</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental Health Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>2.97</td>
<td>4.32</td>
<td>0.002</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ex/VG/Good (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College or more</td>
<td>2.49</td>
<td>3.21</td>
<td>0.006</td>
<td>0.003</td>
</tr>
<tr>
<td>High-school graduate</td>
<td>2.41</td>
<td>3.53</td>
<td>0.015</td>
<td>0.002</td>
</tr>
<tr>
<td>No high-school graduation (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Patient Experience: Getting Need Care</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimorbidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.088</td>
<td>0.45</td>
<td>0.088</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Getting Needed Care</td>
<td>-</td>
<td>1.00</td>
<td>N.S.</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>Patient Experience: Doctor Communication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimorbidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.088</td>
<td>0.45</td>
<td>0.088</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctor Communication</td>
<td>-</td>
<td>1.00</td>
<td>N.S.</td>
<td>N.S.</td>
</tr>
</tbody>
</table>
**Figure 3.1** Cohort Selection and Exclusion

- SEER-Medicare Fee-for-service (FFS) 2002-2013 14,758
- No diagnosis at autopsy 11624
- Prostate cancer only 8496
- Continuous enrollment in Medicare FFS A and B 4660
- AJCC stage ≤T2a 5734
- MCAHPS surveying within 24 months post diagnosis 496
Figure 3.2 Adapted Competing Demands Framework

Clinician domain
Life expectancy, Multimorbidity, cancer risk group

Patient domain
Age, race, marital status, education

Practice ecosystem
Diagnosis year, SEER region, Urologist density, Radiation oncologist density, metro status

Study Outcomes
CHAPTER 4

4 Healthcare Expenditures Associated with Low-value Care among Older Men with Incident Localized Prostate Cancer: Statistical and Machine Learning Approaches

4.1 Abstract

Background: Older adults with incident localized prostate cancer represent a large, medically complex population at risk of low-value care in the form of overtreatment. Conservative management (CM), or non-treatment of prostate cancer with little cancer specific mortality risk, can reduce the adverse effects of overtreatment. Other types of low-value care services are associated with high, compounding “downstream” costs and may be useful for predicting cancer survivorship costs.

Research Objective: Use statistical and machine learning approaches to estimate the association of low-value care use prior to incident prostate cancer diagnosis on average non cancer related total healthcare expenditures 12-24 months after diagnosis. Generalized linear models (GLM) with log-link and gamma distribution were used to estimate total healthcare cost. Predictive feature importances of healthcare costs and non-linear relationships were estimated using the XGBoost Machine Learning (ML) algorithm and SHapley Additive exPlanations (SHAP).

Study Design: Retrospective cohort study.

Population Studied: We used linked Surveillance, Epidemiology, and End Results cancer Registry, Medicare Claims, Census, and Area Health Resource files identify Medicare Fee-for-Service beneficiaries with continuous enrollment in Parts A & B and incident localized prostate cancer from 2005 to 2014 (n=75671). We identified 20 low-value care measures 12 months prior to localized prostate cancer diagnosis using SEER-Medicare registry and claims files. Conservative management was defined as no treatment 0-12 months after diagnosis. Non cancer related total healthcare expenditures were defined as all costs independent of cancer treatment modalities 12-24 months after diagnosis.

Principal Findings: Overall, 25.2% used low-value care services in 12 months prior to cancer diagnosis. Screening for asymptomatic carotid artery stenosis, Magnetic Resonance Imaging (MRI) for low-back pain, and traction for low-back pain were the most frequent low-value care procedures. Overall 18.1% used CM and
use of any low-value care was lower among patients using CM (15.2%), with MRI for low-back pain, traction for low-back pain, and asymptomatic carotid artery stenosis more frequent among patients using treatment. In adjusted GLM models, every 1-unit increase in low-value care use was associated with an $858.82 increase in healthcare expenditures. Multimorbidity (physical condition count), care fragmentation, and conservative management use were the top 3 predictors of healthcare expenditures with the highest absolute mean SHAP values.

**Conclusions:** We found incremental low-value healthcare use before incident prostate cancer diagnosis was significantly and positively associated with long-term non-cancer related costs. Using ML, both low-value care and CM use were highly ranked important features of cost prediction 12-24 months after incident prostate cancer diagnosis.
4.2 Introduction

Low-value care, defined as unnecessary tests or treatment lacking a net clinical benefit to the patient, contribute to avoidable morbidities and excessive healthcare costs.(17) Low-value care is prevalent among elderly patients, affecting 1 in 4 Medicare beneficiaries(18) and accounting for $75.7 - $101.2 billion in annual healthcare costs.(19,20) Use of low-value healthcare services among medically complex patients can lead to cascades of unnecessary down-stream care, leading to costs as high as 10 times the original low-value healthcare service.(21)

Adults with incident localized prostate cancer represent a large, medically complex population at risk of low-value care in the form of overtreatment. Older adults with localized prostate cancer have high rates of multimorbidity and often experience care fragmentation that present barriers to shared decision making and evidence-based recommendations.(25) These adults may be at risk for receiving low-value care, augmenting down-stream costs. Many studies have examined comparative costs between different prostate cancer treatment modalities including conservative management (CM) (32); however, long term costs associated with overtreatment or low-value care remain unknown.

Recently, machine learning algorithms have been used to identify high-cost patients (98), including, for example, high-cost breast cancer patients (99). However, the potential predictive effects of low-value care on non-cancer related total healthcare expenditures among medically complex localized prostate cancer survivors remains little explored.

The primary objective of this paper is to examine whether low-value care is a leading predictor of long-term (12-24 months after diagnosis) non-cancer related expenditures among older adults with incident localized prostate cancer using machine learning algorithms and model agnostic interpretations. We used traditional generalized linear models to examine the association of life expectancy, high- and low- value care use, and care fragmentation on healthcare expenditures 12-24 months after the diagnosis of localized prostate cancer.

4.3 Methods

Design
This study used a retrospective longitudinal design with a 12-month baseline and a 24-month follow-up period. The baseline period comprised the 12-month period before date of localized prostate cancer diagnosis. The follow-up period included the 24-month after localized prostate cancer diagnosis. Healthcare expenditures were calculated during the second half of the follow-up period (12-24 months).

**Data sources**

This study used the Surveillance, Epidemiology, and End Results (SEER)-Medicare linked data files. SEER-Medicare data includes data on all incident prostate cancer cases occurring in 18 regions throughout the United States. Incident prostate cancer, prostate cancer specific clinical information (i.e., Gleason Score and Prostate-specific antigen values), and cancer staging information were derived from the SEER database. Medicare claims were derived from inpatient, outpatient, and other files linked to SEER.

Area Health Resource File (AHRF) files were linked via MEDPAR FIPS state and county codes and were used to calculate radiation oncologist and urologist densities. Census files were linked to calculate county-level median income quartiles.

**Study population**

Patients (age ≥ 66 years) diagnosed with localized prostate cancer during the study period (2005-2014) were included in the study population. Cancer diagnosis was identified using International Classification of Diseases for Oncology (ICD-9). Localized prostate cancer was defined as cancer stage ≤T2a.

*Study inclusion/exclusion criteria:*

Patients with localized prostate cancer were included in the study if alive and continuously enrolled in Medicare Parts A and B throughout the baseline and follow-up periods. Patients with missing cancer stage, PSA values, Gleason Scores were excluded. Patients diagnosed with prostate cancer at autopsy were excluded (Appendix Figure 7.7).

**Dependent variable/target**

*Total healthcare expenditures* were calculated by summing inpatient (MEDPAR), outpatient (Outsaf, NCH), and “other” (durable medical equipment, and home health agency) claims. To compare non-cancer related healthcare expenditures, we subtracted treatment-related costs from total healthcare expenditures. All
expenditures were estimated 12-24 months after localized prostate cancer diagnosis. Consumer price index (CPI) was used to adjust all expenditures to 2016 constant dollars. (100)

**Key Independent variables/features**

*Low-value care*

Various professional societies identify medically unnecessary and/or potentially harmful procedures, tests, and treatments for Choosing Wisely in an international effort to reduce low-value care. (20,22) Low-value healthcare services were operationalized using a claims-based algorithm representing Choosing Wisely campaign recommendations. The Low-value care algorithm included procedure, diagnosis, hospitalization, and BETOS codes using previously published methods. (60–62)

**Other independent variables**

*Conservative management*

Conservative management was estimated within the first 12 months of the 24-month follow-up period using a validated claims-based algorithm. (53) Patients without treatment claims were designated as using conservative management.

*Multimorbidity*

The multimorbidity framework developed by the United States Department of Health and Human Services for guiding programs, practice, and policy guided the selection of chronic conditions as follows: arthritis, asthma, coronary artery disease, congestive heart failure, chronic kidney disease, chronic obstructive pulmonary disease, cardiac arrhythmias, acute myocardial infarction, dementia, diabetes, depression, hepatitis, hyperlipidemia, hypertension, human immunodeficiency virus, osteoporosis, substance abuse, schizophrenia, stroke, anemia, lower limb fracture. (55) We used clinically meaningful cut points (77) of multimorbidity (i.e., >2 chronic conditions) in descriptive analyses and included total counts of mental and physical health conditions in ML models.

*Prostate cancer comorbidity index*

Prostate cancer comorbidity index (PCCI) is a weighted comorbidity index validated in prostate cancer patient populations to predict 5 and 10 year life expectancy. (54) PCCI was calculated during the baseline period to estimate 5 and 10 year life expectancy. PCCI was categorized into 3 groups: 0-8 (> 10-year life expectancy); 9
to 13 (5-10-year life expectancy); and > 13 (< 5-year life expectancy). In GLM models, PCCI total representing >10 years of life expectancy was used as the reference group. In ML models, we included PCCI as a continuous variable.

Managing pre-existing conditions and an incident prostate cancer diagnosis requires the use of limited patient, clinician, and practice ecosystem resources; patients must balance the management of several chronic conditions and participate in shared-decision making to select prostate cancer management within a practice environment with variation in supply and demand of healthcare resources (i.e., supply of clinician specialists). Therefore, a competing demands framework and previous literature reviews were used to guide the selection of other independent variables known to affect low-value care and prostate cancer treatment selection within patient, clinician, and practice ecosystem domains.(20,27,32,101)

Other independent variables included: diagnosis year group (2003-2009 and 2010-2013), low-risk prostate cancer (operationalized as Gleason Score ≤ 6 and PSA test ≤ 10 ng/mL), education (Less than high-school graduation and college graduation per 10,000 persons over age 65), income, urologist, and radiation oncologist quartiles (per 10,000 persons over age 65).

We used a modified version of the Bice-Boxerman continuity of care index to calculate care fragmentation.(56) The continuity of care index represents the concentration of visits per patient among health care providers based on visit number, proportion of encounters to each provider, and total number of visits. We used physician specialty codes representing primary care, oncology, and various specialist visit encounters; where n is the total number of visits, nk is the total number of visits associated with a physician specialty, and k the total number of physician specialty codes.(59) For example, multiple encounters with a single provider would result in a score of zero, however, multiple encounters among several health care providers would result in a score approaching 1. We included many specialties as patients with multimorbidity and/or limited life expectancy may require services from multiple specialists. Care fragmentation was measured during the 12 month baseline period.

All independent variables were calculated in the 12 months before cancer diagnosis.

**Statistical Analysis**
Chi-square tests and t-tests were used to identify significant group differences in low-value care and CM use by categorical variables. Generalized linear models (GLM) with log-link and gamma distribution were used to estimate total healthcare cost. GLM transforms the mean cost directly, reversing the transformation is possible without bias estimation of mean cost. Differences between cost associated with each independent variable of interest and the exponentiated intercept term were reported as the incremental average cost (i.e., mean) associated with independent variables of interest. GLMs were fit using separate unadjusted and adjusted regressions to identify independent and interactive associations of key independent variables and total healthcare expenditures. All statistical tests were 2-sided with a 5% Type I error rate and were completed in SAS (SAS Institute Inc., Cary, NC, USA. Version 9.4, using Windows 10).

**Machine learning**

We used the XGBoost regression, a decision-tree ensemble machine learning algorithm, to determine the association of low-value care to healthcare expenditures.

Model training was conducted on 70% of the original data with 10-fold cross validation based on the original data set. Hyperparameter adjustments were made during training interactions to address overfitting. Healthcare expenditures were log transformed to approximate a normal distribution. Prediction performance was assessed on unseen data by reserving a 30% stratified random sample of the original data (i.e., the test set). Model tuning and test prediction performance were assessed using Residual Mean Squared Error and R2. Hyperparameter tuning parameters codes are available in Supplemental Materials 1.

SHAP is an *additive feature attribution method* that provides consistent, locally accurate, individualized feature attributions based on conditional expectation. SHAP is an improvement over other feature importance methods that utilize model performance metrics (i.e., gain/accuracy) subject to inconsistency bias. SHAP values can be sorted to illustrate feature importance and display cumulative effects of interactions. We used SHAP feature importance to describe the top 10 predictive features of non-cancer related total healthcare cost 12-24 months after incident prostate cancer diagnosis. SHAP summary and partial density plots for key variables were used to illustrate non-linear relationships between important features and non-cancer related total healthcare cost.
XGBoost hyperparameters can be adjusted to fine-tune regularization to control model complexity to optimize predictive performance. (43)

Machine learning analyses were conducted using Python 3.7 with open-source scikit-learn 0.21.3 wrapper interface for XGBoost regressor.

TreeSHAP was used to generate SHAP and SHAP feature plots using Python 3.7 and the SHAP package (0.29.2).

4.4 Results

Study cohort description

The study cohort was predominantly non-Hispanic white (81.3%) (Table 4.1). The median age at diagnosis was 73.5 years and did not differ by year of diagnosis (2003-2009: M=73.8, SD= 5.4; 2010-2013: M=73.2, SD=5.2). Overall 18.1% used CM, defined as no curative treatment within 12 months of incident localized prostate cancer diagnosis.

Among older men with incident localized prostate cancer, 25.2% used LVC in the 12 months prior to cancer diagnosis. Screening for asymptomatic carotid artery stenosis, Magnetic Resonance Imaging (MRI) for low-back pain, and traction for low-back pain were the most frequently reported low-value care procedures. Low-value care use counts ranged from 1 to 6, with 13422, 4105, and 5649 patients using 1, 2, or 3 or more LVC services, respectively.

Cohort characteristics and Low-value care

Use of low-value care was significantly more common in older men aged 75 years or older, with multimorbidity, with less than 10 years of life expectancy, and who used preventative care (Table 4.1). Use of any low-value care was lower among patients using CM (15.2%), with MRI for low-back pain, traction for low-back pain, and asymptomatic carotid artery stenosis more frequent among patients using treatment (Supplemental Table 4.2).

Predictive feature importance of low value care

The predictive feature importance for low-value on healthcare expenditures ranked 7th (Figure 4.1). Absence of low-value care was inversely associated with healthcare expenditures, with a complex linear increase for every additional low-value care service used (Figure 4.2).
Statistical Associations of Low-value care to healthcare expenditures

Average outpatient, inpatient, “other”, and non-treatment related total healthcare expenditures were all significantly higher in low-value care than in no low-value care use groups (Table 4.2). In adjusted models using GLM with log-link function, every 1-unit increase in low-value care use was associated with an $858.82 increase in healthcare expenditures (Table 4.4).

