2020

Development and Validation of Predictive Risk and Multiple Criteria Decision Analysis Models to Evaluate Cardiovascular Outcomes Among Cancer Patients

Purva N. Parab
Virginia Commonwealth University

Development and Validation of Predictive Risk and Multiple Criteria Decision Analysis Models to Evaluate Cardiovascular Outcomes Among Cancer Patients

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

by

Purva Niteen Parab
Bachelor of Pharmacy (B.Pharm), Mumbai University, India, 2016
Master of Science (M.S.), Virginia Commonwealth University School of Pharmacy, Richmond VA, 2018

Advisor:
Pramit A. Nadpara, PhD, MS, BPharm
Assistant Professor
Department of Pharmacotherapy & Outcomes Science

Committee Members:

Norman V. Carroll, PhD, RPh
Professor
Department of Pharmacotherapy & Outcomes Science

Julie A. Patterson, PharmD, PhD
Assistant Professor
Department of Pharmacotherapy & Outcomes Science

Adam Sima, PhD
Affiliate faculty
School of Medicine

D'Arcy P. Mays, PhD
Associate Professor and Chair
Department of Statistical Sciences and Operations Research

Virginia Commonwealth University
Richmond, VA
November 2020

ACKNOWLEDGEMENTS

Success can seldom be achieved without the support of your loved ones and I am no different. There are a lot of people who are responsible for where I am today. My parents being the most important. Without their love and support, I would not have been here. Their faith in me makes me want to be a better person every day. Other than them, my friends here in Richmond, Varsha, Vineet and Palak who became my family in these past years have been a constant support system who have been responsible for keeping me sane all these years. I cannot thank my girl gang Varsha, Aarya and Kimaya enough for being by my side all this while. Varsha who has been a backbone and my cheerleader, Aarya and Kimaya who have seen me grow and have been by my side whether or not we are in the same city. All other friends that I made in the process, Rushabh, Serena, Shwetha and Hrishi have had a huge impact.

I would like to thank Dr.Nadpara, my mentor, and all the other committee members Dr.Carroll, Dr.Patterson, Dr.Sima and Dr.Mays. Dr.Nadpara has been my mentor since masters and has helped me in improving my research skills. Dr.Carroll's critical thinking and attention to detail has been of great help in improving my technical skills. Dr. Patterson's doors have always been open to her students and her immense knowledge has helped greatly in building my topic and making me a better researcher. Dr.Sima's experience with predictive modeling has been very critical in strengthening our study. Dr.Mays has always been available to bounce off some analyses ideas and to answer any statistical questions. Other than these, Dr.McRae has always been available for any concerns that I had and willing to help at any point to make the PhD journey an easier one. I would also like to thank Dr.Wijesinghe and Dr.Arodz for steering me in the right direction

when it came to application of machine learning in healthcare and outcomes research. I would also like to thank Dr.Erin Hickey for her expert opinion in finalizing therapy alternatives and criteria used for the MCDA model. Dr. Salgado and Dr.Holdford have also always had their doors open for any concerns that the students might have.

Finally, I would like to thank all my McGuire 218 colleagues. Vasco and Elena have been my go-to people for any immediate help or even for my venting sessions. Everyone from McGuire 218 who have been a part of this journey have been very helpful in making my research better and helping me grow as a researcher.

All these people have played a major role in supporting me throughout the PhD journey and once again I would like to thank each one of them.

TABLE OF CONTENTS

LIST OF TABLES VI

LIST OF FIGURES IXX

LIST OF ABBREVIATIONS XII

ABSTRACT XIV

CHAPTER 1: INTRODUCTION 1

BACKGROUND 1

LITERATURE REVIEW 3

RATIONALE 11

SPECIFIC AIMS 13

REFERENCES 14

CHAPTER 2: PREDICTIVE RISK MODELING FOR PREDICTION OF CVD RISK AMONG CANCER PATIENTS 20

BACKGROUND 20

CONCEPTUAL FRAMEWORK 27

METHODS AIM 1: TO DEVELOP AND VALIDATE PREDICTIVE RISK MODELS TO ASSESS THE RISK OF CARDIOVASCULAR DISEASES (CVD) AMONG CANCER PATIENTS USING THE MACHINE LEARNING ALGORITHMS 28

RESULTS 39

DISCUSSION 71

APPENDIX 76

REFERENCES 80

CHAPTER 3: BENEFIT-RISK ASSESSMENT OF BREAST CANCER THERAPIES USING MCDA MODEL 84

BACKGROUND 84

METHODS AIM 2: TO ASSESS CARDIOTOXICITY ASSOCIATED WITH TARGETED THERAPIES AS COMPARED TO NON-TARGETED THERAPIES AND DEVELOP A MODEL TO CONDUCT BENEFIT-RISK ASSESSMENT OF THERAPY REGIMENS IN BREAST CANCER PATIENTS 94

RESULTS 110

DISCUSSION 125

REFERENCES 134

CHAPTER 4: CONCLUSION 142

SUMMARY OF FINDINGS 142

FUTURE RESEARCH 145

REFERENCES 148

List Of Tables

Table 1: Summary Of Literature Review.....6

Table 2: Summary Of Literature Review (Aim 1).....23

Table 3: Time Interval Between Cancer And Chronic CVD Diagnoses (N = 12,339).....29

Table 4: Study Variables For Aim 1b.....31

Table 5: Patients With Cancer And CVD Both By Cancer Type In The US From 2005-2015.....39

Table 6: Acute CVD Events In Patients With Cancer In The US From 2005-2015.....39

Table 7: Sociodemographic Characteristics Of Cancer Patients With And Without CVD In The US From 2005-2015.....40

Table 8: Survey Logistic Results For The Likelihood Of Being Diagnosed With CVD Along With Cancer Based On The Sociodemographic Characteristics Associated With Cancer Patients In The US From 2005-2015.....43

Table 9: Baseline Characteristics Of Training And Test Data Splits Used For Prediction.....45

Table 10: Learning Accuracies For Random Forest Model Predicting An Acute CVD Event.....52

Table 11: Learning Accuracies For Random Forest Model Predicting Chronic CVD.....53

Table 12: Variable Important Values For The Random Forest Model Built For Predicting Chronic CVD.....54

Table 13: Variable Important Values For The Random Forest Model Built For Predicting Acute CVD Events.....54

Table 14: Errors Associated With Each Iteration Of Gradient Boosting Algorithm For Predicting Acute CVD Event.....55

Table 15: Errors Associated With Each Iteration Of Gradient Boosting Algorithm For Predicting Chronic CVD Event.....55

Table 16: Comparison Of Acute CVD Prediction Models 61

Table 17: Comparison Of Chronic CVD Prediction Models.....62

Table 18: Accuracy Metrics For Internal Validation Of Acute CVD Events Prediction Model.....63

Table 19: Accuracy Metrics For Internal Validation Of Chronic CVD Events Prediction Model 64

Table 20: Accuracy Metrics For External Validation Of Acute CVD Event Prediction Model 65

Table 21: Accuracy Metrics For External Validation Of Chronic CVD Event Prediction Model.....66

Table 22: Summary Of Literature Review (Aim 2).....90

Table 23: Study Variables For Aim 2a Identified From The FAERS 2005-2015 Database 96