Other Leading predictors of healthcare expenditures

Multimorbidity (physical condition count), care fragmentation, and conservative management use were the top 3 predictors of healthcare expenditures with the highest absolute mean SHAP values (Figure 4.1). Partial density plots display linear and complex relationships with healthcare expenditures (Figure 4.2). Lower (0 to 1) and higher (>99) values of care fragmentation were inversely associated with healthcare expenditure prediction (Figure 4.2). Physical health condition counts displayed more variation lower (0-4) versus higher (>7) chronic condition counts (Figure 4.2). In contrast, CM use was inversely associated with healthcare expenditure prediction (Figure 4.1).

Statistical Associations among other important features

Patients with a life expectancy of greater than 10 years had lower average healthcare expenditures (M=$6941.96, SD=16028.12) versus 5-10 year (M=$12638.58±24002.64) and less than 5 year (M=$21284.75±34633.97) life expectancy groups (M=$8893.27 ±19802.73) (F=1618.68, p<0.001). Healthcare expenditures were also significantly lower among patients using CM (M=$7332.74 ± 18787.79) versus patients without CM use (M= $9237.09 ± 20003.22) (p<0.001).

CM use was associated with a $932.05 decrease in non-treatment related total healthcare expenditures versus treatment for low risk prostate cancer (Table 4.4). Life expectancies less than 10 years were associated with significantly higher, and CM use, significantly lower, non-treatment related healthcare expenditures in adjusted GLM models (Table 4.3).

4.5 Discussion

In this study, we used claims-based algorithms supported by expert consensus to identify low-value care use prior to incident prostate cancer diagnosis. We observed that 1 in 4 older men with incident prostate cancer used low-value care prior to diagnosis and that low value care use was associated with higher healthcare
expenditures, defined as total non-cancer related healthcare expenditures 12-24 months after incident prostate cancer diagnosis.

Low-value care use prior to incident prostate cancer diagnosis was also associated with substantial increases in healthcare expenditures. On average, older patients using low-value care before an incident localized prostate cancer diagnosis were more likely to use treatment (versus CM) and to have higher healthcare expenditures. These associations remained robust after adjustment for prostate cancer-specific life expectancy estimates, low-risk prostate cancer, and practice ecosystem factors such as physician supply and care fragmentation, suggesting that low-value care was a significant driver of cost within 2 years of diagnosis.

Efforts to promote a high-value, patient-centered healthcare system has led to increased attention on excessive spending on low-value care, under-provision of high-value care, and policy approaches to controlling the rising cost of cancer care. Recently, the Centers for Medicare & Medicaid Services (CMS) developed an Oncology Care First Model that establishes performance based payment components based on prospective total cost of care targets among patients completing curative treatment. As portions of cancer care reimbursement are transitioned from fee-for-service (FFS) to capitation payments, identifying the most important features of cost prediction could help providers establish more effective policies and interventions for high cost patients. Up-stream, low-value care, such as PSA screening for men over age 70, are known to contribute to increased short- and long-term costs associated with subsequent biopsies, treatment, and/or additional PSA tests. We found that other non-cancer related low-value care procedures were predictive of downstream expenditures after incident prostate cancer diagnosis. Our results have important implications for predicting patient cost after incident care diagnosis, given that low-value care use prior to incident prostate cancer diagnosis is predictive of higher cost.

CM use, defined as no treatment within one year post diagnosis (102), was one of the highest ranking features in predicting healthcare expenditures using machine learning models. Previous studies report that CM use, including both Active Surveillance and Watchful Waiting, could reduce annual healthcare costs associated with avoidable localized prostate cancer treatment by an estimated 1.2 billion dollars.(64) Other studies have reported 3 year median costs of $1914 and $10588 associated with CM versus low-value treatment,
respectively. We found that CM use was associated with a $1905 reduction in 12-24 month healthcare expenditures independent of treatment-related costs.

High level evidence supports CM for localized prostate cancer among men with low-risk disease and/or limited life expectancy to reduce overtreatment morbidities and excessive healthcare expenditures. Regardless of initial treatment modality, incident prostate cancer patients encounter significant ephemeral anxiety that predisposes patients to treatment without prostate cancer specific mortality benefit. Decisional conflict, adverse mental health effects, and survival expectations associated with diagnosis are independent factors associated with treatment selection regardless of risk grouping or comorbidity burden. Other studies report cascades of downstream care associated with avoidable treatment-related morbidities, competing demand for chronic condition management and preventative care, and compounding costs. We speculate that an incident localized prostate cancer diagnosis among patients using low-value care services may affect other forms of overtreatment or use of unnecessary care leading to higher cancer survivorship costs.

Older adults with multimorbidity, living in an urban locality, and/or specialist dense regions are at risk for low-value care and down-stream cascades of high cumulative costs. Patients with incident prostate cancer and multimorbidity are likely to encounter additional care fragmentation, which may include consultation from medical, urological, radiation, and/or surgical oncologists. Care fragmentation during prostate cancer survivorship is associated with redundant use of healthcare services, PSA testing, and proportional increases in cost. As in previous studies, we found care fragmentation and multimorbidity prior to incident prostate cancer diagnosis to highly ranked predictors of long term healthcare expenditures.

Current Oncology Care Model (OCM) do not address care complexities associated with multimorbidity, care fragmentation, or non-treatment approaches to localized prostate cancer survivorship. Although multidisciplinary approaches have been proposed as a management strategy for prostate cancer survivors, it remains unclear how primary care and oncology professionals can effectively coordinate care in a fragmented FFS environment. More research is needed to identify predictors of high cost consumption of care in order to inform population health management initiatives and allow policymakers to develop tailored interventions to proactively address high-cost patients.
4.6 Strengths and Limitations

Despite calls to identify low-value care use in vulnerable populations, no studies have investigated how use of low-value care in men with incident localized prostate cancer is associated with longer term survivorship costs. We used previously validated algorithms to identify conservative management use, prostate cancer specific mortality index, care fragmentation, and multimorbidity estimates to understand the predictive importance of low-value care use on healthcare cost during survivorship. We also estimated non-treatment related total expenditures 12-24 months after diagnosis to identify long-term costs lacking characterization in the current literature. We applied statistical, machine learning, and novel machine learning interpretative approaches to estimate and predict the impact of low-value care use on healthcare expenditures.

Our results must be interpreted with important limitations. First, this observational study is a claims-based, retrospective analysis, therefore, our results are subject to unobservable variable bias and selection bias of paid Medicare claims. To minimize the proportion of missed claims we included beneficiaries only continuously enrolled in Medicare (A&B, without HMO) throughout the study period. Second, our analysis may not be generalizable to commercial insurance beneficiaries as we only observed Medicare FFS beneficiaries. Our estimates of low-value care use are likely underestimated as methods for identification of low value care in claims data are limited. Social determinants of health (i.e., income, education) were not available at the individual level therefore county level measures were used. We only included measures detectable within a one year period to match the length of the baseline period before diagnosis, likely leading to further underestimation of costs. Although NCCN guidelines recommend CM for patients with low-risk cancer and a life expectancy of less than 10 years, both curative therapies and Active Surveillance are evidenced-based choices for patients with a life expectancy of greater than 10 years.(6) Lastly, we did not account for follow-up costs such as PSA and/or biopsy procedures that are commonly part of routine medical practice for patients using Active Surveillance and/or treatment related monitoring.

4.7 Conclusion

In this study, incremental use of low-value healthcare showed a significant positive association with long-term, non-treatment related costs after adjustments for multiple possible confounders. Use of evidence-based CM after diagnosis was associated with lower non-treatment related total healthcare costs. Using
machine learning, we estimate both low-value care and CM use to be high-ranking features of cost prediction 12-24 months after incident prostate cancer diagnosis. Targeting patients with low-value care use prior to cancer diagnosis using existing claims data could help to reduce low-value care related morbidities by proactively identifying high-cost patients.
| Table 4.1 Patient Characteristics by Low-value Care Use among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671) |
|---------------------------------|--------|
|                                | LVC    | No LVC |
| N                               | %      | N      |
| **ALL**                         | 19,071 | 56,600 |
| Age in Years                    | 400.79 | <0.001 |
| 66-74                           | 12075  | 37474  | 75.6 |
| 75+                             | 8135   | 17987  | 68.9 |
| Race                            | 39.00  | <0.001 |
| White                           | 16701  | 44853  | 72.9 |
| Black                           | 2079   | 6533   | 75.9 |
| Hispanic                        | 352    | 927    | 72.5 |
| Other                           | 1025   | 3014   | 74.6 |
| Marital Status                  | 52.97  | <0.001 |
| Married                         | 1271   | 3881   | 75.3 |
| Unmarried                       | 13908  | 38585  | 73.5 |
| Separated/Divorced/Widowed      | 2495   | 7002   | 73.7 |
| Unknown                         | 2536   | 5993   | 70.3 |
| Income quintiles                | 38.85  | <0.001 |
| First                           | 3746   | 10938  | 74.5 |
| Second                          | 3841   | 10912  | 74.0 |
| Third                           | 3909   | 10860  | 73.5 |
| Four                            | 3992   | 10726  | 72.9 |
| Fifth                           | 4283   | 10785  | 71.6 |
| Education quintiles             | 51.10  | <0.001 |
| First                           | 4184   | 10545  | 71.6 |
| Second                          | 4062   | 10748  | 72.6 |
| Third                           | 3962   | 10890  | 73.3 |
| Four                            | 3872   | 10941  | 73.9 |
| Fifth                           | 3712   | 11146  | 75.0 |
| Urologist density quintiles     | 49.73  | <0.001 |
| First                           | 3809   | 11262  | 74.7 |
| Second                          | 4265   | 11114  | 72.3 |
| Third                           | 3889   | 10562  | 73.1 |
| Four                            | 4349   | 11206  | 72.0 |
| Fifth                           | 3880   | 11231  | 74.3 |

Note: Based on 75671 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare part A & Part B, diagnosed with incident localized prostate cancer between 2005 and 2014. SEER= Surveillance, Epidemiology and End Results cancer Registry, CM= Conservative management.
Table 4.1: Continued

<table>
<thead>
<tr>
<th>Radiation oncologist quintiles</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>3773</td>
<td>24.5</td>
<td>11641</td>
</tr>
<tr>
<td>Second</td>
<td>4239</td>
<td>29.0</td>
<td>10365</td>
</tr>
<tr>
<td>Third</td>
<td>3739</td>
<td>25.2</td>
<td>11071</td>
</tr>
<tr>
<td>Four</td>
<td>4667</td>
<td>29.9</td>
<td>10960</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multimorbidity</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-2 Chronic conditions</td>
<td>4052</td>
<td>12.1</td>
<td>29314</td>
</tr>
<tr>
<td>&gt;2 Chronic conditions</td>
<td>16158</td>
<td>38.2</td>
<td>26147</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prostate Cancer Comorbidity Index</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10 or more years life expectancy</td>
<td>11676</td>
<td>20.2</td>
<td>46207</td>
</tr>
<tr>
<td>5-10 years life expectancy</td>
<td>5290</td>
<td>42.6</td>
<td>7140</td>
</tr>
<tr>
<td>5 or less years life expectancy</td>
<td>3244</td>
<td>60.5</td>
<td>2114</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conservative management</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservative management use</td>
<td>3070</td>
<td>22.5</td>
<td>10592</td>
</tr>
<tr>
<td>No conservative management use</td>
<td>17140</td>
<td>27.6</td>
<td>44869</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preventative A1c test</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A1c test</td>
<td>3604</td>
<td>35.9</td>
<td>6442</td>
</tr>
<tr>
<td>No A1c test</td>
<td>16606</td>
<td>25.3</td>
<td>49019</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preventative flu</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Influenza vaccination</td>
<td>9630</td>
<td>31.7</td>
<td>20756</td>
</tr>
<tr>
<td>No influenza vaccination</td>
<td>10580</td>
<td>23.4</td>
<td>34705</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preventative lipid screen</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lipid test</td>
<td>6270</td>
<td>30.2</td>
<td>14520</td>
</tr>
<tr>
<td>No lipid test</td>
<td>13940</td>
<td>25.4</td>
<td>40941</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metro</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Metro county</td>
<td>17337</td>
<td>27.2</td>
<td>46301</td>
</tr>
<tr>
<td>Non-metro county</td>
<td>2855</td>
<td>23.9</td>
<td>9074</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SEER Region</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td>4499</td>
<td>29.0</td>
<td>11020</td>
</tr>
<tr>
<td>South</td>
<td>5018</td>
<td>25.9</td>
<td>14345</td>
</tr>
<tr>
<td>North Central</td>
<td>2533</td>
<td>29.6</td>
<td>6030</td>
</tr>
<tr>
<td>West</td>
<td>8160</td>
<td>25.3</td>
<td>24066</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year of diagnosis</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2009</td>
<td>9651</td>
<td>29.1</td>
<td>23567</td>
</tr>
<tr>
<td>≥2009</td>
<td>10559</td>
<td>24.9</td>
<td>31894</td>
</tr>
</tbody>
</table>

Note: Based on 75,671 older (age ≥66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare part A & Part B, diagnosed with incident localized prostate cancer between 2005 and 2014. SEER= Surveillance, Epidemiology and End Results cancer Registry, CM= Conservative management.
**Table 4.2** Healthcare Expenditures by Low-value Care Use among Elderly Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer SEER-Medicare, 2005-2014 (n=75671)

<table>
<thead>
<tr>
<th></th>
<th>LVC</th>
<th>No LVC</th>
<th>CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=19071</td>
<td>n=56600</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean (SD)</strong></td>
<td><strong>Mean (SD)</strong></td>
<td><strong>CI</strong></td>
<td><strong>p-value</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Total healthcare expenditures</strong></td>
<td>12589.57 (24028.45)</td>
<td>7647.83 (17988.73)</td>
<td>8752.17, 9034.37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Outpatient expenditures</strong></td>
<td>6467.20 (8740.10)</td>
<td>4063.00 (6110.23)</td>
<td>4619.42, 4718.42</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Inpatient expenditures</strong></td>
<td>5276.75 (18032.29)</td>
<td>3114.84 (13664.18)</td>
<td>3553.42, 3765.97</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Other expenditures</strong></td>
<td>845.69 (2996.22)</td>
<td>469.98 (2120.18)</td>
<td>547.74, 803.17</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: Based on 75671 older (age >66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare part A & Part B, diagnosed with incident localized prostate cancer between 2005 and 2014. Healthcare expenditures for patients completing curative treatment (i.e., “No CM”) do not include treatment costs. CI= Confidence interval, SEER= Surveillance, Epidemiology and End Results cancer Registry, CM= Conservative management.
<table>
<thead>
<tr>
<th></th>
<th>CM</th>
<th>No CM</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>13662</td>
<td>62009</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>(Mean (SD))</td>
<td>Mean (SD)</td>
<td></td>
</tr>
<tr>
<td><strong>Total healthcare expenditures</strong></td>
<td>7332.74 (18787.79)</td>
<td>9237.09 (20003.22)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Outpatient expenditures</strong></td>
<td>3681.65 (7334.92)</td>
<td>4886.44 (6840.01)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Inpatient expenditures</strong></td>
<td>3111.29 (13781.13)</td>
<td>3780.53 (15151.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Other expenditures</strong></td>
<td>539.8 (2438.32)</td>
<td>570.16 (2363.56)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Based on 75671 older (age >66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare part A & Part B, diagnosed with incident localized prostate cancer between 2005 and 2014. Healthcare expenditures for patients completing curative treatment (i.e., “No CM”) do not include treatment costs. CI= Confidence interval, SEER= Surveillance, Epidemiology and End Results cancer Registry, CM= Conservative management.
Table 4.4 Parameter estimates of PCCI categories from unadjusted and adjusted generalized linear models on 12-24 months healthcare expenditures among elderly Medicare fee-for-service beneficiaries with incident localized prostate cancer diagnosis using SEER-Medicare data, 2005-2015 (n=75671).

<table>
<thead>
<tr>
<th>Parameter Estimate (SE)</th>
<th>Unadjusted GLM</th>
<th>Change #</th>
<th>Parameter Estimate (SE)</th>
<th>Change #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66-74</td>
<td>8.949 (0.010)</td>
<td>7700.19</td>
<td>8.416 (0.055)</td>
<td>$4518.79</td>
</tr>
<tr>
<td>75 and over</td>
<td>0.370 (0.017)</td>
<td>3447.64</td>
<td>0.247 (0.018)</td>
<td>$1266.07</td>
</tr>
<tr>
<td><strong>Prostate cancer comorbidity score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 or more years</td>
<td>8.845 (0.009)</td>
<td>6939.60</td>
<td>8.416 (0.055)</td>
<td>$4518.79</td>
</tr>
<tr>
<td>Less than 5 years</td>
<td>1.120 (0.031)</td>
<td>14329.27</td>
<td>0.895 (0.033)</td>
<td>$6540.21</td>
</tr>
<tr>
<td>Between 5 and 10 years</td>
<td>-0.599 (0.022)</td>
<td>5692.54</td>
<td>0.470 (0.023)</td>
<td>$2711.25</td>
</tr>
<tr>
<td><strong>Conservative management</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>9.131 (0.009)</td>
<td>7331.97</td>
<td>8.416 (0.055)</td>
<td>$4518.79</td>
</tr>
<tr>
<td>Yes</td>
<td>-0.231 (0.021)</td>
<td>-1905.28</td>
<td>-0.219 (0.022)</td>
<td>$-932.05</td>
</tr>
<tr>
<td><strong>Low-value care†</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.956 (0.009)</td>
<td>7754.28</td>
<td>8.416 (0.055)</td>
<td>$4518.79</td>
<td></td>
</tr>
<tr>
<td><strong>Total Low-value care</strong></td>
<td>0.312 (0.013)</td>
<td>2839.27</td>
<td>0.174 (0.012)</td>
<td>$858.82</td>
</tr>
<tr>
<td><strong>Low-risk prostate cancer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>9.141 (0.010)</td>
<td>9320.77</td>
<td>8.416 (0.055)</td>
<td>$4518.79</td>
</tr>
<tr>
<td>Yes</td>
<td>-0.140 (0.017)</td>
<td>-1217.68</td>
<td>-0.023 (0.018)</td>
<td>NS</td>
</tr>
<tr>
<td><strong>Care fragmentation</strong></td>
<td>.010 (0.000)</td>
<td>52.87</td>
<td>.006 (0.00)</td>
<td>$27.19</td>
</tr>
</tbody>
</table>

Note: Based on 75671 older (age > 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare part A & Part B, diagnosed with incident localized prostate cancer between 2005 and 2014. Healthcare expenditures for patients completing curative treatment (i.e., “No CM”) do not include treatment costs. Total healthcare expenditures include inpatient, outpatient, durable medical equipment, and home health agency costs with treatment-related cost subtracted.

† Sum of low-value care procedures in 12 months before prostate cancer diagnosis.
Change # was calculated by difference between the 1) exponentiation of the model intercept term and 2) the sum of the intercept and the variable parameter estimate.
Compare to omitted category
SE= Standard Error, SEER= Surveillance, Epidemiology and End Results cancer Registry, CM= Conservative management.
**Figure 4.1** SHAP Feature Importance and Summary Plot of Low-value Care and Conservative Management Use on Non-Treatment Related Total Healthcare Expenditures among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671)

(A) Mean SHAP values and (B) SHAP summary plot in descending order of log-dollars. Based on 75671 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident localized prostate cancer (Stage ≤T2a) between 2005 and 2014.
Figure 4.2 SHAP Partial Density Plots of Selected Features on Non-Treatment Related Total Healthcare Expenditures among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671)

SHAP Partial Dependence plots (PDP) of SHAP values (log-dollars) values by (A.) Multimorbidity (Physical conditions only), (B.) Care fragmentation, (C) Low value care total prior to incident prostate cancer diagnosis, and (D.) Conservative Management. Based on 75671 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident localized prostate cancer (Stage ≤T2a) between 2005 and 2014.

Physical condition total = Total count of multimorbidity physical conditions.
Care fragmentation = Bice-Boxerman continuity of care index to calculate care fragmentation during the 12 month baseline period (See Methods).
Low value care = Sum of low value care prior to incident prostate cancer diagnosis.
Conservative Management = Yes (1)/No (0).
SHAP = Shapley Additive ePlanations.
SEER= Surveillance, Epidemiology, and End Results cancer Registry.
Chapter 5

5 Summary and Conclusion

5.1 Summary of Findings and Discussion

The goal of this dissertation is to highlight healthcare practices that provide limited clinical benefit, potentially harmful effects, and a significant economic burden among older men with incident localized prostate cancer. This dissertation also highlights the role of patient experiences in receiving high value care for incident prostate cancer. Older men with localized prostate cancer represent a large, medically complex and growing population. Although prostate cancer is heterogeneous at the population level, high-level randomized evidence concluded a lack of significant survival benefit of treatment (surgery, radiation, chemotherapy, hormone therapy, cryotherapy) among men with intermediate grade disease and 15 year life expectancy. Conservative management (Active Surveillance and Watchful Waiting) within the context of cancer risk group, life expectancy estimates, and personal values has been promoted as high-value care as significant quality decrement and high costs of overtreatment can be avoided. Despite evidence-based guidelines recommending conservative management for a decade, many men with low-risk prostate cancer and limited life expectancy received treatment for prostate cancer. Higher rates of treatment use among fee-for-service (58%) versus integrated healthcare beneficiaries (20%) highlight disparities in evidenced-based healthcare delivery.

This dissertation focused on three related studies: 1) examine the leading predictors of low-value healthcare practice (i.e., prostate cancer treatment for low-risk prostate cancer) among older men; 2) assess the role of patient-reported experience with care on high-value prostate cancer management; and 3) Estimate the association of high-value care on non-cancer related healthcare expenditures. In all of these studies an emphasis is placed on multimorbidity because of high prevalence and negative clinical, humanistic, and economic effects throughout cancer survivorship. We specifically examined the associations of patient-reported experiences with evidence-based treatment selection and low-value care to reflect the standards of the Institute of Medicines (IOM) and The National Quality Task Force aim to promote patient-centered care and reductions in inappropriate, wasteful healthcare practices among adults managing chronic conditions.
This dissertation filled a knowledge gap through a comprehensive analysis of clinical and non-clinical factors that drive low-value cancer treatment among incident low-risk prostate cancer. A novel feature of this dissertation is the use of machine learning algorithms and the application of model agnostic interpretable machine learning techniques. This dissertation also made a unique contribution and clarified the role of life expectancy on the complex decision making process of choosing from several treatments (surgery, radiation, cryotherapy, and or chemotherapy and hormone therapy) or conservative management. We distinguished between comorbidity and life expectancy to estimate the independent contribution of life expectancy on treatment choice. This dissertation is a series of firsts for its measurement of care fragmentation during cancer survivorship, use of machine learning algorithms and unlocking the “black box” prediction with interpretable machine learning techniques, and assessing the role of patient-reported experiences on high-value care among older men with incident prostate cancer.

5.2 Implications and Suggestions for Future Research

A common theme that emerged from our studies is the negative effect of multimorbidity on outcomes (i.e., low-value cancer treatment, high-value conservative management, and non-cancer related healthcare expenditures). This finding is not surprising given the voluminous data and robust evidence on the negative health effects of multimorbidity on health outcomes in a variety of settings and populations. Adults with multimorbidity and incident cancer are at high-risk for poor quality of both cancer and non-cancer care (37,107) as current non-cancer clinical guidelines are developed within a single-disease framework.(108) Men with multimorbidity and incident cancer may not opt for conservative management because of a preference for immediate cure (ie, “take care of it”)(87) and may have a preference to avoid one more condition that requires long-term management. In addition, men are likely to experience fear of non-treatment regret,(88) emotional distress,(89) and anxiety.(90) The implementation of OCM which provides a bundled payment for all care during an episodic treatment will inherit the complex management requirements of men with multimorbidity and incident prostate cancer. Value-based payment mechanisms are needed to support the conservative management of incident localized prostate cancer with multimorbidity, especially among men with limited life expectancy, who benefit more from non-cancer chronic condition management versus curative treatment and
other potential cascades of low-value care or increased care fragmentation without the benefit of reduced prostate cancer specific mortality.

Our results suggest focusing on reducing care fragmentation, such as broader system reforms to increase care continuity between primary care physicians and oncologists, could help reduce treatment selection among patients with low-risk prostate cancer. Considering our results, we speculate that care fragmentation is likely to contribute to differing treatment rates for low-risk prostate cancer between Medicare FFS and Veterans Administration beneficiaries (approximately 37%). (46)

For a subset of individuals older men with localized prostate cancer we observed that patient experiences with timely care was important in treatment choice. Timely access to care for men with localized prostate cancer is not limited to initial diagnosis of prostate cancer, but the opportunity and ease by which a patient is able to access and utilize needed services along the continuum of care throughout survivorship. (93) Choices for elderly localized prostate cancer patients involve selecting curative and non-curative treatments with trade-offs in efficacy, potential adverse quality of life effects, and competing risk mortality. Timeliness of care domains used in this dissertation, such as perceived barriers to appointment scheduling, are fundamental to shared decision-making among multiple health-care providers that significantly influence treatment choice for men with localized prostate cancer. (94–96) We speculate that adults with higher timely care ratings may choose CM if they have a favorable perception of health-care system capacity to provide services once a need is detected. Our results suggest that addressing specific modifiable barriers to timely care along the cancer continuum for older adults with localized prostate cancer and limited life expectancy could reduce the adverse effects of overtreatment on health outcomes and costs.

A noteworthy finding is the superior performance of machine learning algorithms over traditional statistical models, specifically logistic regression. This dissertation confirmed flexibility, predictive accuracy, and ability of machine algorithms to handle multicollinear variables. With machine learning, model misspecification can be avoided through the use of cross-validation model tuning on “training” data before applying the algorithm to a “test” or new data set. Once validated, interpretable machine learning approaches, such as SHAP, can be applied to reveal the complexity and non-linearity of associations using feature importance and SHAP summary and partial density plots. In our dissertation, machine learning analyses
revealed significant variation among individuals (i.e. patient-level) of estimated outcomes at specific feature values. Our results highlight the need to target individuals with non-linear, complex feature combinations pushing them toward low-value outcomes that are otherwise unobservable with “on average” traditional statistical approaches.

The application of machine learning algorithms to predict low- and high-value care are in alignment with the National Quality Forum created a Technology Evaluation Framework.(16) These organizations promote the use of advanced technologies, such as machine learning algorithms, to identify solutions to address high-value healthcare practices. The framework seeks to strengthen health literacy to complement shared decision making processes to facilitate evidence-based medicine while reducing propensity for low-value services. Our study demonstrates that machine learning and interpretive approaches can be leveraged to proactively identify patients who may benefit from interventions, for both patients and physicians, to facilitate shared decision making regarding prostate cancer care and reduce the use of low-value care.

Our results from the analysis on the association of low- and high-value care to non-cancer related healthcare expenditures suggest the development and implementation of targeted payment reforms could reduce the economic burden to the payers, specifically Medicare. Current Oncology Care Model (OCM) are voluntary five-year bundled payment programs developed by the Center for Medicare & Medicaid Innovation (CMMI) to facilitate high-value, lower-cost healthcare through improved care coordination and episodic or bundled reimbursement. OCM incentivizes providers to lower the total cost of care for patients throughout the treatment episodes.(109) Although multidisciplinary approaches have been proposed as a management strategy for prostate cancer survivors, it remains unclear how primary care and oncology specialist delivery coordinated care in a fragmented FFS environment. Our research highlights how the application of machine learning can be used to identify patients at risk of low-value care. Healthcare entities can leverage large patient data sets, claims registries, and electronic medical records to continuously re-train machine learning models to improve predictive accuracy and provide personalized medicine for many diseases. However, the implementation of such technologies will require a robust technical investment from healthcare entities to develop specific algorithms for the patient population they serve. Policy makers should consider mechanisms.
to support healthcare organizations in the development of predictive “precision” population health technologies.

5.3 Strengths and Limitations

Our dissertation was a series of many first. We used both multimorbidity and prostate cancer-specific of life expectancy measures to understand the independent contributions of comorbidity and prostate cancer specific mortality on treatment selection among men with localized prostate cancer. We used robust machine learning and statistical methods to explore novel predictive and statistical associations of low-value care, patient-reported experience measures, and care fragmentation on overtreatment and cost outcomes among older adults with multimorbidity. Our research supports nationally recognized strategic objectives for improving patient-centered care while identifying predictors to reduce inappropriate low-value care. We used cancer registry, claims, MCAHPS surveys, and other socio-demographic data linkages to incorporate measures of multimorbidity, care fragmentation, patient-reported experience measures, social determinants of health, and low-value care in machine learning and statistical models. Although educational attainment may not be modifiable among older adults, initiatives such as “health in all policies” by the World Health Organization and the Centers for Disease and Prevention Control suggest the integration and consideration of community health policy(73) as social, economic, and physical environments have a significant impact on the health of an individual. With these considerations, we developed machine learning and statistical approaches using a competing demands framework in all of our study aims.

Due to the retrospective nature of this work our findings are not causal. Generalizability to younger patients, health Maintenance Organization, and integrated healthcare networks may not be appropriate as our studies included only fee-for-service beneficiaries. Measurement of low-value care in claims data requires the estimation of clinical scenarios using inclusion and exclusion criteria that are diagnosis and/or procedure specific, and often include a temporal component. We could only select procedures meeting criteria for low-value care that could be operationalized accurately in claims data, therefore, we are likely to underestimate the burden of low-value care.
5.5 Conclusion

Our studies reveal that low-value cancer treatment is highly-prevalent as only 2 in 5 men received conservative management. In predictive models of low-value care (i.e., overtreatment) and high-value care (i.e., conservative management) we did not find a life expectancy of 10 or more years to be predictive. These results indicate that treatment choice among older adults, within the context of life expectancy, do not reflect evidence-based guidelines for prostate cancer treatment. Our findings suggest a need for broad implementation of interventions to improve patient and physician education during the shared-decision process.
6 References

15. Institute of Medicine (US) Committee on Quality of Health Care in America. Executi...Crossing the Quality Chasm - NCBI Bookshelf. 2001;


7 Appendices

Supplemental Figure 7.0 Aim 1 Inclusion & Exclusion Criteria

- SEER-Medicare Fee-for-service (FFS) Prostate Cancer Diagnoses 2005-2014 502156
- No HMO, Continuous Enrollment in Medicare FFS A & B 13229
- Not Diagnosed at Autopsy 132207
- Age ≥ 66 at Diagnosis 123033
- AJCC stage ≤T2a 98287
- Gleason Score & PSA Value Available 75671
- Low-Risk Prostate Cancer (Gleason Score=6, PSA≤10 ng/mL) 13870
Supplemental Figure 7.1 Feature Importance of Treatment Prediction among Fee-for-Service Medicare Beneficiaries with Incident Low-risk Prostate Cancer using SEER Cancer Registry, 2009-2014 (n=13870)

Mean SHAP values in descending order of log-odds. Based on 13870 older (age ≥ 66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare Part A & Part B, diagnosed with incident low-risk prostate cancer (Gleason 6, PSA <10, Stage ≤T2a) between 2009 and 2014.