Table 24: Treatment Alternatives Considered In The MCDA Model.....99

Table 25: Literature Sources For Model Inputs..... 103

Table 26: Model Inputs For The MCDA Model – Performance Table 104

Table 27: Values Used For Sensitivity Analyses Of The MCDA Model..... 109

Table 28:Descriptive Characteristics Associated With Cardiotoxic Events In Breast Cancer Patients In The US From FAERS 2005-2015 Files..... 110

Table 29: 2x2 Table For Those Reporting A Cardiovascular Adverse Event Across The Breast Cancer Therapy Groups In The US From FAERS 2005-2015 Files..... 112

Table 30: Unadjusted Reporting Odds Ratios For The Association Between Targeted Therapy And Cardiovascular Adverse Event In Breast Cancer Patients In The US From FAERS 2005-2015 Files..... 112

Table 31: Reporting Odds Ratios On Adjusting For Other Covariates Identified From The FAERS 2005-2015 Files..... 113

Table 32: Scores Assigned To Criteria Used In The MCDA Model For HER 2 Positive Therapy Regimen Alternatives..... 117

Table 33: Weights Assigned To Criteria Used In The MCDA Model 122

Table 34:Different Sets Of Weights Used To Conduct Sensitivity Analyses To Assess The Robustness Of The MCDA Model 125

LIST OF FIGURES

Figure 1: PRISMA Flowchart Of Literature Review 5

Figure 2: PRISMA Flowchart Of Literature Review (Aim 1) 22

Figure 3: ROC For Acute CVD Prediction Using A Regression Model.....50

Figure 4: ROC For Chronic CVD Prediction Using A Regression Model 51

Figure 5: Learning Curves For A Random Forest Model Predicting An Acute CVD Event.....52

Figure 6: Learning Curves For A Random Forest Model Predicting Chronic CVD.....53

Figure 7: ROC Obtained With A Gradient Boosting Model For Acute CVD Event 56

Figure 8: ROC Obtained With A Gradient Boosting Model For Chronic CVD.....56

Figure 9: Variable Importance Plot Based On The Gradient Boosting Model For Acute CVD Events 57

Figure 10: Variable Importance Plot Based On The Gradient Boosting Model For Chronic CVD Events 58

Figure 11: Learning Curves For The Deep Learning Acute CVD Prediction Model 60

Figure 12: Learning Curves For The Deep Learning Chronic CVD Prediction Model.....60

Figure 13: Learning Curves For Internal Validation Of Random Forest Model Using South And West Census Regions For Acute CVD Events Prediction.....63

Figure 14: Learning Curves For Internal Validation Of Random Forest Model Using South And West Census Regions For Chronic CVD Prediction 64

Figure 15: Heat Map Depicting Prediction Metrics For Acute CVD Events Models.....67

Figure 16: Heat Map Depicting Prediction Metrics For Chronic CVD Models 67

Figure 17: Web-Based Application To Predict The Risk Of Chronic CVD In Cancer Patients.....69

Figure 18: Web-Based Application To Predict The Risk Of Acute CVD In Cancer Patients.....70

Figure 19: PRISMA Flowchart Of Literature Review (Aim 2) 89

Figure 20: Effects Tree For The MCDA Model 101

Figure 21: Process Of Developing MCDA Model 108

Figure 22: Performance Of HER 2 Positive Therapy Regimen Alternatives On Each Criterion Of The MCDA Model Identified From The Literature 116

Figure 23: Radar Plots For Each HER 2 Positive Therapy Regimen Alternative Of The MCDA Model Using Literature Inputs And Scores Assigned 118

Figure 24: Criterion Preference Values Representing The Relative Importance Of Criteria In The MCDA Model 119

Figure 25: Criterion Value Functions Across The Therapy Regimen Alternatives 120

Figure 26: Criterion Preference Values Representing The Relative Importance Of Criteria In The MCDA Model 121

LIST OF ABBREVIATIONS

AHRQ	Agency for Healthcare Research and Quality
CVD	Cardiovascular diseases
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
DTC	Differentiated Thyroid Cancer
RR	Relative Risk
ADT	Androgen Deprivation Therapy
VEGFR-TKI	Vascular Endothelial Growth Factor Receptor Tyrosine Kinase Inhibitors
HR	Hazard Ratio
CT	Computed Tomography
ML	Machine Learning
MCDA	Multiple Criteria Decision Analysis
SDOH	Social Determinants of Health
WHO	World Health Organization
ED	Emergency Department
CAC	Coronary Artery Calcifications
ICC	Intraclass Correlation Coefficient
ICU	Intensive Care Unit
EMR	Electronic Medical Records
CHD	Carcinoid Heart Disease
MEPS	Medical Expenditure Panel Survey
US	United States
CCS	Clinical Classification Software
ICD	International Classification of Diseases
BMI	Body Mass Index
USC	Usual Source Care

RF	Random Forest
VIP	Variable Importance Plot
SGD	Stochastic Gradient Descent
ROC AUC	Receiver Operating Characteristic Area Under the Curve
KNN	K-nearest Neighbor
UI	User Interface
ER	Emergency Room
FPL	Federal Poverty Line
OR	Odds Ratios
PDC	Proportion Drug Covered
GEE	Generalized Estimating Equations
HER2	Human Epidermal Growth Factor Receptor 2
HTA	Health Technology Assessment
PDA	Portfolio Decision Analysis
SDM	Shared Decision making
PFS	Progression Free Survival
FAERS	FDA Adverse Event Reporting System
MedDRA	Medical Dictionary for Regulatory Activities
PT	Preferred Terms
ROR	Reporting Odds Ratios
CI	Confidence Intervals
QOL	Quality of Life
QALY	Quality Adjusted Life Years
OS	Overall Survival
CM	Cardiovascular Mortality
CH	Cardiovascular Hospitalization
CE	Cardiovascular Events
D	Diarrhea

PN	Peripheral Neuropathy
FN	Febrile Neutropenia
SVM	Support Vector Machines

ABSTRACT

DEVELOPMENT AND VALIDATION OF PREDICTIVE RISK AND MULTIPLE CRITERIA DECISION ANALYSIS MODELS TO EVALUATE CARDIOVASCULAR OUTCOMES AMONG CANCER PATIENTS

By Purva N. Parab, B.Pharm, M.S.

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

Virginia Commonwealth University, 2020
Advisor: Pramit A. Nadpara, Ph.D.
Assistant Professor, Department of Pharmacotherapy & Outcomes Science

Objective: The objectives of our study were to characterize the study population with cancer and cardiovascular diseases (CVD) both as compared to those without and to build a predictive model using machine learning (ML) algorithms that can predict the risk of CVD in cancer patients. In addition, our objective was also to evaluate characteristics associated with cardiotoxic adverse events of breast cancer therapies and develop a multiple criteria decision analysis (MCDA) model to conduct benefit-risk assessment of breast cancer therapy regimens. **Methods:** We used Medical Expenditure Panel Survey (MEPS) and FDA Adverse Events Reporting System (FAERS) 2005-2015 files along with literature evidence for our study. We used MEPS database to train our predictive models using ML algorithms such as random forest (RF), gradient boosting and deep learning and compared these to standard regression models. Separate predictive models were built for chronic and acute CVD. We characterized the population with both cancer and CVD and those with cancer therapy associated cardiotoxic adverse events using multinomial logistic models . FAERS and literature evidence were also used to build the MCDA model to rank the breast cancer therapy regimens given the benefits and the risks involved in the treatment alternatives. **Results:** Our study sample consisted of 44,217 cancer patients