Age group included 66-74 and 75 and over categories.
Care fragmentation = Bice-Boxerman continuity of care index to calculate care fragmentation during the 12 month baseline period (See Methods).
SEER= Surveillance, Epidemiology, and End Results cancer Registry.
Appendix 7.2 Unadjusted and Adjusted Odds Ratios and 95% Confidence Intervals (CI) of PCCI, Mental Health, Timeliness of Care, Low-risk Prostate Cancer, and Education on Conservative Management Use in among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (N = 496).

<table>
<thead>
<tr>
<th>Prostate Cancer Comorbidity Index</th>
<th>UOR [95% CI]</th>
<th>p-value</th>
<th>AOR [95% CI]</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 to 10 years life expectancy</td>
<td>0.58 [0.38 - 0.90]</td>
<td>0.014</td>
<td>0.48 [0.29 - 0.78]</td>
<td>0.003</td>
</tr>
<tr>
<td>&lt; 5 years life expectancy</td>
<td>0.58 [0.33 - 1.03]</td>
<td>0.062</td>
<td>0.42 [0.21 - 0.83]</td>
<td>0.012</td>
</tr>
<tr>
<td>&gt; 10 years life expectancy (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Getting Care Quickly</td>
<td>1.15 [1.05 - 1.27]</td>
<td>0.003</td>
<td>1.20 [1.09 - 1.34]</td>
<td>0.001</td>
</tr>
<tr>
<td>Low-risk prostate cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.41 [0.95 - 2.08]</td>
<td>0.088</td>
<td>1.65 [1.07 - 2.58]</td>
<td>0.024</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental Health</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>2.97 [1.50 - 5.90]</td>
<td>0.002</td>
<td>5.54 [2.33 - 13.2]</td>
<td>0.001</td>
</tr>
<tr>
<td>Excellent/Very Good (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College or more</td>
<td>2.49 [1.30 - 4.78]</td>
<td>0.006</td>
<td>3.28 [1.52 - 7.10]</td>
<td>0.002</td>
</tr>
<tr>
<td>High-school graduate</td>
<td>2.41 [1.18 - 4.92]</td>
<td>0.015</td>
<td>3.57 [1.60 - 7.96]</td>
<td>0.002</td>
</tr>
<tr>
<td>No high-school grad. (Ref.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Appendix 7.3 Unadjusted and Adjusted Odds Ratios and 95% Confidence Intervals of PCCI, Multimorbidity, Mental Health, Timeliness of Care, Low-risk Prostate Cancer, and Education on Conservative Management Use up to 24 Months after Incident Prostate Cancer among Fee-for-Service Medicare Beneficiaries using Linked SEER Cancer Registry with MCAHPS, 2002-2013 (N = 496).

<table>
<thead>
<tr>
<th></th>
<th>UOR [95% CI]</th>
<th>AOR [95% CI]</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multimorbidity Model</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multimorbidity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.52 [0.35 - 0.77]</td>
<td>0.42 [0.27 - 0.65]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Getting Care Quickly</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.13 [1.02 - 1.24]</td>
<td>1.21 [1.09 - 1.35]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low-risk prostate cancer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.46 [0.97 - 2.18]</td>
<td>1.76 [1.14 - 2.72]</td>
<td>0.01</td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mental Health</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>3.44 [1.73 - 6.81]</td>
<td>4.32 [1.86 - 10.1]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Excellent/Very Good (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College or more</td>
<td>2.76 [1.35 - 5.65]</td>
<td>3.21 [1.50 - 6.89]</td>
<td>0.003</td>
</tr>
<tr>
<td>High-school graduate</td>
<td>3.06 [1.42 - 6.63]</td>
<td>3.53 [1.59 - 7.83]</td>
<td>0.002</td>
</tr>
<tr>
<td>No high-school graduation (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PCCI Model</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostate Cancer Comorbidity Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 to 10 years life expectancy</td>
<td>0.52 [0.33 - 0.82]</td>
<td>0.39 [0.30 - 0.70]</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt; 5 years life expectancy</td>
<td>0.56 [0.31 - 1.01]</td>
<td>0.34 [0.16 - 0.73]</td>
<td>0.005</td>
</tr>
<tr>
<td>&gt;10 years life expectancy (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Getting Care Quickly</strong></td>
<td></td>
<td>1.20 [1.08 - 1.34]</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>Low-risk prostate cancer</strong></td>
<td></td>
<td>1.76 [1.11 - 2.79]</td>
<td>0.017</td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mental Health</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair/Poor</td>
<td>7.59 [3.07 - 18.8]</td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Excellent/Very Good (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-school graduate</td>
<td>0.26 [0.11 - 0.60]</td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td>No high-school grad. (Ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Based on 496 older (age ≥66 years) Fee-for-Service Medicare beneficiaries, with continuous enrollment in Medicare part A & Part B, diagnosed with incident localized prostate cancer between 2003 and 2013. Adjusted for age, race, marital status, income, education, health status, urologist density, radiation oncologist density, SEER region, geography, diagnostic year, and low-risk prostate cancer.

UOR= Unadjusted Odds Ratio, AOR= Adjusted Odds Ratio, N.S.= Not significant, CI= Confidence interval, Ref= Reference group, SEER= Surveillance, Epidemiology and End Results cancer Registry, MCAHPS= Medicare Consumer Assessment of Healthcare Providers and System surveys, CM= Conservative management, PCCI= Prostate Cancer Comorbidity Index. Statistically significant results displayed.
### Appendix 7.4 Prostate cancer treatment codes used within 12 months of prostate cancer diagnosis to Identify Conservative Management Use

<table>
<thead>
<tr>
<th>Treatment Type</th>
<th>Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Radical prostatectomy</strong></td>
<td>55840, 55842, 55845, 55866, 55810, 55812, 55815 and 60.62, 17.42, 60.5, 60.4, and 60.3 (40.3, 40.53, and 40.59 for lymph node dissection)</td>
</tr>
<tr>
<td><strong>Brachytherapy</strong></td>
<td>76873, 55859, 76965, 55860, 55875, 55876, 76873, 76965, 77326, 77327, 77328, 7761, 77762, 77763, 77799, 77776, 77777, 77778, and 60.99, 92.27, 92.28, and 92.29, 77781, 77782, 77783, 77784, 77785, 77786, 77787, 77789, 77790, Q3001, A9527, C1715, C1716, C1717, C1719, C1728, C2616, C2634, C2635, C2636, C2637, C2638, C2639, C2640, C2641, C2642, C2643, C2698, C2699, C9725</td>
</tr>
<tr>
<td><strong>External Beam Radiotherapy</strong></td>
<td>77305, 77310, 77315, 77321, 77371, 77372, 77373, 77402, 77403, 77404, 77406, 77407, 77408, 77409, 77411, 77412, 77413, 77414, 77416, 77422, 77423, 92.24, 92.26, 77301, 77418, 0073T, 77380, 77381, 77520, 77522, 77523, and 77525, 77301, 77338, 77761, 77762, 77763, 77789, 77427</td>
</tr>
<tr>
<td><strong>Cryotherapy</strong></td>
<td>50250, 50593, 55873, 55873, 60.62, C2618, G0160, G0161</td>
</tr>
<tr>
<td><strong>Androgen Deprivation Therapy</strong></td>
<td>J1050, J1051 (Progesterone), J1950 J9217 J9218 J9219, C9430 (Lupron), J9165 (DES), J9202 (Zoladex), J3315 (Trelstar), J9225 CPT 11981(Vantas), all injections ICD-99.24, S0175 (Flutamide), J8999 (Bicalutamide/Flutamide/Nilutamide) S0165, J0128, C9216, (Abarelix), J9155 (Degarelix), S9560 (any hormone/adt), G0356 (any ADT)</td>
</tr>
<tr>
<td><strong>Conservative management</strong></td>
<td>No treatment within one year of prostate cancer diagnosis</td>
</tr>
</tbody>
</table>
### Appendix 7.5 Patient Characteristics by Conservative Management among Fee-for-Service Medicare Beneficiaries with Incident Localized Prostate Cancer using Linked SEER Cancer Registry, 2005-2014 (n=75671)

<table>
<thead>
<tr>
<th></th>
<th>CM</th>
<th>No CM</th>
<th><strong>X²</strong></th>
<th><strong>p-value</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ALL</strong></td>
<td>13662 18.1%</td>
<td>62009 82.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age in Years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66-74</td>
<td>8424 17.0%</td>
<td>41125 83.0%</td>
<td>107.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>75+</td>
<td>5238 20.1%</td>
<td>20884 79.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td>145.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>White</td>
<td>10741 17.4%</td>
<td>50813 82.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1941 22.5%</td>
<td>6671 77.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>188 14.7%</td>
<td>1091 85.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>750 18.6%</td>
<td>3289 81.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Marital Status</strong></td>
<td></td>
<td></td>
<td>813.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Married</td>
<td>8190 15.6%</td>
<td>44303 84.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmarried</td>
<td>997 19.4%</td>
<td>4155 80.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separated/Divorced/Widowed</td>
<td>2204 23.2%</td>
<td>7293 76.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>2271 26.6%</td>
<td>6258 73.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Income quintiles</strong></td>
<td></td>
<td></td>
<td>35.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>First</td>
<td>2772 18.9%</td>
<td>11912 81.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>2620 17.8%</td>
<td>12133 82.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>2696 18.3%</td>
<td>12073 81.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four</td>
<td>2558 17.4%</td>
<td>12160 82.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fifth</td>
<td>2642 17.5%</td>
<td>12426 82.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Education quintiles</strong></td>
<td></td>
<td></td>
<td>102.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>First</td>
<td>2604 17.7%</td>
<td>12125 82.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>2461 16.6%</td>
<td>12349 83.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>2496 16.8%</td>
<td>12356 83.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four</td>
<td>2758 18.6%</td>
<td>12055 81.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fifth</td>
<td>2984 20.1%</td>
<td>11874 79.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Urologist density quintiles</strong></td>
<td></td>
<td></td>
<td>33.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>First</td>
<td>2509 16.6%</td>
<td>12562 83.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>2820 18.3%</td>
<td>12559 81.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third</td>
<td>2689 18.6%</td>
<td>11762 81.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four</td>
<td>2817 18.1%</td>
<td>12738 81.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fifth</td>
<td>2798 18.5%</td>
<td>12313 81.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continued Appendix 7.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation oncologist quintiles</td>
<td>58.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>First</strong></td>
<td>2631</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12783</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>82.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Second</strong></td>
<td>2421</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12183</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>83.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Third</strong></td>
<td>2756</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12054</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>81.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Four</strong></td>
<td>2960</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>81.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fifth</strong></td>
<td>2865</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12247</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>81.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low-risk prostate cancer</strong></td>
<td>2600</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS &gt;6 and/or PSA ≥10ng/mL</td>
<td>6245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>42513</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS 6 and PSA &lt;10ng/mL</td>
<td>7417</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19496</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Multimorbidity</strong></td>
<td>752.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2 Chronic conditions</td>
<td>7465</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>22.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25901</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;2 Chronic conditions</td>
<td>6197</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prostate Cancer Comorbidity Index</strong></td>
<td>67.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 or more years life expectancy</td>
<td>10780</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>47103</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-10 years life expectancy</td>
<td>1927</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10503</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 or less years life expectancy</td>
<td>955</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4403</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low-value care</strong></td>
<td>206.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-value care</td>
<td>3070</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No low-value care</td>
<td>10592</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44869</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preventative A1c test</strong></td>
<td>38.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1c test</td>
<td>1590</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8456</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No a1c test</td>
<td>12072</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>53553</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preventative flu</strong></td>
<td>422.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Influenza vaccination</td>
<td>4420</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25966</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No influenza vaccination</td>
<td>9242</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36043</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preventative lipid screen</strong></td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lipid test</td>
<td>3161</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17629</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No lipid test</td>
<td>10501</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44380</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Metro</strong></td>
<td>14.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metro county</td>
<td>11586</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>52052</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-metro county</td>
<td>2047</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9882</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SEER Region</strong></td>
<td>458.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>2137</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13382</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>3296</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16067</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Central</td>
<td>1370</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7193</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>6859</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25367</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Year of diagnosis</strong></td>
<td>198.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2009</td>
<td>6737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26481</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>79.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td>N</td>
<td>LVC%</td>
<td>CM</td>
<td>N</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Preoperative chest radiography</td>
<td>2010</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>280</td>
<td>13.9</td>
<td>1730</td>
<td>86.1</td>
</tr>
<tr>
<td>No</td>
<td>13382</td>
<td>18.2</td>
<td>60279</td>
<td>81.8</td>
</tr>
<tr>
<td>Preoperative echocardiography</td>
<td>1008</td>
<td>1.33%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>140</td>
<td>13.9</td>
<td>868</td>
<td>86.1</td>
</tr>
<tr>
<td>No</td>
<td>13522</td>
<td>18.1</td>
<td>61141</td>
<td>81.9</td>
</tr>
<tr>
<td>Arthroscopic surgery for knee osteoarthritis</td>
<td>33</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Hypercoagulability testing for patients with DVT</td>
<td>78</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>13</td>
<td>16.7</td>
<td>65</td>
<td>83.3</td>
</tr>
<tr>
<td>No</td>
<td>13649</td>
<td>18.1</td>
<td>61944</td>
<td>81.9</td>
</tr>
<tr>
<td>Stress echocardiography for CAD or risk assessment</td>
<td>1916</td>
<td>2.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>288</td>
<td>15.0</td>
<td>1628</td>
<td>85.0</td>
</tr>
<tr>
<td>No</td>
<td>13374</td>
<td>18.1</td>
<td>60381</td>
<td>81.9</td>
</tr>
<tr>
<td>Laminectomy or spinal fusion</td>
<td>348</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>55</td>
<td>15.8</td>
<td>293</td>
<td>84.2</td>
</tr>
<tr>
<td>No</td>
<td>13607</td>
<td>18.1</td>
<td>61716</td>
<td>81.9</td>
</tr>
<tr>
<td>Fiberoptic laryngoscopy for sinusitis</td>
<td>636</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>94</td>
<td>14.8</td>
<td>542</td>
<td>85.2</td>
</tr>
<tr>
<td>No</td>
<td>13568</td>
<td>18.1</td>
<td>61467</td>
<td>81.9</td>
</tr>
<tr>
<td>Routine monitoring of digoxin in CHF patients</td>
<td>613</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>95</td>
<td>15.5</td>
<td>518</td>
<td>84.5</td>
</tr>
<tr>
<td>No</td>
<td>13567</td>
<td>18.1</td>
<td>61491</td>
<td>81.9</td>
</tr>
<tr>
<td>EEG or imaging monitoring in patients with syncope</td>
<td>1388</td>
<td>1.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>247</td>
<td>17.8</td>
<td>1141</td>
<td>82.2</td>
</tr>
<tr>
<td>No</td>
<td>13415</td>
<td>18.1</td>
<td>60868</td>
<td>81.9</td>
</tr>
<tr>
<td>Serological tests for helicobacter pylori</td>
<td>732</td>
<td>0.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>108</td>
<td>14.8</td>
<td>624</td>
<td>85.2</td>
</tr>
<tr>
<td>No</td>
<td>13554</td>
<td>18.1</td>
<td>61385</td>
<td>81.9</td>
</tr>
<tr>
<td>Condition</td>
<td>Alternatives</td>
<td>Yes</td>
<td>No</td>
<td>P-value</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>--------------</td>
<td>-----</td>
<td>----</td>
<td>---------</td>
</tr>
<tr>
<td>MRI in individuals with traumatic brain injury</td>
<td>16</td>
<td>0.02</td>
<td>0.005</td>
<td>0.942</td>
</tr>
<tr>
<td>Traction for low back pain</td>
<td>5759</td>
<td>7.61</td>
<td>5.850</td>
<td>0.016*</td>
</tr>
<tr>
<td>MRI Lumbar Spine for Low Back Pain</td>
<td>5759</td>
<td>7.61</td>
<td>25.889</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Immunoglobulin (IgG, IgE) tests for evaluation of allergy</td>
<td>269</td>
<td>0.36</td>
<td>1.444</td>
<td>0.230</td>
</tr>
<tr>
<td>IVC filters to prevent PE</td>
<td>41</td>
<td>0.05</td>
<td>0.059</td>
<td>0.808</td>
</tr>
<tr>
<td>Renal angioplasty or stent</td>
<td>43</td>
<td>0.06</td>
<td>1.201</td>
<td>0.273</td>
</tr>
<tr>
<td>Vitamin D screening</td>
<td>462</td>
<td>0.61</td>
<td>2.648</td>
<td>0.104</td>
</tr>
<tr>
<td>Head imaging for uncomplicated headache</td>
<td>332</td>
<td>0.44</td>
<td>0.018</td>
<td>0.893</td>
</tr>
<tr>
<td>Imaging for plantar fasciitis</td>
<td>37</td>
<td>0.05</td>
<td>0.516</td>
<td>0.473</td>
</tr>
</tbody>
</table>
Supplemental Figure 7.7 Aim 3 Inclusion & Exclusion Criteria