identified using MEPS 2005-2015 files out of which 12,339 (28.7%) patients were also diagnosed with CVD. Age, marital status, education and employment status were the sociodemographic characteristics that differed significantly across cancer patients with and without CVD. We observed that most of the ML models for chronic (RF c-statistic: 0.9872, gradient boosting c-statistic: 0.7608, deep learning c-statistic: 0.7662) and acute CVD (RF c-statistic: 0.9738, gradient boosting c-statistic: 0.7853, deep learning c-statistic: 0.8267) were more accurate than the standard regression models for chronic (standard regression model c-statistic: 0.7641, GLM net model c-statistic: 0.7349) and acute (standard regression model c-statistic: 0.7534, GLM net model c-statistic: 0.7853) CVD. We then used the most accurate RF model to build a web-based application that could predict CVD risk. We then identified 35,630,544 breast cancer patients using FAERS dataset. Our findings suggest that breast cancer patients receiving targeted therapies were more likely to be diagnosed with CVD as compared to those who were receiving conventional therapies (OR = 1.213, 95% CI = 1.180, 1.247). On conducting MCDA, we found that the breast cancer therapy regimen 3 with trastuzumab, cyclophosphamide/ carboplatin and a taxane (paclitaxel/ docetaxel) was the most preferred therapy alternative given the benefits and the risks associated with each of the alternatives. **Conclusion:** Our study thus evaluated the use of newer analytical techniques such as ML algorithms and MCDA to evaluate certain outcomes. Our study suggests that ML algorithms were more accurate in predicting CVD risk in cancer patients. In addition, our MCDA model suggested that the breast cancer therapy regimen with trastuzumab, cyclophosphamide/ carboplatin and a taxane was the most preferred alternative considering the survival and adverse events benefits and risks.

CHAPTER 1: INTRODUCTION

BACKGROUND:

Cancer is one of the most prevalent diseases in the country.[1] It has been listed as one of the priority conditions by the Agency for Healthcare Research and Quality (AHRQ). The mean survival rate of cancer patients is around 67%.[1] The financial burden of cancer survivors is high too. The AHRQ estimates that the direct medical costs for cancer in the US in 2014 were $87.7 billion. On an average, cancer patients pay around $3664 to $8115 out of pocket annually.[2] Cancer patients incur significant indirect costs as well. Considering lost productivity value, the human capital approach is estimated to increase from $115.8 billion in 2000 to $147.6 billion in 2020, a 27.5% increase due only to population growth and aging.[3] Thus, along with the direct costs, cancer patients have financial burden due to significant indirect costs and loss of productivity as well. Comorbidities associated with the cancer condition are one of the major reasons for increased financial burden and a reduced quality of life.[4]

Cardiovascular diseases have been identified as one of the most commonly associated comorbidities with cancer patients.[5] Cardiovascular diseases (CVD) and cancer are the two leading causes of death worldwide.[6] They further increase the burden on patients if CVD occurs as a comorbidity along with cancer. In cancer patients, the CVD mortality rate has increased by 20-30% in recent years, whereas the cancer mortality rate has decreased by 20-30%.[7] This makes CVD an even bigger concern in cancer patients. CVD and cancer share various similarities and possible interactions, including a number of similar risk factors like age, tobacco use, diet and lack of physical activity.[8] Other than the lifestyle risk factors, there are also certain cancer therapies that put patients at a higher risk of developing CVD. There have been certain cancer therapies such as 5-fluorouracil,

taxanes, cyclophosphamides, trastuzumab, tamoxifen, bevacizumab and certain anthracyclines that are proven to be very cardiotoxic.[9] In addition to the existing therapies, novel anticancer therapeutics associated with higher survival outcomes are also associated with a higher cardiotoxic potential making the cardiovascular implications of cancer therapies increasingly important.[10] Other than targeted therapies, radiotherapy also impacts cardiac health and leads to certain cardiotoxicities.[10,11] Building a predictive model including cancer therapies and cancer type along with lifestyle factors would make predicting CVD risk in cancer patients more accurate. With increasing number of predictors, the accuracy of standard regression model goes on decreasing due to added variance. As the variance increases the standard error increases as well further reducing the accuracy pf the regression model. Predictive models using machine learning (ML) algorithms can incorporate a very high number of predictors in the model and predict the risk of CVD better.[12,13] In addition, CVD has been identified as one of the leading causes of mortality among cancer patients.[14] There have been certain studies in the field of cardio-oncology looking at strategies for management of CVD risk for such cardiac events and certain clinical outcomes leading to a need for developing cardio-oncology guidelines.[15–19] Currently, these guidelines are continuously revised by the American College of Cardiology.[20] CVDs identified along with cancer are coronary artery disease, valvular heart disease, heart failure and arrhythmias. Identifying the risk of these CVD conditions in cancer patients would help in managing the condition more efficiently.[21] Due to the increasing reports of cardiotoxicity associated with the newer cancer drugs in the market, there has been a growing interest in ways to prevent the cardiovascular events and manage cardiovascular health of cancer patients. There are studies conducted in literature

looking at the risk of CVD in cancer patients and strategies to contain the risk. Lifestyle factors, certain drug combinations, evidence-based medicine, adherence to CVD medications and higher utilization of CVD screening services are some of the strategies suggested to reduce the risk of CVD among cancer patients.[15,16] Studies looking at cardiovascular outcomes in cancer patients are summarized below.

LITERATURE REVIEW:

Search strategy and literature summary evaluating the CVD risk associated with cancer patients using the conventional regression approach:

A literature review was conducted using PubMed/Medline to summarize the existing evidence with respect to cardiovascular outcomes among patients with cancer. Risk for cardiovascular diseases (CVD) among patients with cancer and the benefit-risk tradeoff between survival and cardiotoxic outcomes associated with cancer therapies were the outcomes that were assessed for in the literature. The following search strategy using a combination of MeSH terms was used: (((((((("Cardiovascular Diseases"[Mesh]) AND ("Cardiovascular Diseases/adverse effects"[Mesh] OR "Cardiovascular Diseases/complications"[Mesh] OR "Cardiovascular Diseases/drug therapy"[Mesh] OR "Cardiovascular Diseases/epidemiology"[Mesh] OR "Cardiovascular Diseases/mortality"[Mesh] OR "Cardiovascular Diseases/organization and administration"[Mesh] OR "Cardiovascular Diseases/prevention and control"[Mesh] OR "Cardiovascular Diseases/radiotherapy"[Mesh] OR "Cardiovascular Diseases/statistics and numerical data"[Mesh]))) AND "Neoplasms"[Majr]) AND humans) AND English)) AND observational studies) The titles and abstracts were then screened for their eligibility using the following inclusion/exclusion criteria.

We included studies that report cardiovascular risk/ mortality among patients with cancer or cancer survivors. We also included studies that looked at interventions to reduce the risk of cardiovascular conditions in patients with cancer and reported these outcomes. We only included studies that were conducted in humans and published in English. We excluded studies that focused on cancer outcomes rather than cardiovascular outcomes. We also excluded studies that were not looking at cardiovascular conditions as a cancer comorbidity. We excluded studies that only reported clinical/ physiological outcomes and not any other health outcomes. We excluded studies conducted in pediatric population and those that were only narrative reviews

The search strategy resulted into 527 studies. On applying the inclusion/exclusion criteria, 513 articles were excluded. The final literature review included 15 studies.[22-36] Figure 1 below includes a PRISMA flowchart of studies that were included in the literature review. All the studies in literature have been published post 2009 which coincides with the timeline of majority of targeted therapies booming in the market. The number of studies looking at cardiovascular conditions in patients with cancer has increased in the recent past.