SEER-Medicare Fee-for-service (FFS) Prostate Cancer Diagnoses 2005-2014 502156

No HMO, Continuous Enrollment in Medicare FFS A & B 13229

Not Diagnosed at Autopsy 132207

Age $\geq 66$ at Diagnosis 123033

AJCC stage $\leq T2a$ 98287

Gleason Score & PSA Value Available 75671
Appendix 7.8 Aim 1 python codes

import pandas as pd
import numpy as np

feature_names for SHAP analyses below

use the list to select a subset of the original DataFrame

X = dataset[feature_names]

print the first 5 rows
X.head()

# Instantiate Y with dependent/label var, X with indepdent/feature vars
y = dataset['trmt']

# Split data into train and test data sets; 80/30
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.30, random_state=21)

from collections import Counter

# count # of examples in each class
counter = Counter(y)

# estimate scale_pos_weight value, assuming the class labels are 0 and 1. Modify according to your use-case
weight = counter[0] / counter[1]

Calibration before tune

from sklearn.calibration import CalibratedClassifierCV
import seaborn as sns

fig, ax = plt.subplots(1, figsize=(12, 6))

# Create an uncorrected classifier.
clf = xgc
clf.fit(X_train, y_train)
y_test_predict_proba = clf.predict_proba(X_test)[:, 1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)

plt.plot(mean_predicted_value, fraction_of_positives, 's-', color='red', label='Uncalibrated')

# Create a corrected classifier.
clf_sigmoid = CalibratedClassifierCV(clf, cv=3, method='sigmoid')
clf_sigmoid.fit(X_train, y_train)
y_test_predict_proba = clf_sigmoid.predict_proba(X_test)[:1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)

plt.plot(mean_predicted_value, fraction_of_positives, 's-', label='Calibrated (Platt)')
plt.plot([0, 1], [0, 1], '--', color='gray')

dsns.despine(left=True, bottom=True)
plt.gca().xaxis.set_ticks_position('none')
plt.gca().yaxis.set_ticks_position('none')
plt.gca().legend()
plt.title(r'$XGClassifier$ Sample Calibration Curve', fontsize=20); pass

d from sklearn.metrics import roc_auc_score, accuracy_score, recall_score, precision_score, f1_score, cohen_kappa_score

# Performance for train
train_y_pred = clf_sigmoid.predict(X_train)
auc = roc_auc_score(y_train, train_y_pred)
print("AUC performance for train : ", auc)

# Calc bal_accuracy, recall, and average positive prediction score
recall = recall_score(y_train, train_y_pred)
print("Recall performance for train : ", recall)
f1 = f1_score(y_train, train_y_pred)
print("F1 performance for train : ", f1)
prcsn = precision_score(y_train, train_y_pred, labels=np.unique(train_y_pred))
print("Precision performance for train : ", prcsn)

# Performance for train
y_pred = clf_sigmoid.predict(X_test)
auc_t = roc_auc_score(y_test, y_pred)
print("AUC performance for test : ", auc_t)

# Calc bal_accuracy, recall, and average positive prediction score
recall_t = recall_score(y_test, y_pred)
print("Recall performance for test : ", recall_t)
f1_t = f1_score(y_test, y_pred)
print("F1 performance for test : ", f1_t)
prcsn_t = precision_score(y_test, y_pred, labels=np.unique(y_pred))
print("Precision performance for test : ", prcsn_t)

AUC performance for train : 0.6787989250338197
Recall performance for train : 0.9956783454236765
F1 performance for train : 0.8606497231672336
Precision performance for train : 0.7578712406015038
AUC performance for test : 0.5588647894037931
Recall performance for test: 0.958139534883721
F1 performance for test: 0.8089412475456879
Precision performance for test: 0.6999477260846837

import xgboost as xgb # XGBoost
# Preliminary model w/o tuning or cross-validation
clf_xgb = xgb.XGBClassifier(objective='binary:logistic', scale_pos_weight=weight, seed=42)
clf_xgb.fit(X_train, y_train, verbose=True,
           ## the next three arguments set up early stopping.
           early_stopping_rounds=10,
           eval_metric=['auc'],
           eval_set=[(X_test, y_test)])

XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
               colsample_bynode=1, colsample_bytree=1, gamma=0, gpu_id=-1,
               importance_type='gain', interaction_constraints='',
               learning_rate=0.300000012, max_delta_step=0, max_depth=6,
               min_child_weight=1, missing=nan, monotone_constraints='()',
               n_estimators=100, n_jobs=0, num_parallel_tree=1, random_state=42,
               reg_alpha=0, reg_lambda=1, scale_pos_weight=0.4955790381712314,
               seed=42, subsample=1, tree_method='exact', validate_parameters=1,
               verbosity=None)

# predict probabilities
xgb_probs = clf_xgb.predict_proba(X_test)
# keep probabilities for the positive outcome only
xgb_probs = xgb_probs[:, 1]

import matplotlib.pyplot as plt
prediction = xgb_probs

plt.figure(figsize=(15, 7))
plt.hist(prediction[y_test==0], bins=50, label='Negatives')
plt.hist(prediction[y_test==1], bins=50, label='Positives', alpha=0.7, color='r')
plt.xlabel('Probability of being Positive Class', fontsize=25)
plt.ylabel('Number of records in each bucket', fontsize=25)
plt.legend(fontsize=15)
plt.tick_params(axis='both', labelsize=25, pad=5)
plt.show()

# Performance for train
train_y_pred = clf_xgb.predict(X_train)
auc = roc_auc_score(y_train, train_y_pred)
print("AUC performance for train : ", auc)

# Calc bal_accuracy, recall, and average positive prediction score
recall = recall_score(y_train, train_y_pred)
print("Recall performance for train : ", recall)
f1 = f1_score(y_train, train_y_pred)
print("F1 performance for train : ", f1)
prcsn = precision_score(y_train, train_y_pred, labels=np.unique(train_y_pred))
print("Precision performance for train : ", prcsn)

# Performance for train
y_pred = clf_xgb.predict(X_test)
auc_t = roc_auc_score(y_test, y_pred)
print("AUC performance for test : ", auc_t)

# Calc bal_accuracy, recall, and average positive prediction score
recall_t = recall_score(y_test, y_pred)
print("Recall performance for test : ", recall_t)
f1_t = f1_score(y_test, y_pred)
print("F1 performance for test : ", f1_t)
prcsn_t = precision_score(y_test, y_pred, labels=np.unique(y_pred))
print("Precision performance for test : ", prcsn_t)
kp_t = cohen_kappa_score(y_test, y_pred)
print("Cohen Kappa for test : ", kp_t)

AUC performance for train : 0.7268423776363454
Recall performance for train : 0.7471832072850748
F1 performance for train : 0.7892076948157809
Precision performance for train : 0.8362411470029366
AUC performance for test : 0.6110343978606433
Recall performance for test : 0.6715563506261181
F1 performance for test : 0.7101778282255014
Precision performance for test : 0.753512645523886
Cohen Kappa for test : 0.2101154238033613

from sklearn.model_selection import GridSearchCV # cross validation and tuning
from sklearn.model_selection import RepeatedStratifiedKFold

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

## ROUND 0
param_grid = {
    'scale_pos_weight': [weight]
}

## NOTE: To speed up cross validation, and to further prevent overfitting
## we are only using a random subset of the data (90%) and are only
## using a random subset of the features (columns) (50%) per tree.
optimal_params_0 = GridSearchCV(
estimator=xgb.XGBClassifier(objective='binary:logistic',
    seed=42,
    subsample=0.9,
    colsample_bytree=0.5),
    param_grid=param_grid,
    scoring='roc_auc', ## see https://scikit-learn.org/stable/modules/model_evaluation.html#scoring-parameter
    verbose=0,  # NOTE: If you want to see what Grid Search is doing, set verbose=2
    n_jobs = -1,
    cv = cv)
)

optimal_params_0.fit(X_train,
    y_train,
    early_stopping_rounds=10,
    eval_metric=['auc'],
    eval_set=[(X_test, y_test)],
    verbose=False)

print(optimal_params_0.best_params_)
{'scale_pos_weight': 0.4955790381712314}

# Performance for train
train_y_pred = optimal_params_0.predict(X_train)
auc = roc_auc_score(y_train, train_y_pred)
recall = recall_score(y_train, train_y_pred)
print("AUC for train : ", auc)
print("Recall for train : ", recall)
AUC for train : 0.6834020319039793
Recall for train : 0.7076709368729742

plot_confusion_matrix(optimal_params_0,
    X_test,
    y_test,
    values_format='d',
    display_labels=['No Treatment', 'Treatment'])
<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x1799dd47fa0>

png
png

# Performance for train
train_y_pred = optimal_params_0.predict(X_train)
auc = roc_auc_score(y_train, train_y_pred)
print("AUC performance for train : ", auc)
# Calc bal_accuracy, recall, and average positive prediction score
recall = recall_score(y_train, train_y_pred)

# Calc bal_accuracy, recall, and average positive prediction score
print("Recall performance for train : ", recall)
f1 = f1_score(y_train, train_y_pred)
print("F1 performance for train : ", f1)
prcsn = precision_score(y_train, train_y_pred,labels=np.unique(train_y_pred))
print("Precision performance for train : ", prcsn)

# Performance for train
y_pred = optimal_params_0.predict(X_test)
auc_t = roc_auc_score(y_test, y_pred)
print("AUC performance for test : ", auc_t)

# Calc bal_accuracy, recall, and average positive prediction score
recall_t = recall_score(y_test, y_pred)
print("Recall performance for test : ", recall_t)
f1_t = f1_score(y_test, y_pred)
print("F1 performance for test : ", f1_t)
prcsn_t = precision_score(y_test, y_pred,labels=np.unique(y_pred))
print("Precision performance for test : ", prcsn_t)

AUC performance for train : 0.6834020319039793
Recall performance for train : 0.7076709368729742
F1 performance for train : 0.7538018906699548
Precision performance for train : 0.8063665142455153
AUC performance for test : 0.5978129739102193
Recall performance for test : 0.643649373881932
F1 performance for test : 0.6911256242796773
Precision performance for test : 0.7461634176690171

# Use recommended XGboost approach or tune for best performance metric.
weight = optimal_params_0.best_params_['scale_pos_weight']

## ROUND 1

n_estimators = [20, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900]
learning_rate = [0.0001, 0.001, 0.01, 0.1, 0.2]
scale_pos_weight = [weight]
param_grid_1 = dict(scale_pos_weight=scale_pos_weight,
 learning_rate=learning_rate,
 n_estimators=n_estimators)

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

optimal_params_1 = GridSearchCV(
estimator=xgb.XGBClassifier(objective='binary:logistic',
 seed=42,
 subsample=0.9,
 colsample_bytree=0.5),
 param_grid=param_grid_1,
 scoring='roc_auc', ## see https://scikit-learn.org/stable/modules/model_evaluation.html#scoring-parameter
verbose=0, # NOTE: If you want to see what Grid Search is doing, set verbose=2
n_jobs = -1,
cv = cv
)

optimal_params_1.fit(X_train,
    y_train,
    early_stopping_rounds=10,
    eval_metric=['auc'],
    eval_set=[(X_test, y_test)],
    verbose=False)

print(optimal_params_1.best_params_)
{'learning_rate': 0.01, 'n_estimators': 200, 'scale_pos_weight': 0.4955790381712314}

# summarize results
print("Best: %f using %s" % (optimal_params_1.best_score_, optimal_params_1.best_params_))
means = optimal_params_1.cv_results_['mean_test_score']
stds = optimal_params_1.cv_results_['std_test_score']
params = optimal_params_1.cv_results_['params']
for mean, stdev, param in zip(means, stds, params):
    print("%f (%f) with: %r" % (mean, stdev, param))

# Performance for train
train_y_pred = optimal_params_1.predict(X_train)
auc = roc_auc_score(y_train, train_y_pred)
print("Performance for train : ", auc)

# Performance for test
y_pred_rd1 = optimal_params_1.predict(X_test)
auc_t_rd1 = roc_auc_score(y_test, y_pred_rd1)
print("AUC performance for test : ", auc_t_rd1)

# Calc bal_accuracy, recall, and average positive prediction score
recall_t_rd1 = recall_score(y_test, y_pred_rd1)
print("Recall performance for test : ", recall_t_rd1)
aps_t_rd1 = average_precision_score(y_test, y_pred)
print("APS performance for test: ", aps_t_rd1)

f1_t_rd1 = f1_score(y_test, y_pred_rd1)
print("F1 performance for test : ", f1_t_rd1)
prcsn_t_rd1 = precision_score(y_test, y_pred_rd1,labels=np.unique(y_pred_rd1))
print("Precision performance for test ", prcsn_t_rd1)
accry_t = accuracy_score(y_test, y_pred_rd1)
print("Accuracy performance for test : ", accry_t)

Performance for train : 0.6850635083482383
AUC performance for test : 0.6114555640824836
Recall performance for test : 0.6826475849731664
APS performance for test: 0.7196331537267628
F1 performance for test: 0.7158131682611143
Precision performance for test: 0.7523659305993691
Accuracy performance for test: 0.6359048305695746

learning_rate = optimal_params_1.best_params_['learning_rate']
n_estimators = optimal_params_1.best_params_['n_estimators']

## ROUND 2
from sklearn.model_selection import RepeatedStratifiedKFold

param_grid_2 = {
    'max_depth': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14],
    'min_child_weight': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14],
    'scale_pos_weight': [weight],
    'learning_rate': [learning_rate],
    'n_estimators': [n_estimators]
}

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

optimal_params_2 = GridSearchCV(
    estimator=xgb.XGBClassifier(objective='binary:logistic',
    seed=42,
    subsample=0.9,
    colsample_bytree=0.5),
    param_grid=param_grid_2,
    scoring='roc_auc', ## see https://scikit-learn.org/stable/modules/model_evaluation.html#scoring-parameter
    verbose=3, # NOTE: If you want to see what Grid Search is doing, set verbose=2
    n_jobs = -1,
    cv = cv
)

optimal_params_2.fit(X_train, y_train,
    early_stopping_rounds=10,
    eval_metric=['auc'],
    eval_set=[(X_test, y_test)],
    verbose=False)

print(optimal_params_2.best_params_)

{'learning_rate': 0.01, 'max_depth': 12, 'min_child_weight': 10, 'n_estimators': 200, 'scale_pos_weight': 0.4955790381712314}