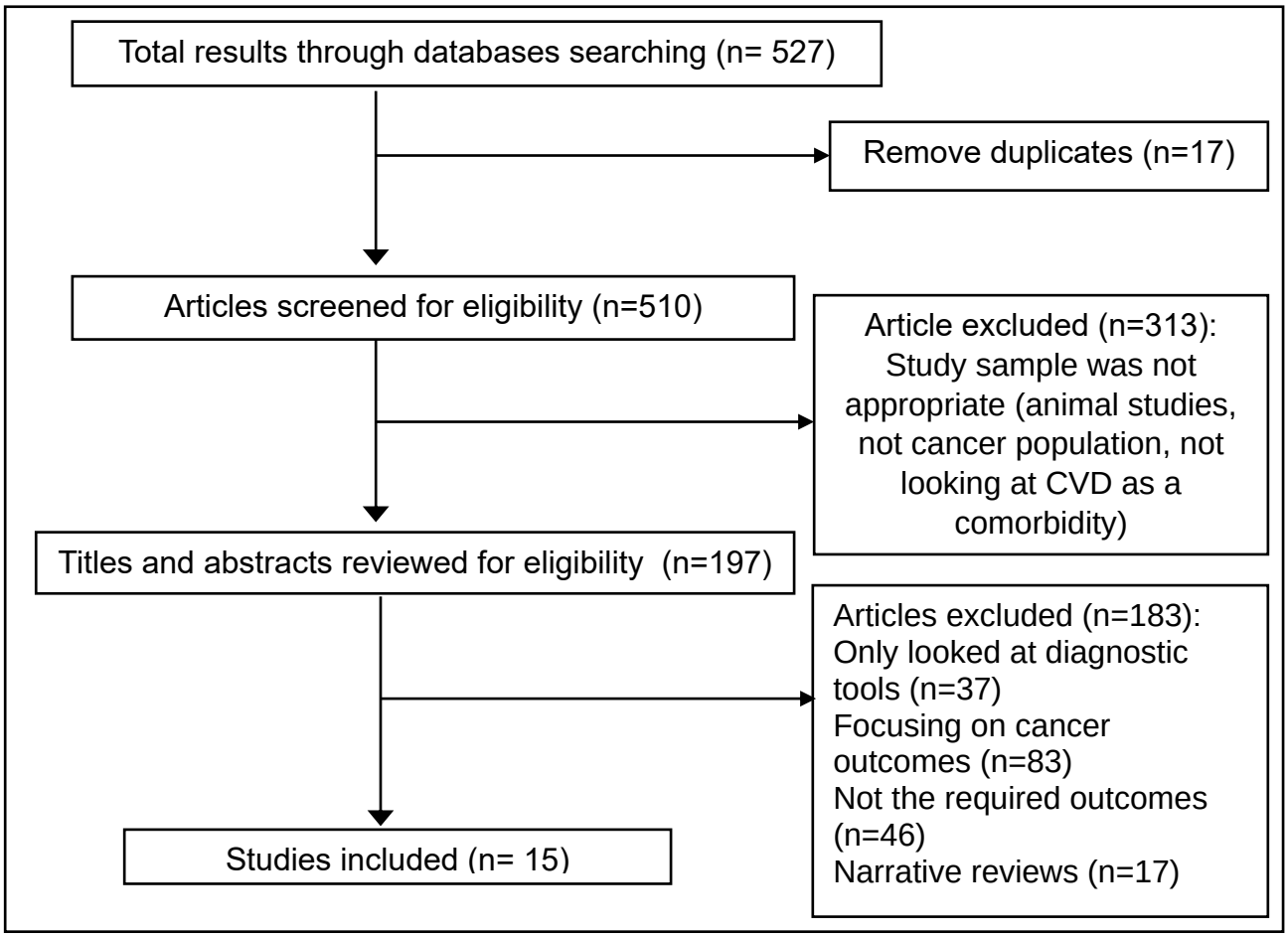


Figure 1: PRISMA Flowchart of Literature Review

Table 1 below summarizes the studies that were included in the literature review. There were 15 studies that were summarized for the literature review to identify the gaps in the literature. Majority of these studies have either looked at one specific type of cancer or have looked at cardiotoxicity associated with one specific cancer therapy.

Table 1: Summary of Literature Review

Study	Study Objective
Sturgeon *et al.,* 2019[22]	To characterize CVD mortality risk for multiple cancer sites, with respect to (i) continuous calendar year, (ii) age at diagnosis, and (iii) follow-up time after diagnosis
Winther *et al.,* 2018[23]	To examine the risk of CVD among cancer survivors with diabetes
Pajamaki et al., 2018[24]	To evaluate long-term cardiovascular morbidity and mortality in patients treated for differentiated thyroid cancer (DTC)
Berkman *et al.,* 2017[25]	To assess racial differences in 20-year cardiovascular mortality risk among cancer survivors
Grazziotin et al., 2017[26]	To measure the incidence of trastuzumab-related cardiotoxicity in patients with breast cancer
Khosrow-Khavar et al., 2017[27]	To determine the association of aromatase inhibitors with the increased risk of cardiovascular events
Santoni et al., 2017[28]	To evaluate the incidence and relative risk (RR) of developing all-grade and high-grade cardiotoxicity in cancer patients receiving targeted agents
Armenian *et al*., 2016[29]	To examine the impact of cardiovascular risk factors in survivors of adult-onset cancer
Bhakta *et al.,* 2016[30]	To estimate the cumulative burden of cardiovascular mortality in survivors of Hodgkin's lymphoma
Okoye et al., 2016[31]	To quantify CVD risk and receipt of primary preventive care among patients with head and neck squamous cell carcinoma treated with radiotherapy or chemoradiotherapy
O'Farrell *et al.*, 2015[32]	To examine the association between risk of CVD and the duration and type of androgen-deprivation therapy in men with advanced prostate cancer
Hesselink *et al.,* 2013[33]	To study the risk of CVD mortality in patients with advanced thyroid carcinoma
Mulrooney *et al.,* 2009[34]	To assess the incidences and risks of cardiac outcomes in cancer survivors
Efstathiou et al., 2009[35]	To assess the relationship between cardiovascular mortality and androgen deprivation therapy for locally advanced prostate cancer patients
Tsai et al., 2007[36]	To investigate the association between androgen deprivation therapy (ADT) and the risk of cardiovascular mortality

Studies conducted by Grazziotin et al.,[26] Khosrow-Khavar et al.,[27] Santoni et al.,[28] Okoye et al.,[31] O'Farrell et al.,[32] Mulrooney et al.,[34] Efstathiou et al.[35] and Tsai et al.[36] mainly looked at one specific cancer treatment related cardiotoxicity or cardiac events. Pajamaki et al.,[24] Bhakta et al.[30] and Hesselink et al.[33] looked at a small sample size of one specific cancer type. There were only four studies conducted by Sturgeon et al.,[22] Winther et al.,[23] Berkman et al.,[25] and Armenian et al.[29] that looked at the overall cardiovascular risk/ mortality risk in cancer survivors.