# Performance for train
train_y_pred = optimal_params_2.predict(X_train)
auc = roc_auc_score(y_train, train_y_pred)
print("Performance for train : ", auc)
# Performance for test

```python
y_pred_rd2 = optimal_params_2.predict(X_test)
auc_t_rd2 = roc_auc_score(y_test, y_pred_rd2)
print("AUC performance for test : ", auc_t_rd2)

aps_t_rd2 = average_precision_score(y_test, y_pred_rd2)
print("APS performance for test: ", aps_t_rd2)
```

# Calc bal_accuracy, recall, and average positive prediction score

```python
recall_t_rd2 = recall_score(y_test, y_pred_rd2)
print("Recall performance for test : ", recall_t_rd2)

f1_t_rd2 = f1_score(y_test, y_pred_rd2)
print("F1 performance for test : ", f1_t_rd2)
presn_t_rd2 = precision_score(y_test, y_pred_rd2, labels=np.unique(y_pred_rd2))
print("Precision performance for test : ", presn_t_rd2)
accry_t = accuracy_score(y_test, y_pred_rd2)
print("Accuracy performance for test : ", accry_t)
```

Performance for train : 0.7176366178513388
AUC performance for test : 0.618159650285361
APS performance for test: 0.7306195103281976
Recall performance for test : 0.67269585098
F1 performance for test : 0.713371276797571
Precision performance for test : 0.7592891760904685
Accuracy performance for test : 0.6368661379476087

min_child_weight = optimal_params_2.best_params_['min_child_weight']
max_depth = optimal_params_2.best_params_['max_depth']

## ROUND 3

```python
param_grid_3 = {
    'max_depth': [max_depth],
    'min_child_weight': [min_child_weight],
    'scale_pos_weight': [weight],
    'learning_rate': [learning_rate],
    'n_estimators': [n_estimators],
    'gamma': [0, 0.001, 0.002, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5]
}
```

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

```python
optimal_params_3 = GridSearchCV(  
estimator=xgb.XGBClassifier(objective='binary:logistic',  
    seed=42,  
    subsample=0.9,  
    colsample_bytree=0.5),  
    param_grid=param_grid_3,  
    scoring='roc_auc',  
    ## see https://scikit-learn.org/stable/modules/model_evaluation.html#scori
```
ng-parameter
    verbose=0, # NOTE: If you want to see what Grid Search is doing, set verbose=2
    n_jobs = -1,
    cv = cv
)

optimal_params_3.fit(X_train,
    y_train,
    early_stopping_rounds=10,
    eval_metric=['auc'],
    eval_set=[(X_test, y_test)],
    verbose=False)

print(optimal_params_3.best_params_)
{'gamma': 0.01, 'learning_rate': 0.01, 'max_depth': 12, 'min_child_weight': 10, 'n_estimators': 200, 'scale_pos_weight': 0.4955790381712314}

# Performance for train
train_y_pred = optimal_params_3.predict(X_train)

auc = roc_auc_score(y_train, train_y_pred)
print("Performance for train : ", auc)

# Performance for test
y_pred_rd3 = optimal_params_3.predict(X_test)

auc_t_rd3 = roc_auc_score(y_test, y_pred_rd3)
print("AUC performance for test : ", auc_t_rd3)

# average precision score
aps_t_rd3 = average_precision_score(y_test, y_pred_rd3)
print("APS performance for test: ", aps_t_rd3)

# Calc bal_accuracy, recall, and average positive prediction score
recall_t_rd3 = recall_score(y_test, y_pred_rd3)
print("Recall performance for test : ", recall_t_rd3)

f1_t_rd3 = f1_score(y_test, y_pred_rd3)
print("F1 performance for test : ", f1_t_rd3)
prcsn_t_rd3 = precision_score(y_test, y_pred_rd3,labels=np.unique(y_pred_rd3))
print("Precision performance for test : ", prcsn_t_rd3)
accry_t = accuracy_score(y_test, y_pred_rd3)
print("Accuracy performance for test : ", accry_t)

Performance for train : 0.7176366178513388
AUC performance for test : 0.6183385411619263
APS performance for test: 0.7307162431227368
Recall performance for test : 0.6729874776386404
F1 performance for test : 0.7135811836115326
Precision performance for test : 0.7593863544610416
Accuracy performance for test : 0.6371064647921173
## ROUND 4

param_grid_4 = {
    'max_depth': [max_depth],
    'min_child_weight': [min_child_weight],
    'scale_pos_weight': [weight],
    'learning_rate': [learning_rate],
    'n_estimators': [n_estimators],
    'gamma': [gamma],
    'reg_lambda': [1, 2, 3, 4, 5, 10, 50],
    'reg_alpha': [1, 2, 3, 4, 5, 10, 50]
}

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

optimal_params_4 = GridSearchCV(
    estimator=xgb.XGBClassifier(objective='binary:logistic',
    seed=42,
    subsample=0.9,
    colsample_bytree=0.5),
    param_grid=param_grid_4,
    scoring='roc_auc',
    verbose=0,
    # NOTE: If you want to see what Grid Search is doing, set verbose=2
    n_jobs=-1,
    cv=cv)

optimal_params_4.fit(X_train,
    y_train,
    early_stopping_rounds=10,
    eval_metric=['auc'],
    eval_set=[(X_test, y_test)],
    verbose=False)

print(optimal_params_4.best_params_)

{'gamma': 0.01, 'learning_rate': 0.01, 'max_depth': 12, 'min_child_weight': 10, 'n_estimators': 200, 'reg_alpha': 1, 'reg_lambda': 3, 'scale_pos_weight': 0.4955790381712314}

# Performance for train
train_y_pred = optimal_params_4.predict(X_train)
auc = roc_auc_score(y_train, train_y_pred)
print("Performance for train : ", auc)

# Performance for test
y_pred_rd4 = optimal_params_4.predict(X_test)
auc_t_rd4 = roc_auc_score(y_test, y_pred_rd4)
print("AUC performance for test : ", auc_t_rd4)

# Calc bal_accuracy, recall, and average positive prediction score
recall_t_rd4 = recall_score(y_test, y_pred_rd4)
print("Recall performance for test : ", recall_t_rd4)

# average precision score
aps_t_rd4 = average_precision_score(y_test, y_pred_rd4)
print("APS performance for test: ", aps_t_rd4)

f1_t_rd4 = f1_score(y_test, y_pred_rd4)
print("F1 performance for test : ", f1_t_rd4)

prcsn_t_rd4 = precision_score(y_test, y_pred_rd4,labels=np.unique(y_pred_rd4))
print("Precision performance for test : ", prcsn_t_rd4)

accry_t = accuracy_score(y_test, y_pred_rd4)
print("Accuracy performance for test : ", accry_t)

Performance for train :  0.6860186303260307
AUC performance for test :  0.6065291241156949
Recall performance for test :  0.6654740608228981
APS performance for test:  0.7242139745672891
F1 performance for test :  0.7054807509956381
Precision performance for test :  0.7506053268765133
Accuracy performance for test :  0.6267724104782504

reg_lambda = optimal_params_4.best_params_['reg_lambda']
reg_alpha = optimal_params_4.best_params_['reg_alpha']

## ROUND 5

param_grid_5 = {
    'max_depth': [max_depth],
    'min_child_weight': [min_child_weight],
    'scale_pos_weight': [weight],
    'learning_rate': [learning_rate],
    'n_estimators': [n_estimators],
    'gamma': [gamma],
    'reg_lambda': [reg_lambda],
    'reg_alpha': [reg_alpha],
    'subsample': [0.3, 0.5, 0.7, 0.9],
    'colsample_bytree': [0.3, 0.5, 0.7, 0.9]
}

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

optimal_params_5 = GridSearchCV(
    estimator=xgb.XGBClassifier(objective='binary:logistic',
                                seed=42),
    param_grid=param_grid_5,
    scoring='roc_auc',
    )

# The Grid Search process

```
ng-parameter
    verbose=0, # NOTE: If you want to see what Grid Search is doing, set verbose=2
    n_jobs = -1,
    cv = cv
)

optimal_params_5.fit(X_train,
    y_train,
    early_stopping_rounds=10,
    eval_metric=['auc'],
    eval_set=[(X_test, y_test)],
    verbose=False)

print(optimal_params_5.best_params_)
{'colsample_bytree': 0.5, 'gamma': 0.01, 'learning_rate': 0.01, 'max_depth': 12, 'min_child_weight': 1.0, 'n_estimators': 200, 'reg_alpha': 1, 'reg_lambda': 3, 'scale_pos_weight': 0.4955790381712314, 'subsample': 0.9}

# Performance for train
train_y_pred = optimal_params_5.predict(X_train)

auc = roc_auc_score(y_train, train_y_pred)
print("Performance for train : ", auc)

# Performance for test
y_pred_rd5 = optimal_params_5.predict(X_test)

auc_t_rd5 = roc_auc_score(y_test, y_pred_rd5)
print("AUC performance for test : ", auc_t_rd5)

# average precision score
aps_t_rd5 = average_precision_score(y_test, y_pred_rd5)
print("APS performance for test : ", aps_t_rd5)

# Calc bal_accuracy, recall, and average positive prediction score
recall_t_rd5 = recall_score(y_test, y_pred_rd5)

print("Recall performance for test : ", recall_t_rd5)

f1_t_rd5 = f1_score(y_test, y_pred_rd5)
print("F1 performance for test : ", f1_t_rd5)

prcsn_t_rd5 = precision_score(y_test, y_pred_rd5,labels=np.unique(y_pred_rd5))

print("Precision performance for test : ", prcsn_t_rd5)

accry_t = accuracy_score(y_test, y_pred_rd5)

print("Accuracy performance for test : ", accry_t)

Performance for train : 0.6860186303260307
AUC performance for test : 0.6065291241156949
APS performance for test : 0.7242139745672891
Recall performance for test : 0.6654740608228981
F1 performance for test : 0.7054807509956381
Precision performance for test : 0.7506053268765133
Accuracy performance for test : 0.6267724104782504
```
subsample = optimal_params_5.best_params_['subsample']
colsample_bytree = optimal_params_5.best_params_['colsample_bytree']

## ROUND 6

param_grid_6 = {
    'max_depth': [max_depth],
    'min_child_weight': [min_child_weight],
    'scale_pos_weight': [weight],
    'learning_rate': [learning_rate],
    'n_estimators': [n_estimators],
    'gamma': [gamma],
    'reg_lambda': [reg_lambda],
    'reg_alpha': [reg_alpha],
    'subsample': [subsample],
    'colsample_bytree': [colsample_bytree]
}

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

optimal_params_6 = GridSearchCV(
    estimator=xgb.XGBClassifier(objective='binary:logistic',
                                 seed=42),
    param_grid=param_grid_6,
    scoring='roc_auc', ## see https://scikit-learn.org/stable/modules/model_evaluation.html#scoring-parameter
    verbose=0, # NOTE: If you want to see what Grid Search is doing, set verbose=2
    n_jobs = -1,
    cv = cv
)

optimal_params_6.fit(X_train,
                      y_train,
                      early_stopping_rounds=10,
                      eval_metric=['auc'],
                      eval_set=[(X_test, y_test)],
                      verbose=False)

print(optimal_params_6.best_params_)

{'colsample_bytree': 0.5, 'gamma': 0.01, 'learning_rate': 0.01, 'max_depth': 12, 'min_child_weight': 10, 'n_estimators': 200, 'reg_alpha': 1, 'reg_lambda': 3, 'scale_pos_weight': 0.4955790381712314, 'subsample': 0.9}

print('Best: {optimal_params_6.best_score_} using {optimal_params_6.best_params_}')

Best: 0.6860864943064999 using {'colsample_bytree': 0.5, 'gamma': 0.01, 'learning_rate': 0.01, 'max_depth': 12, 'min_child_weight': 10, 'n_estimators': 200, 'reg_alpha': 1, 'reg_lambda': 3, 'scale_pos_weight': 0.4955790381712314, 'subsample': 0.9}

from sklearn.model_selection import cross_validate
xgbcl = optimal_params_6.best_estimator_

kfold = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)

# refit the model on k-folds to get stable avg error metrics
scores = cross_validate(estimator=xgbcl, X=X_train, y=y_train, cv=kfold, n_jobs=-1, scoring=['accuracy', 'roc_auc', 'average_precision', 'precision', 'recall', 'f1', 'neg_log_loss'])

print('Training 5-fold Cross Validation Results:
')
print('AUC: ', scores['test_roc_auc'].mean())
print('Average precision: ', scores['test_average_precision'].mean())
print('Accuracy: ', scores['test_accuracy'].mean())
print('Precision: ', scores['test_precision'].mean())
print('Recall: ', scores['test_recall'].mean())
print('F1: ', scores['test_f1'].mean())
print('Neg.LogLoss: ', scores['test_neg_log_loss'].mean(), '
')

Training 5-fold Cross Validation Results:

AUC: 0.6876457369330226
Average precision: 0.7995369612229754
Accuracy: 0.6535853497120976
Precision: 0.763820977300308
Recall: 0.6962997945567067
F1: 0.728415533297554
Neg.LogLoss: -0.6369035750460006

import sklearn.metrics as metrics

# Fit the final model
xgbcl.fit(X_train, y_train)

# Generate predictions against our training and test data
pred_train = xgbcl.predict(X_train)
proba_train = xgbcl.predict_proba(X_train)
pred_test = xgbcl.predict(X_test)
proba_test = xgbcl.predict_proba(X_test)

# Print model report
print("Classification report (Test): \
")
print(metrics.classification_report(y_test, pred_test))
print("Confusion matrix (Test): \
")
print(metrics.confusion_matrix(y_test, pred_test)/len(y_test))

print ('Train Accuracy:', metrics.accuracy_score(y_train, pred_train))
print ('Test Accuracy:', metrics.accuracy_score(y_test, pred_test))

print ('Train AUC:', metrics.roc_auc_score(y_train, proba_train[:,1]))
print ('Test AUC:', metrics.roc_auc_score(y_test, proba_test[:,1]))
# calculate the fpr and tpr for all thresholds of the classification
train_fpr, train_tpr, train_threshold = metrics.roc_curve(y_train, proba_train[:,1])
test_fpr, test_tpr, test_threshold = metrics.roc_curve(y_test, proba_test[:,1])

train_roc_auc = metrics.auc(train_fpr, train_tpr)
test_roc_auc = metrics.auc(test_fpr, test_tpr)

import matplotlib.pyplot as plt
fig, ax = plt.subplots(figsize=[7,5])
plt.title('Receiver Operating Characteristic')
plt.plot(train_fpr, train_tpr, 'b', label = 'Train AUC = %0.2f % train_roc_auc')
plt.plot(test_fpr, test_tpr, 'g', label = 'Test AUC = %0.2f % test_roc_auc')
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1], 'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('True Positive Rate')
plt.xlabel('False Positive Rate')
plt.show()

# plot feature importance
xgb.plot_importance(xgbcl, importance_type='gain');

Classification report (Test):

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
<th>f1-score</th>
<th>support</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.44</td>
<td>0.55</td>
<td>0.49</td>
<td>1366</td>
</tr>
<tr>
<td>1.0</td>
<td>0.75</td>
<td>0.67</td>
<td>0.71</td>
<td>2795</td>
</tr>
</tbody>
</table>

accuracy: 0.63 4161
macro avg: 0.60 0.61 0.60 4161
weighted avg: 0.65 0.63 0.63 4161

Confusion matrix (Test):

[[0.17976448 0.14852199]
 [0.2247056  0.44700793]]