Grazziotin et al., suggested that the incidence of trastuzumab-related cardiotoxicity in patients with metastatic breast cancer patients was 75% whereas in early stage breast cancer patients was 45.7%. This study was however restricted to trastuzumab-related cardiotoxicity in cancer patients.[26]

Khosrow-Khavar et al. conducted a meta-analysis looking at the cardiotoxicity associated with aromatase inhibitors and tamoxifen in postmenopausal women with breast cancer. The study found that there was a 19% increased risk of cardiovascular events associated with aromatase inhibitors as compared to those using Tamoxifen. The results were also suggestive of a small protective effect of Tamoxifen.[27]

Santoni et al., conducted a meta-analysis of available clinical trials looking at the incidence and relative risks of developing cardiotoxicities in cancer patients receiving targeted therapies. The highest relative risk (RR) of high-grade cardiac events was observed in Vandetinib (RR = 7.71). Grouping by drug category, highest risk of cardiotoxicity was associated with anti-VEGFR-TKIs (RR = 5.62). Targeted agents were correlated with a significant increase in the risk of cardiotoxicity. The study was restricted to patients receiving targeted therapies and cardiotoxicities associated with these.[28]

Okoye et al. looked at quantifying risk factors of CVD among head and neck squamous cell carcinoma patients receiving radiotherapy or chemoradiotherapy. Hypertension, smoking and diabetes were the comorbidities that were identified as the most prevalent CVD risk factors. This study however only focused on a very small sample of head and neck cancers receiving radiotherapy and focused majorly on the risk factors of CVD.[31]

O'Farrell et al. looked at the association of androgen deprivation therapy (ADT) with the risk and timing of CVD in men with prostate cancer. CVD risk was highest during the first 6 months of ADT in men who experienced two or more cardiovascular events before therapy with a Hazard Ratio (HR) of 1.91.[32]

Mulrooney et al. looked at the incidence rates of cardiac outcomes among adult cancer survivors. Higher hazard ratios (HR) were associated with congestive heart failure (HR = 5.9), myocardial infarction (HR = 5.0), pericardial disease (HR = 6.3) and valvular abnormalities (HR = 4.8) in cancer survivors as compared to those without cancer. The study also focused mainly on the association of cardiac events with the exposure to cancer therapies such as dose of anthracyclines and the extent of radiation exposure. Higher doses and exposures were associated with worse cardiac events. This study was however restricted to those who had a cancer diagnosis under the age of 21years.[34]

Efstathiou et al. and Tsai et al. looked at cardiovascular mortality after ADT. The prior study was conducted in patients with localized cancer whereas the later was in patients with locally advanced cancer. In the study conducted by Efstathiou et al. the use of ADT was significantly associated with increased risk of cardiovascular mortality (HR =

2.6). In the later study however, there was no significant increase in treatment-related cardiovascular mortality.[35,36]

All of these studies have majorly only focused on cancer therapy related cardiotoxicities. There have been three studies focusing only on Hodgkin lymphoma and thyroid cancer. Bhakta et al. conducted a study assessing the cumulative burden of cardiovascular morbidity in Hodgkin lymphoma survivors. The study suggested that at 50years of age, the cumulative incidence of survivors experiencing at least one grade 3-5 (more severe) cardiovascular condition was 45.5%. Myocardial infarction was one of the major contributors to the excess cardiovascular burden.[30]

Pajamaki et al. and Hesselink et al. looked at cardiovascular mortality and morbidity in thyroid cancer patients. Both the studies reported similar findings, where patients with thyroid cancer reported a higher cardiovascular morbidity as compared to those without cancer. Pajamaki et al. reported a higher HR of 1.16 associated with morbidity due to any CVD event, whereas a lower HR of 0.73 associated with cardiovascular mortality in thyroid cancer patients.[24] Hesselink et al. reported increased risk of cardiovascular mortality (HR = 3.35) among patients with thyroid cancer. The results with respect to cardiovascular mortality in specific cancers are thus inconclusive.[33]

Sturgeon et al, suggested in their study that CVD mortality risk was highest in survivors diagnosed at <35 years of age with the risk being highest within first year after cancer diagnosis (standardized mortality ratio = 3.93, 95% confidence intervals = 3.89-3.97). This risk remained elevated throughout follow-up compared to general population. This study however, only focused on CVD mortality rather than CVD diagnosis risk. In

addition, the only factors studied to evaluate the association were age at diagnosis and follow-up after diagnosis.[22]

Winther et al., suggested that compared to patients without cancer, those with cancer are 3.6 times more likely to develop a cardiovascular event. This hazard of cardiovascular event is even higher (HR = 8.7) in patients with both cancer and diabetes as compared to those without cancer and diabetes. Comorbidities thus add to the burden of CVD in patients with cancer.[23]

Berkman et al. looked at racial differences in 20-year cardiovascular mortality risk in cancer survivors. Black survivors had higher risks for CVD mortality (HR = 2.13) compared to white survivors. The increased risk of CVD persisted at 5-years (HR = 2.38) and 20-years (HR = 2.31).[25]

Armenian et al. examined the impact of cardiovascular risk factors (hypertension, diabetes, dyslipidemia) on long-term CVD risk in cancer survivors. Cancer survivors with two or more risk factors had the highest risk (incidence rate ratio = 1.83) of CVD when compared to noncancer controls. The magnitude of CVD risk varied depending on the number of cardiovascular risk factors present. This was the most relevant article for our research since it looked at multiple cancers and examined the impact of multiple factors on the risk of CVD in cancer patients.[29]

Gaps in the literature:

There are several studies in the literature that have looked at the risk of cardiovascular diseases among cancer patients. Majority of these studies have focused only on cardiotoxicities associated with specific cancer therapies or have been restricted

to a single type of cancer. The studies that looked at multiple cancers and CVD risk factors have also focused majorly on one of the risk factors, such as comorbidities, racial background or a finite number of CVD risk factors. Thus, it has been established in the literature that independently there are multiple risk factors including lifestyle, sociodemographic and certain therapy related characteristics that put cancer patients at a higher risk of CVD as compared to general population. There has not been an attempt to build a predictive risk model which could predict the risk of CVD in cancer patients by incorporating all of these factors and their interactions. There is not a clear understanding of CVD risk in cancer population on considering all the risk factors and their interactions. These predictive models could help in assessing the CVD risk in cancer population more accurately than the standard regression techniques. CVD is becoming a major concern among patients with cancer due to high CVD mortality and morbidity rates. A machine learning (ML) model in this case would be very beneficial. In addition, there has not been an attempt to build a benefit-risk model to account for survival and cardiotoxic outcomes of cancer therapies. Accounting for these together in a model would help in assigning a value to each therapy and guide the decision-making process.

RATIONALE:

As mentioned earlier, there is not a clear understanding of CVD risk in cancer population on a more generalizable scale. There is still a lack of a predictive risk model which could incorporate all of these risk factors in a single model and predict the future risk of CVD. It has been established in the literature that there are multiple shared risk factors including lifestyle, sociodemographic and certain therapy related characteristics that put cancer patients at a higher risk of CVD as compared to general population.