Train Accuracy: 0.6966731898238747
Test Accuracy: 0.6267724104782504

Train AUC: 0.755213509518965
Test AUC: 0.6567936626008062
from sklearn.calibration import CalibratedClassifierCV

fig, ax = plt.subplots(1, figsize=(12, 6))

# Create an uncorrected classifier.
clf = xgc
clf.fit(X_train, y_train)
y_test_predict_proba = clf.predict_proba(X_test)[:, 1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)
plt.plot(mean_predicted_value, fraction_of_positives, 's-', color='red', label='Uncalibrated')

# Create a corrected classifier.
clf_sigmoid = CalibratedClassifierCV(clf, cv=3, method='sigmoid')
clf_sigmoid.fit(X_train, y_train)
y_test_predict_proba = clf_sigmoid.predict_proba(X_test)[:, 1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)
plt.plot(mean_predicted_value, fraction_of_positives, 's-', label='Calibrated (Platt)')
plt.plot([0, 1], [0, 1], '--', color='gray')

# Create a corrected classifier.
clf_sigmoid_tune = CalibratedClassifierCV(xgbcl, cv=3, method='sigmoid')
clf_sigmoid_tune.fit(X_train, y_train)
y_test_predict_proba = clf_sigmoid_tune.predict_proba(X_test)[:, 1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)
plt.plot(mean_predicted_value, fraction_of_positives, 's-', label='Calibrated (Platt), After tune')
plt.plot([0, 1], [0, 1], '--', color='green')

sns.despine(left=True, bottom=True)
plt.gca().xaxis.set_ticks_position('none')
plt.gca().yaxis.set_ticks_position('none')
plt.legend()
plt.title(r"$XGClassifier$ Sample Calibration Curve", fontsize=20); pass

png

import sklearn.metrics as metrics

# Fit the final model
clf_sigmoid_tune.fit(X_train, y_train)

# Generate predictions against our training and test data
pred_train = clf_sigmoid_tune.predict(X_train)
proba_train = clf_sigmoid_tune.predict_proba(X_train)
pred_test = clf_sigmoid_tune.predict(X_test)
proba_test = clf_sigmoid_tune.predict_proba(X_test)
# Print model report
print("Classification report (Test): \
")
print(metrics.classification_report(y_test, pred_test))
print("Confusion matrix (Test): \
")
print(metrics.confusion_matrix(y_test, pred_test)/len(y_test))

print ("Train Accuracy:'; metrics.accuracy_score(y_train, pred_train))
print ('Test Accuracy:', metrics.accuracy_score(y_test, pred_test))

print ("Train AUC:'; metrics.roc_auc_score(y_train, proba_train[:,1]))
print ('Test AUC:', metrics.roc_auc_score(y_test, proba_test[:,1]))

print ("Train PR-AUC:'; metrics.average_precision_score(y_train, proba_train[:,1]))
print ('Test PR-AUC:', metrics.average_precision_score(y_test, proba_test[:,1]))

# calculate the fpr and tpr for all thresholds of the classification
train_fpr, train_tpr, train_threshold = metrics.roc_curve(y_train, proba_train[:,1])
test_fpr, test_tpr, test_threshold = metrics.roc_curve(y_test, proba_test[:,1])

train_roc_auc = metrics.auc(train_fpr, train_tpr)
test_roc_auc = metrics.auc(test_fpr, test_tpr)

import matplotlib.pyplot as plt
fig, ax = plt.subplots(figsize=[7,5])
plt.title('Receiver Operating Characteristic')
plt.plot(train_fpr, train_tpr, 'b', label = 'Train AUC = %0.2f' % train_roc_auc)
plt.plot(test_fpr, test_tpr, 'g', label = 'Test AUC = %0.2f' % test_roc_auc)
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1],'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('True Positive Rate')
plt.xlabel('False Positive Rate')
plt.show()

# plot feature importance
xgb.plot_importance(xgbcl, importance_type='gain');

Classification report (Test):

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
<th>f1-score</th>
<th>support</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.60</td>
<td>0.23</td>
<td>0.33</td>
<td>1366</td>
</tr>
<tr>
<td>1.0</td>
<td>0.71</td>
<td>0.92</td>
<td>0.80</td>
<td>2795</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>accuracy</th>
<th>macro avg</th>
<th>weighted avg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.70</td>
<td>0.65</td>
<td>0.67</td>
</tr>
</tbody>
</table>

4161
Confusion matrix (Test):

\[
\begin{bmatrix}
0.0752223 & 0.25306417 \\
0.05118962 & 0.62052391
\end{bmatrix}
\]

Train Accuracy: 0.7289113193943764
Test Accuracy: 0.695746214852199

Train AUC: 0.7814297155324872
Test AUC: 0.6598860127240391

Train PR-AUC: 0.8725053236929292
Test PR-AUC: 0.7798344142467664

# predicted probabilities using tuning and calibration curve
final_model = clf.sigmoid_tune.fit(X_train, np.ravel(y_train))

predictions = final_model.predict(X_test)
preds = final_model.predict_proba(X_test)

# Create a dataframe for the probabilities of treatment
preds_df = pd.DataFrame(preds[:,1], columns = ['prob_treatment'])

print("Accuracy Score Before and After Thresholding: {}, {}".format(accuracy_score(y_test, predictions), accuracy_score(y_test, roc_predictions)))
print("Precision Score Before and After Thresholding: {}, {}".format(precision_score(y_test, predictions), precision_score(y_test, roc_predictions)))
print("Recall Score Before and After Thresholding: {}, {}".format(recall_score(y_test, predictions), recall_score(y_test, roc_predictions)))
print("F1 Score Before and After Thresholding: {}, {}".format(f1_score(y_test, predictions), f1_score(y_test, roc_predictions)))
print("Roc Score Before and After Thresholding: {}, {}".format(roc_auc_score(y_test, predictions), roc_auc_score(y_test, roc_predictions)))

Accuracy Score Before and After Thresholding: 0.695746214852199, 0.6534486902186974
Precision Score Before and After Thresholding: 0.7103163686382393, 0.7420393559928443
Recall Score Before and After Thresholding: 0.9237924865831842, 0.7420393559928443
F1 Score Before and After Thresholding: 0.9237924865831842, 0.7420393559928443
Roc Score Before and After Thresholding: 0.5764643252828073, 0.607104539847092

from sklearn.calibration import CalibratedClassifierCV

fig, ax = plt.subplots(1, figsize=(12, 6))

# Create an uncorrected classifier.
clf = xgbclf
clf.fit(X_train, y_train)
y_test_predict_proba = clf.predict_proba(X_test)[:, 1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)

plt.plot(mean_predicted_value, fraction_of_positives, 's-', color='red', label='Uncalibrated')
# Create a corrected classifier.
clf_sigmoid = CalibratedClassifierCV(clf, cv=3, method='sigmoid')
clf_sigmoid.fit(X_train, y_train)
y_test_predict_proba = clf_sigmoid.predict_proba(X_test)[:, 1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)

plt.plot(mean_predicted_value, fraction_of_positives, 's-', label='Calibrated (Platt)')
plt.plot([0, 1], [0, 1], '--', color='gray')

# Create a corrected classifier.
clf_sigmoid_tune = CalibratedClassifierCV(optimal_params_6.best_estimator_, cv=3, method='sigmoid')
clf_sigmoid_tune.fit(X_train, y_train)
y_test_predict_proba = clf_sigmoid_tune.predict_proba(X_test)[:, 1]
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, y_test_predict_proba, n_bins=10)

plt.plot(mean_predicted_value, fraction_of_positives, 's-', label='Calibrated (Platt), After tune')
plt.plot([0, 1], [0, 1], '--', color='green')

# Create a corrected classifier.
fraction_of_positives, mean_predicted_value = calibration_curve(y_test, roc_predictions, n_bins=10)

plt.plot(mean_predicted_value, fraction_of_positives, 's-', label='Calibrated (Platt), After tune & threshold')
plt.plot([0, 1], [0, 1], '--', color='yellow')
sns.despine(left=True, bottom=True)
plt.gca().xaxis.set_ticks_position('none')
plt.gca().yaxis.set_ticks_position('none')
plt.gca().legend()
plt.title("$XGClassifier$ Sample Calibration Curve", fontsize=20); pass

png
png

import shap
shap_values_list = []
shap_values_list_1000 = []
for calibrated_classifier in clf_sigmoid_tune.calibrated_classifiers_:
    explainer = shap.TreeExplainer(calibrated_classifier.base_estimator)
    shap_values = explainer.shap_values(X_train)
    shap_values_1000 = explainer.shap_values(X_test[:1000])
    shap_values_list_1000.append(shap_values_1000)
    shap_values_list.append(shap_values)
    shap_values = np.array(shap_values_list).sum(axis=0) / len(shap_values_list)

First summary plot is tune
# Shap values for positive class
shap.summary_plot(shap_values, X_train, max_display=15, cmap=plt.get_cmap("cool"))

# Shap values for positive class
shap.summary_plot(shap_values, X_train, max_display=15, plot_type="bar", cmap=plt.get_cmap("cool"))

# Shap values for positive class
shap.summary_plot(shap_values, X_train, max_display=50, plot_type="bar", cmap=plt.get_cmap("cool"))

shap.dependence_plot("Married", shap_values, X_train.values, feature_names=feature_names,interaction_index=None, cmap=plt.get_cmap("cool"))

shap.dependence_plot("Age group 66-74", shap_values, X_train.values, feature_names=feature_names,interaction_index=None, cmap=plt.get_cmap("cool"))

shap.dependence_plot("Care fragmentation", shap_values, X_train.values, feature_names=feature_names,interaction_index=None, cmap=plt.get_cmap("cool"))

shap.dependence_plot("Median income", shap_values, X_train.values, feature_names=feature_names,interaction_index=None, cmap=plt.get_cmap("cool"))

shap.dependence_plot("Life expectancy over 10 years", shap_values, X_train.values, feature_names=feature_names,interaction_index=None, cmap=plt.get_cmap("cool"))

shap.dependence_plot("Multimorbidity group", shap_values, X_train.values, feature_names=feature_names,interaction_index=None, cmap=plt.get_cmap("cool"))
Appendix 7.8 Aim 3 python codes

```python
import pandas as pd
import numpy as np
from scipy.stats import stats, randint
import random
from sklearn.model_selection import GridSearchCV, cross_val_score, learning_curve
from sklearn.model_selection import train_test_split
import xgboost as xgb
from collections import Counter
import itertools
import matplotlib.pyplot as plt
import seaborn as sns

X_vars = [
    "Age at diagnosis",
    "Hispanic",
    "Metro locality",
    "Physical condition total","Mental health condition total",
    "A1c screening","Influenza vaccination","Lipid screening",
    "SEER Northcentral","SEER Northeast","SEER South","SEER West",
    "White race","Black race","Other race",
    "Psychology visit","Primary Care visit",
    "Median income",
    "Urologist density","Radiation Oncologist density",
    "College education","Less than highschool education",
    "PCCI",
    "Married","Separated/Divorced/Widowed","Unmarried","Marital status unknown",
    "Low-risk prostate cancer",
    "Low value care (sum)",
    "Conservative management",
    "Care fragmentation",
    "Diagnostic year 2005",
    'Diagnostic year 2006',
    'Diagnostic year 2007',
    'Diagnostic year 2008',
    'Diagnostic year 2009',
    'Diagnostic year 2010',
    'Diagnostic year 2011',
    'Diagnostic year 2012',
    'Diagnostic year 2013',
    'Diagnostic year 2014'
]

# use the list to select a subset of the original DataFrame
X = dataset[X_vars]
```
y = dataset['notx_12_24maft']
y.head()

0   3404.767935
1   1613.494700
2  54756.131876
3    143.520000
4  3225.969687
Name: notx_12_24maft, dtype: float64

import matplotlib.pyplot as plt
import seaborn as sns; sns.set()

# Organize Data
SR_y = pd.Series(y, name="y (Target Vector Distribution)")

# Plot Data
fig, ax = plt.subplots()
sns.distplot(SR_y, bins=25, color="g", ax=ax)
plt.show()

# X_train, y_train

XGB regressor with grid search

import numpy as np
import pandas as pd
from sklearn import preprocessing
import xgboost as xgb
from xgboost.sklearn import XGBRegressor
import datetime
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import r2_score

y_train_full = dataset['notx_12_24maft']
x_train_full = dataset.drop(['notx_12_24maft'], axis=1)

y_train_full = np.log1p(y_train_full)

Confirm y_ln distribution

import matplotlib.pyplot as plt
import seaborn as sns; sns.set()

# Organize Data
SR_y_ln = pd.Series(y_train_full, name="y (Target Vector Distribution)")

# Plot Data
fig, ax = plt.subplots()
sns.distplot(SR_y_ln, bins=25, color="g", ax=ax)
plt.show()

# Split, check distribution of train/test features
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(x_train_full, y_train_full, test_size=0.3, random_state=1)

print(X_train.shape)
print(X_test.shape)
print(y_train.shape)
print(y_test.shape)

(52969, 45)
(22702, 45)
(52969,)
(22702,)

import xgboost as xgb
from sklearn.metrics import r2_score, mean_squared_error, mean_absolute_error

# set cross-validation and defaults
cv=10
seed=33

#
n_estimators = [500,670,675]
learning_rate = [0.008,0.009,0.01]

param_grid_1 = dict(learning_rate=learning_rate,
                     n_estimators=n_estimators)

xgb_model = XGBRegressor(objective = 'reg:squarederror',
                          seed=seed)

optimal_params_1 = GridSearchCV(
    estimator = xgb_model,
    param_grid = param_grid_1,
    scoring = 'r2', # R2 tune
    scoring = 'neg_mean_squared_error', # MSE
    cv = cv,
    n_jobs = -1,
    verbose = 1
)

optimal_params_1.fit(X_train,
                      y_train,
                      early_stopping_rounds=10,
                      eval_set=[(X_test, y_test)],
verbose=False)

print("Learning rate: ", optimal_params_1.best_params_['learning_rate'])
print("n_estimators: ", optimal_params_1.best_params_['n_estimators'])

Fitting 10 folds for each of 9 candidates, totalling 90 fits

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 12 concurrent workers.
[Parallel(n_jobs=-1)]: Done  26 tasks    | elapsed:  3.8min
[Parallel(n_jobs=-1)]: Done  90 out of  90 | elapsed: 10.1min finished

Learning rate: 0.009
n_estimators: 670

# summarize results
print("Best: %f using %s" % (optimal_params_1.best_score_, optimal_params_1.best_params_))
means = optimal_params_1.cv_results_['mean_test_score']
stds = optimal_params_1.cv_results_['std_test_score']
params = optimal_params_1.cv_results_['params']
for mean, stdev, param in zip(means, stds, params):
    print("%f (%f) with: %r" % (mean, stdev, param))
Best: 0.344432 using {'learning_rate': 0.009, 'n_estimators': 670}
0.339254 (0.025079) with: {'learning_rate': 0.008, 'n_estimators': 500}
0.344367 (0.024457) with: {'learning_rate': 0.008, 'n_estimators': 670}
0.344394 (0.024453) with: {'learning_rate': 0.008, 'n_estimators': 675}
0.342533 (0.024821) with: {'learning_rate': 0.009, 'n_estimators': 500}
0.344432 (0.024475) with: {'learning_rate': 0.009, 'n_estimators': 670}
0.344430 (0.024477) with: {'learning_rate': 0.009, 'n_estimators': 675}
0.343783 (0.024534) with: {'learning_rate': 0.01, 'n_estimators': 500}
0.344329 (0.024393) with: {'learning_rate': 0.01, 'n_estimators': 670}
0.344329 (0.024393) with: {'learning_rate': 0.01, 'n_estimators': 675}

y_train_pred = optimal_params_1.best_estimator_.predict(X_train)
y_test_pred = optimal_params_1.best_estimator_.predict(X_test)

print("XGBregressor evaluating result: ")
print("Train MAE: ", sklearn.metrics.mean_absolute_error(y_train, y_train_pred))
print("Train RMSE: ", np.sqrt(sklearn.metrics.mean_squared_error(y_train, y_train_pred)))
print("Train R2: ", np.sqrt(sklearn.metrics.r2_score(y_train, y_train_pred)))
print("Test MAE: ", sklearn.metrics.mean_absolute_error(y_test, y_test_pred))
print("Test RMSE: ", np.sqrt(sklearn.metrics.mean_squared_error(y_test, y_test_pred)))
print("Test R2: ", np.sqrt(sklearn.metrics.r2_score(y_test, y_test_pred)))