Machine learning algorithms have proven to be more accurate and precise in situations where the number of predictors and interactions within these are large. There are only two studies that looked at predicting clinical deterioration (cardiac arrest) or cardiotoxicity in cancer patients using the machine learning approach.[37,38] These have however looked at a very small sample of a specific cancer type and have not incorporated appropriate algorithms that could handle interactions within predictors. These have also looked at a specific CVD outcome as opposed to the overall CVD risk in cancer patients. CVD is becoming a major concern among patients with cancer due to high CVD mortality and morbidity rates. Building a predictive risk model by considering majority of predictors would help in quantifying the probability of getting diagnosed with CVD better. An understanding of the underlying probability of CVD risk beforehand is necessary to tailor interventions and manage patients more efficiently. However, the studies so far have used a simpler approach to estimate the risk of CVD in cancer patients. As seen from the above literature, there is also a lack of evidence of studies that incorporate the benefits and risks of cancer therapies together in a model to assign values and make the decision process more transparent. Targeted therapies as suggested are associated with a higher survival as well as a higher cardiotoxic potential which makes choosing the therapy alternative difficult. Most of these targeted therapies associated with a higher survival and cardiotoxic profile are used in breast cancer patients. Most of the clinical studies conducted looking at these outcomes have also been in breast cancer patients. We would thus assume that breast cancer therapies would have the highest benefit-risk trade off and restrict the Multiple Criteria Decision Analysis (MCDA) model to have breast cancer therapies as alternatives. Current studies have only looked at patient preferences for treatment outcomes, there are

no studies in the literature that look at value assessment of cancer therapies from a benefit-risk perspective. Developing a MCDA model would help in assigning a value to these breast cancer therapies by quantifying the risks and the benefits and making the decision process more transparent and easier.

SPECIFIC AIMS:

1. To develop and validate predictive risk models to assess the risk of cardiovascular diseases (CVD) among cancer patients using machine learning algorithms
    a. To compare sociodemographic characteristics of cancer patients with CVD to those without CVD
    b. To build predictive risk models using random forest, gradient boosting and deep learning algorithms and compare the accuracy to standard regression models
    c. To validate the predictive risk models using cross validation techniques and evaluate the model fit on a varied sample
    d. To create an interactive web-based application using R-shiny to predict the risk of CVD among cancer patients using the most accurate model identified
2. To assess cardiotoxicity associated with targeted therapies as compared to non-targeted therapies and develop a model to conduct benefit-risk assessment of therapy regimens in breast cancer patients
    a. To describe the cancer therapy characteristics associated with the cardiotoxic adverse events in breast cancer patients receiving targeted therapies as compared to those receiving non-targeted therapies and evaluate the drug-event association using a disproportionality analysis

b. To develop a Multi Criteria Decision Analysis (MCDA) model to conduct benefit-risk assessment of breast cancer therapy regimens

c. To conduct sensitivity analyses to assess the MCDA model performance and uncertainty in the model

REFERENCES:

1. Surveillance, Epidemiology, and End Results Program - Cancer Incidence.

2. American Cancer Society - Economic Impact of Cancer. https://www.cancer.org/cancer/cancer-basics/economic-impact-of-cancer.html%0A%0A. Accessed March 25, 2019.

3. Yabroff KR, Lund J, Kepka D, Mariotto A. Economic burden of cancer in the United States: estimates, projections, and future research. *Cancer Epidemiol Biomarkers Prev*. 2011;20(10):2006-2014. doi:10.1158/1055-9965.EPI-11-0650

4. Jones SMW, Chennupati S, Nguyen T, Fedorenko C, Ramsey SD. Comorbidity is associated with higher risk of fi nancial burden in Medicare bene fi ciaries with cancer but not heart disease or diabetes.

5. Liu D, Ma Z, Yang J, et al. Prevalence and prognosis significance of cardiovascular disease in cancer patients : a population-based study. 2019;11(18):7948-7960.

6. MedicalXpress: Cardiology and Oncology.

7. Herrmann J. From trends to transformation: where cardio-oncology is to make a difference. *Eur Heart J*. 2019;40(48):3898-3900. doi:10.1093/eurheartj/ehz781

8. Koene RJ, Prizment AE, Blaes A, Konety SH. Shared Risk Factors in Cardiovascular Disease and Cancer. *Circulation*. 2016;133(11):1104-1114. doi:10.1161/CIRCULATIONAHA.115.020406

9. Florescu M, Cinteza M, Dragos V. Chemotherapy-induced Cardiotoxicity. *Maedica - a J Clin Med*. 2013;8(1):59-67.

10. Gm W, Ar L, Mt H, Gg H. Cardiotoxicity Following Cancer Treatment. 2017;86(June 2016):3-9.

11. Mozdzanowska D, Woźniewski M. Radiotherapy and anthracyclines – cardiovascular toxicity. doi:10.5114/wo.2014.40108

12. Royer J, Wu E AR. Use of machine learning in health economics and outcomes research (HEOR)- Optimizing predictive models. In: ; 2019.

13. Rahimian F, Salimi-khorshidi G, Payberah AH, et al. Predicting the risk of emergency admission with machine learning : Development and validation using linked electronic health records. 2018;779:1-18.

14. Almuwaqqat Zakaria et. al. Statins and cancer survivors: the need for structured guidelines.

15. Abe J, Sood AK, Martin JF. Editorial : Cardio-Oncology : From Bench to Bedside. 2019;6(April):1-2. doi:10.3389/fcvm.2019.00037

16. Totzeck M, Schuler M, Stuschke M, Heusch G, Rassaf T. Cardio-oncology - strategies for management of cancer-therapy related cardiovascular disease. *Int J Cardiol*. 2019;280:163-175. doi:10.1016/j.ijcard.2019.01.038

17. Hussain Y, Drill E, Dang CT, Liu JE, Steingart RM, Yu AF. Cardiac outcomes of trastuzumab therapy in patients with HER2-positive breast cancer and reduced left ventricular ejection fraction. *Breast Cancer Res Treat*. February 2019. doi:10.1007/s10549-019-05139-6

18. Narayan V, Ky B. Common Cardiovascular Complications of Cancer Therapy: Epidemiology, Risk Prediction, and Prevention. *Annu Rev Med*. 2018;69:97-111. doi:10.1146/annurev-med-041316-090622

19. Mehta LS, Watson KE, Barac A, et al. Cardiovascular Disease and Breast Cancer: Where These Entities Intersect: A Scientific Statement From the American Heart Association. *Circulation*. 2018;137(8):e30-e66. doi:10.1161/CIR.0000000000000556

20. Latest Advances in CardioOncology. American College of Cardiology https://www.acc.org/membership/sections-and-councils/cardio-oncology/section-updates/2019/12/17/24/42/keep-up-on-the-latest-advances-in-cardio-oncology-with-jacc-cardiooncology accessed on: 08/25/2020

21. Herrmann J, Lerman A, Sandhu NP, Villarraga HR, Mulvagh SL, Kohli M. Evaluation and management of patients with heart disease and cancer: cardio-oncology. *Mayo Clin Proc*. 2014;89(9):1287-1306. doi:10.1016/j.mayocp.2014.05.013

22. Sturgeon KM, Deng L, Bluethmann SM, et al. A population-based study of cardiovascular disease mortality risk in US cancer patients. *Eur Heart J*. 2019;40(48):3889-3897. doi:10.1093/eurheartj/ehz766

23. Winther JF, Bhatia S, Cederkvist L, Gudmundsdottir T. Risk of Cardiovascular Disease Among Nordic Childhood Cancer Survivors With Diabetes Mellitus : A Report From Adult Life After Childhood Cancer in Scandinavia. 2018:4393-4400. doi:10.1002/cncr.31696

24. Pajamäki N, Metso S, Hakala T, et al. term cardiovascular morbidity and mortality in patients treated for differentiated thyroid cancer. 2018;(November 2017):303-310. doi:10.1111/cen.13519

25. Berkman AM, Brewster AM, Jones LW, et al. Among Childhood and Young Adult Cancer Survivors. 2017;6(3):414-421. doi:10.1089/jayao.2017.0024

26. Grazziotin L, Picon P. Observational study of trastuzumab-related cardiotoxicity in early and metastatic breast cancer. *Journal of Oncology Pharmacy Practice*. 2017;23(4):264-272. doi:10.1177/1078155216639755

27. Khosrow-Khavar F, Filion KB, Bouganim N, Suissa S, Azoulay L. Aromatase inhibitors and the risk of cardiovascular outcomes in women with breast cancer: A population-based cohort study. *Circulation*. 2017:549-559. doi:10.1161/CIRCULATIONAHA.119.044750

28. Santoni M, Guerra F, Conti A, et al. Incidence and risk of cardiotoxicity in cancer patients treated with targeted therapies. *Cancer Treat Rev*. 2017;59:123-131. doi:10.1016/j.ctrv.2017.07.006

29. Armenian SH, Xu L, Ky B, et al. Cardiovascular Disease Among Survivors of Adult-Onset Cancer: A Community-Based Retrospective Cohort Study. *J Clin Oncol*. 2016;34(10):1122-1130. doi:10.1200/JCO.2015.64.0409

30. Bhakta N, Liu Q, Yeo F, et al. Cumulative burden of cardiovascular morbidity in paediatric , adolescent , and young adult survivors of Hodgkin ' s lymphoma : an analysis from the St Jude Lifetime Cohort Study. *Lancet Oncol*. 2016;17(9):1325-1334. doi:10.1016/S1470-2045(16)30215-7

31. Okoye CC, Bucher J, Tatsuoka C, et al. Patients Treated with Radiotherapy. 2016;39(3):527-532. doi:10.1002/hed.24646.*Cardiovascular*

32. Farrell SO, Garmo H, Holmberg L, et al. Risk and Timing of Cardiovascular Disease After Androgen- Deprivation Therapy in Men With Prostate Cancer. *Journal Of Clinical Oncology* 2015;33(11). doi:10.1200/JCO.2014.59.1792

33. Hesselink ENK, Kamphuisen PW, Plukker JTM, Links TP, Lefrandt JD. Long-Term Cardiovascular Mortality in Patients With Differentiated Thyroid Carcinoma : An Observational Study. *Journal Of Clinical Oncology* 2013;31(32). doi:10.1200/JCO.2013.49.1043

34. Mulrooney DA, Yeazel MW, Kawashima T, et al. Cardiac Outcomes in a Cohort of Adult Survivors of Childhood and Adolescent Cancer : Retrospective Analysis of the Childhood Cancer Survivor Study Cohort. *BMJ*. 2009 :1-11. doi:10.1136/bmj.b4606

35. Efstathiou JA, Bae K, Shipley WU, et al. J OURNAL OF C LINICAL O NCOLOGY Cardiovascular Mortality After Androgen Deprivation Therapy for Locally Advanced Prostate Cancer : RTOG 85-31. 2020;27(1):11-14. doi:10.1200/JCO.2009.12.3752

36. Tsai HK, Amico AVD, Sadetsky N, Chen M, Carroll PR. Androgen Deprivation Therapy for Localized Prostate Cancer and the Risk of Cardiovascular Mortality.

2007;99(20). doi:10.1093/jnci/djm168

37. Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model. 2016:1-12. doi:10.1371/journal.pone.0161401

38. Dranitsaris G, Rayson ÆD, Vincent ÆM, et al. The development of a predictive model to estimate cardiotoxic risk for patients with metastatic breast cancer receiving anthracyclines. 2008:443-450. doi:10.1007/s10549-007-9803-5

CHAPTER 2: PREDICTIVE RISK MODELING FOR PREDICTION OF CVD RISK AMONG CANCER PATIENTS

BACKGROUND:

The increasing number of risk factors associated with CVD in cancer patients and the interactions within these have made it essential to build a predictive risk model which can incorporate all the predictors. Literature suggests that when the number of predictors is large along with interactions within these, the standard regression models are not accurate. A machine learning (ML) approach would build more accurate predictive models.[1,2] A predictive risk model would help in guiding decision-making by presenting all the factors that put cancer patients at a higher risk of developing CVD. The growing availability of nationally representative datasets, together with advances in ML offer new opportunities for development of novel risk prediction models that are better at predicting risk. Such models have been shown to outperform standard statistical models particularly when the number of predictors is high and relationships more complex.[3] A simple regression model predicts current risk based on the available data where as a predictive model built using a ML approach could predict the future risk based on the available data. A study conducted by Rahimian et al., looked at the prediction of risk models for ED visits using the ML approach and compared the model to the standard hazard models.[2] The authors concluded that, ML approach produced more robust findings for a longer time horizon. Literature suggests that in general population, ML significantly improves the accuracy of cardiovascular risk prediction.[4] Compared to the established risk prediction algorithm, ML algorithms such as random forest, gradient boosting and neural networks improved prediction. The superiority of ML has thus been established in predicting the risk of CVD in general population.[4] However, none of the studies conducted so far have used

the ML approach to build a risk model for CVD amongst cancer patients. As stated earlier, cancer patients are at an even higher risk of developing CVD and the risk factors differ significantly than the general population. The predictive risk model would differ than that obtained in the general population. Developing a predictive risk model specifically in cancer patients would thus be beneficial in improving the accuracy of CVD risk prediction. It would eventually increase the number of patients identified who could benefit from preventive treatment and avoid unnecessary cardiac complications. The results from our study might also guide physicians in decision-making and building cardio-oncology guidelines. Characterizing the patients that have cancer and CVD diagnoses both and incorporating for interactions within the predictive model would help in tailoring CVD interventions by managing risk efficiently in cancer patients. Developing a dynamic web-based application based on the most accurate ML algorithm that would predict the real time risk of CVD in cancer patients would also help the physicians planning the treatment regimen and prognosis of their cancer patients.

Literature Review:

The following search strategy using a combination of MeSH terms was used: (((((("Machine Learning"[MeSH] OR "Neural Networks (Computer)"[Mesh] OR "Algorithms"[Mesh:noexp]))) OR ("Machine learning" OR "Neural networks" OR "Neural network" OR "Network model" OR "Network models" OR "Deep learning" OR "Random Forest" OR "Gradient Boosting" OR "Algorithm" OR "Algorithms")))) AND ((("Angina" OR "Myocardial Infarction" OR "Coronary Artery Disease" OR "Congestive Heart Failure")) OR "Heart Diseases"[Mesh])) AND "Neoplasms"[Mesh]. The titles and abstracts were then screened for their eligibility using the following inclusion/exclusion criteria.

We included studies that looked at application of ML in predicting cardiovascular outcomes in cancer patients. We restricted our studies to those conducted in humans and published in English. We excluded studies that focused on cancer outcomes rather than cardiovascular outcomes or that did not look at CVD as a comorbidity. We also excluded studies that just looked at ML algorithms as a diagnostic/ imaging tool for malignant tumors. We excluded studies that looked at pediatric population or were narrative reviews not looking at any particular outcome.

The search strategy resulted into 228 studies. On applying the inclusion/exclusion criteria, 222 articles were excluded. The final literature review included six studies.[5 – 10] Figure 2 below includes a PRISMA flowchart of studies that were included in the literature review.