XGBregressor evaluating result:
Train MAE: 1.1343200343381408
Train RMSE: 1.5691095294270383
Train R2: 0.6387264523125636
Test MAE: 1.1727194767606826
Test RMSE: 1.627370752580571
Test R2: 0.5810544023298005

```python
learning_rate = optimal_params_1.best_params_['learning_rate']
n_estimators = optimal_params_1.best_params_['n_estimators']

param_grid_2 = {
    'learning_rate':learning_rate,
    'n_estimators':n_estimators,
    'max_depth':[4,5,6],
    'min_child_weight':[11,12,13]
}

xgb_model = XGBRegressor(objective = 'reg:squarederror',
                        seed=seed)

optimal_params_2 = GridSearchCV(
    estimator = xgb_model,
    param_grid = param_grid_2,
    scoring = 'r2',
    cv = cv,
    n_jobs = -1,
    verbose = 1)

optimal_params_2.fit(X_train,
                      y_train,
                      early_stopping_rounds=10,
                      eval_set=[(X_test, y_test)],
                      verbose=False)

print("max_depth: ", optimal_params_2.best_params_['max_depth'])
print("min_child_weight: ",optimal_params_2.best_params_['min_child_weight'])

Fitting 10 folds for each of 9 candidates, totalling 90 fits

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 12 concurrent workers.
[Parallel(n_jobs=-1)]: Done 26 tasks      | elapsed: 2.9min
[Parallel(n_jobs=-1)]: Done 90 out of 90 | elapsed: 9.2min finished

max_depth: 5
min_child_weight: 12

max_depth = optimal_params_2.best_params_['max_depth']
min_child_weight = optimal_params_2.best_params_['min_child_weight']

y_train_pred = optimal_params_2.best_estimator_.predict(X_train)
y_test_pred = optimal_params_2.best_estimator_.predict(X_test)
```
print('XGBregressor evaluating result:
Train MAE:  1.1557842730348065
Train RMSE:  1.6082375631598405
Train R2:  0.6148797875288853
Test MAE:  1.1723220292261036
Test RMSE:  1.6260654683448836
Test R2:  0.5819676548322498

# gamma tune

param_grid_3 = {'learning_rate':[learning_rate],
                'n_estimators':[n_estimators],
                'max_depth':[max_depth],
                'min_child_weight':[min_child_weight],
                'gamma':[0.003,0.004,0.005,0.006,0.007]}

xgb_model = XGBRegressor(objective = 'reg:squarederror',
                          seed=seed)

optimal_params_3 = GridSearchCV(
estimator = xgb_model,
param_grid = param_grid_3,
scoring = 'r2',  #MAE
scoring = 'neg_mean_squared_error',  #MSE
cv = cv,
n_jobs = -1,
verbose = 1
)

optimal_params_3.fit(X_train,
y_train,
early_stopping_rounds=10,
eval_set=[(X_test, y_test)],
verbose=False)

print("Gamma: ", optimal_params_3.best_params_['gamma'])

Fitting 10 folds for each of 5 candidates, totalling 50 fits

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 12 concurrent workers.
[Parallel(n_jobs=-1)]: Done 26 tasks  | elapsed: 3.5min
Gamma: 0.003

gamma = optimal_params_3.best_params_['gamma']

y_train_pred = optimal_params_3.best_estimator_.predict(X_train)
y_test_pred = optimal_params_3.best_estimator_.predict(X_test)

print('XGBregressor evaluating result: ')
print("Train MAE: ", sklearn.metrics.mean_absolute_error(y_train, y_train_pred))
print("Train RMSE: ", np.sqrt(sklearn.metrics.mean_squared_error(y_train, y_train_pred)))
print("Train R2: ", np.sqrt(sklearn.metrics.r2_score(y_train, y_train_pred)))
print("Test MAE: ", sklearn.metrics.mean_absolute_error(y_test, y_test_pred))
print("Test RMSE: ", np.sqrt(sklearn.metrics.mean_squared_error(y_test, y_test_pred)))
print("Test R2: ", np.sqrt(sklearn.metrics.r2_score(y_test, y_test_pred)))

XGBregressor evaluating result:
Train MAE: 1.1557842730348065
Train RMSE: 1.6082375631598405
Train R2: 0.6148797875288853
Test MAE: 1.1723220292261036
Test RMSE: 1.6260654683448836
Test R2: 0.5819676548322498

# max_delta_step

param_grid_4 = {'learning_rate':[learning_rate],
                'n_estimators':[n_estimators],
                'max_depth':[max_depth],
                'min_child_weight':[min_child_weight],
                'gamma':[gamma],
                'max_delta_step':[10,11,12],
                }

xgb_model = XGBRegressor(objective = 'reg:squarederror',
                          seed=seed)

optimal_params_4 = GridSearchCV(  estimator = xgb_model,
                                param_grid = param_grid_4,
                                scoring = 'r2', #MAE
                                scoring = 'neg_mean_squared_error', #MSE
                                cv = cv,
                                n_jobs = -1,
                                verbose = 1
                              )

optimal_params_4.fit(X_train,
                      y_train,
                      early_stopping_rounds=10,
eval_set=[(X_test, y_test)],
verbose=False)

print("max_delta_step: ", optimal_params_4.best_params_['max_delta_step'])

Fitting 10 folds for each of 5 candidates, totalling 50 fits

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 12 concurrent workers.
[Parallel(n_jobs=-1)]: Done 26 tasks    | elapsed:  3.6min
[Parallel(n_jobs=-1)]: Done 50 out of  50 | elapsed:  5.5min finished

max_delta_step: 11

max_delta_step = optimal_params_4.best_params_['max_delta_step']

y_train_pred = optimal_params_4.best_estimator_.predict(X_train)
y_test_pred = optimal_params_4.best_estimator_.predict(X_test)

print('XGBregressor evaluating result:')
print('Train MAE: ', sklearn.metrics.mean_absolute_error(y_train, y_train_pred))
print('Train RMSE: ', np.sqrt(sklearn.metrics.mean_squared_error(y_train, y_train_pred)))
print('Train R2: ', np.sqrt(sklearn.metrics.r2_score(y_train, y_train_pred)))
print('Test MAE: ', sklearn.metrics.mean_absolute_error(y_test, y_test_pred))
print('Test RMSE: ', np.sqrt(sklearn.metrics.mean_squared_error(y_test, y_test_pred)))
print('Test R2: ', np.sqrt(sklearn.metrics.r2_score(y_test, y_test_pred)))

XGBregressor evaluating result:
Train MAE: 1.1557842730348065
Train RMSE: 1.6082375631598405
Train R2: 0.6148797875288853
Test MAE: 1.1723220292261036
Test RMSE: 1.6260654683448836
Test R2: 0.5819676548322498

# reg_lambda, alpha

param_grid_5 = {'learning_rate':[learning_rate],
    'n_estimators':[n_estimators],
    'max_depth':[max_depth],
    'min_child_weight':[min_child_weight],
    'gamma':[gamma],
    'max_delta_step':[max_delta_step],
    'reg_lambda':[1,2,3,4,5],
    'reg_alpha':[1,2,3,4,5]}

xgb_model = XGBRegressor(objective = 'reg:squarederror',
    seed=seed)

optimal_params_5 = GridSearchCV(
estimator = xgb_model,
param_grid = param_grid_5,
scoring = 'r2',
#scoring = 'neg_mean_absolute_error', #MAE
#scoring = 'neg_mean_squared_error', #MSE
cv = cv,
n_jobs = -1,
verbose = 1
)

optimal_params_5.fit(X_train,
    y_train,
    early_stopping_rounds=10,
    eval_set=[(X_test, y_test)],
    verbose=False
)

print("Reg_lambda: ", optimal_params_5.best_params_['reg_lambda'])
print("Reg_alpha: ", optimal_params_5.best_params_['reg_alpha'])

Fitting 10 folds for each of 25 candidates, totalling 250 fits

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 12 concurrent workers.
[Parallel(n_jobs=-1)]: Done 26 tasks    | elapsed:  3.8min
[Parallel(n_jobs=-1)]: Done 176 tasks   | elapsed: 18.5min
[Parallel(n_jobs=-1)]: Done 250 out of 250 | elapsed: 25.6min finished

Reg_lambda: 2
Reg_alpha: 1

reg_lambda = optimal_params_5.best_params_['reg_lambda']
reg_alpha = optimal_params_5.best_params_['reg_alpha']
y_train_pred = optimal_params_5.best_estimator_.predict(X_train)
y_test_pred = optimal_params_5.best_estimator_.predict(X_test)

print('XGBregressor evaluating result:')
print("Train MAE: ", sklearn.metrics.mean_absolute_error(y_train, y_train_pred))
print("Train RMSE: ", np.sqrt(sklearn.metrics.mean_squared_error(y_train, y_train_pred)))
print("Train R2: ", np.sqrt(sklearn.metrics.r2_score(y_train, y_train_pred)))
print("Test MAE: ", sklearn.metrics.mean_absolute_error(y_test, y_test_pred))
print("Test RMSE: ", np.sqrt(sklearn.metrics.mean_squared_error(y_test, y_test_pred)))
print("Test R2: ", np.sqrt(sklearn.metrics.r2_score(y_test, y_test_pred)))

XGBregressor evaluating result:
Train MAE:  1.1564939892972832
Train RMSE:  1.6098665862703325
Train R2:  0.613853885972447
Test MAE:  1.1725238251572232
Test RMSE:  1.6260069515321633
Test R2:  0.5820085458766149
# subsample, colsample_bytree

```
param_grid_6 = {
    'learning_rate': [learning_rate],
    'n_estimators': [n_estimators],
    'max_depth': [max_depth],
    'min_child_weight': [min_child_weight],
    'gamma': [gamma],
    'max_delta_step': [max_delta_step],
    'reg_lambda': [reg_lambda],
    'reg_alpha': [reg_alpha],
    'colsample_bytree': [0.5, 0.7, 0.9],
    'subsample': [0.5, 0.7, 0.9]
}

xgb_model = XGBRegressor(objective = 'reg:squarederror', seed=seed)

optimal_params_6 = GridSearchCV(
    estimator = xgb_model,
    param_grid = param_grid_6,
    scoring = 'r2',
    #scoring = 'neg_mean_absolute_error', #MAE
    #scoring = 'neg_mean_squared_error', #MSE
    cv = cv,
    n_jobs = -1,
    verbose = 1
)

optimal_params_6.fit(X_train,
    y_train,
    early_stopping_rounds=10,
    eval_set=[[X_test, y_test]],
    verbose=False)

print("colsample_bytree: ", optimal_params_6.best_params_['colsample_bytree'])
print("subsample: ", optimal_params_6.best_params_['subsample'])

Fitting 10 folds for each of 9 candidates, totalling 90 fits

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 12 concurrent workers.
[Parallel(n_jobs=-1)]: Done 26 tasks    | elapsed: 2.6min
[Parallel(n_jobs=-1)]: Done 90 out of 90 | elapsed: 8.2min finished

colsample_bytree: 0.7
subsample: 0.7

colsample_bytree = optimal_params_6.best_params_['colsample_bytree']
subsample = optimal_params_6.best_params_['subsample']
```
```python
y_train_pred = optimal_params_6.best_estimator_.predict(X_train)
y_test_pred = optimal_params_6.best_estimator_.predict(X_test)

print('XGBregressor evaluating result:')
print('Train MAE: ', sklearn.metrics.mean_absolute_error(y_train, y_train_pred))
print('Train RMSE: ', np.sqrt(sklearn.metrics.mean_squared_error(y_train, y_train_pred)))
print('Train R2: ', np.sqrt(sklearn.metrics.r2_score(y_train, y_train_pred)))
print('Test MAE: ', sklearn.metrics.mean_absolute_error(y_test, y_test_pred))
print('Test RMSE: ', np.sqrt(sklearn.metrics.mean_squared_error(y_test, y_test_pred)))
print('Test R2: ', np.sqrt(sklearn.metrics.r2_score(y_test, y_test_pred)))

model= XGBRegressor(booster='gbtree',
    learning_rate= learning_rate,
    reg_alpha=reg_alpha,
    colsample_bytree= colsample_bytree,
    gamma= gamma,
    reg_lambda= reg_lambda,
    max_delta_step= max_delta_step,
    max_depth= max_depth,
    min_child_weight= min_child_weight,
    n_estimators= n_estimators,
    sampling_method= 'uniform',
    subsample= subsample,
    seed=seed)

train = model.fit(X_train, y_train)

XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1,
    colsample_bynode=1, colsample_bytree=0.7, gamma=0.003, gpu_id=-1,
    importance_type='gain', interaction_constraints='',
    learning_rate=0.009, max_delta_step=11, max_depth=5,
    min_child_weight=12, missing=nan, monotone_constraints='()',
    n_estimators=670, n_jobs=0, num_parallel_tree=1, random_state=33,
    reg_alpha=1, reg_lambda=2, sampling_method='uniform',
    scale_pos_weight=1, seed=33, subsample=0.7, tree_method='exact',
    validate_parameters=1, verbosity=None)

variables = [
    "Age at diagnosis",
    "Hispanic",
    "Metro locality",
    "Physical condition total","Mental health condition total",
    "A1c screening","Influenza vaccination","Lipid screening",
]```
"SEER Northcentral","SEER Northeast","SEER South","SEER West",
"White race","Black race","Other race",
"Psychology visit","Primary Care visit",
"Median income",
"Urologist density","Radiation Oncologist density",
"College education","Less than highschool education",
"PCCI",
"Married","Separated/Divorced/Widowed","Unmarried","Marital status unknown",
"Low-risk prostate cancer",
"Low value care (sum)",
"Conservative management",
"Care fragmentation"
]

import shap
explainer = shap.TreeExplainer(model)

shap_values = explainer.shap_values(X_train)

# Shap values for positive class
shap.summary_plot(shap_values, X_train, max_display=10)

# Shap values for positive class
shap.summary_plot(shap_values, X_train, max_display=10, plot_type="bar", cmap=plt.get_cmap("cool"))

shap.dependence_plot("Physical condition total", shap_values, X_train, interaction_index = None)

shap.dependence_plot("Care fragmentation", shap_values, X_train, interaction_index = None)

shap.dependence_plot("Low value care (sum)", shap_values, X_train, interaction_index = None)

shap.dependence_plot("Conservative management", shap_values, X_train, interaction_index = None)

shap.dependence_plot("crfrg_10x", shap_values, X_train, interaction_index = "sum_lowval_12mbef")

shap.dependence_plot("mm_pcci", shap_values, X_train, interaction_index = "sum_lowval_12mbef")

shap.dependence_plot("mm_pcci", shap_values, X_train, interaction_index = "crfrg_10x")

shap.dependence_plot("Low value care (sum)", shap_values, X_train, interaction_index = "Physical condition total")