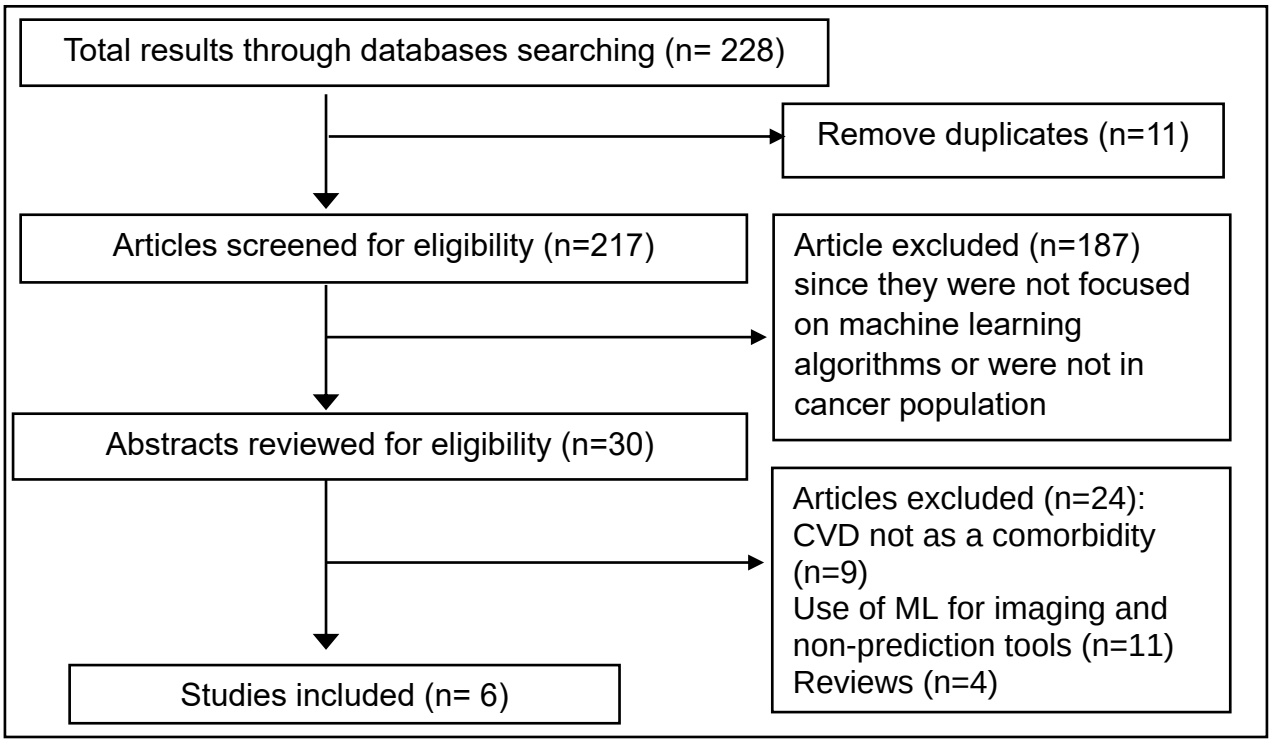


Figure 2: PRISMA Flowchart of Literature Review (Aim 1)

Table 2 below summarizes the studies that were included in the literature review. There were 6 studies that were summarized for the literature review to identify the gaps in the literature. Majority of these studies have looked at one specific type of cancer or have been diagnostic tools for cardiovascular outcomes in cancer patients. The number of studies looking at cardiovascular conditions in patients with cancer using ML algorithms has increased in the recent past due to the increasing popularity of the method.

Table 2: Summary of Literature Review (Aim 1)

Study	**Study Objective**
Gernaat *et al.,* 2018[5]	To apply a new deep learning algorithm for automated quantification of coronary artery calcifications among breast cancer patients
Lessmann et al., 2018[6]	To apply convolutional neural network method to automatically detect coronary artery, aorta and cardiac valve calcifications among lung cancer patients
Hu et al., 2016[7]	To develop a prediction model using a neural network to predict clinical deterioration (cardiac arrest) in adult hematologic malignant patients
Takx *et al.,* 2014[8]	To determine the reliability of automated coronary artery calcification scoring in lung cancer screening population
Dranitsaris et al., 2008[9]	To develop a predictive model to estimate cardiotoxic risk for patients with breast cancer receiving anthracyclines
Van Gerven et al., 2007[10]	To predict the development of carcinoid heart disease in neuroendocrine malignant tumor patients using noisy-threshold classifier

There have been very few studies in the literature that have used ML algorithms to predict cardiovascular outcomes in cancer patients. Most of the studies that have used ML algorithms for prediction have been in general population. In cancer patients, the use of ML so far has been extensively as an imaging/ diagnostic tool to evaluate malignant

tumors. However, there are 6 studies that have been summarized in table 2 that evaluated a ML tool to predict cardiovascular outcomes in cancer patients.

There were two studies conducted by Gernaat et al. and Lessmann et al. that looked at automatic calcification scoring in cancer patients. Both these studies used a similar deep learning approach to quantify coronary artery calcifications (CAC) on radiotherapy computed tomography (CT) scans. The CT scans were evaluated to develop a deep learning algorithm using two convolutional neural networks. This algorithm was trained to automatically score the calcifications in the CT scans.[5,6] The study conducted by Gernaat et al. focused on breast cancer patients, the algorithm was compared against manual scoring and it was concluded that the automatic scoring tool showed high reproducibility (proportion of agreement = 0.90, Intraclass correlation coefficient [ICC] = 0.99) and was quicker than the conventional approaches to quantify calcifications.[5] Lessmann et al., looked at lung cancer patients and the algorithm enabled reliable automatic quantification of calcification (sensitivity = 92%) for lung cancer patients.[6]

Hu et al. developed a prediction tool using a neural network model that would increase the predictive accuracy of detecting clinical deterioration (ICU transfer and cardiac arrest) among hospitalized patients with hematologic malignancies. EMR records of patients were used to build the model. The algorithm was trained on 565 (50%) hospitalized patients and cross-validation was performed on a separate 25% of the sample. Overall sensitivity of the neural network was found to be 84% whereas specificity was found to be 98%. The model correctly identified 7.6% of admissions to be clinically deteriorated whereas 92.4% did not result in clinical deterioration.[7]

Another study conducted by Takx et al. determined the agreement and reliability of such fully automated coronary artery calcium scoring in lung cancer screening population. There were 1749 CT scans that were analyzed to estimate the reliability as compared to manual scoring. Fully automated coronary calcium scoring in a lung cancer screening setting was found to be feasible with acceptable reliability and agreement (ICC = 0.90).[8]

A study conducted by Dranitsaris et al. to develop a predictive model was the most relevant to our study aim. However, this study was restricted to breast cancer patients receiving anthracyclines and focused on treatment induced cardiotoxicity. A model was built based on generalized estimating equations logistic regression using the ML approach. This model was trained on breast cancer patients from a randomized controlled trial. The initial list of 20 predictors was reduced to six to be included in the final predictive model which estimated that prior anthracycline exposure was associated with a higher risk for cardiotoxicity on controlling for other factors.[9]

Van Gerven at al. looked at predicting the development of carcinoid heart disease (CHD) among neuroendocrine tumor patients. There were 54 cases of patients enrolled from the Netherlands Cancer Institute that were analyzed to develop the model using noisy-threshold classifier as compared to naïve-Bayes classifier, logistic regression and decision-tree learning algorithm. The noisy-threshold classifier showed the best classification accuracy of 72% correctly classified cases.[10] Other studies in the literature using ML algorithm have been on a noncancer population.

Gaps in the literature:

There is a lack of a predictive risk model which could incorporate all these risk factors in a single model and predict the future risk of CVD. As mentioned in the literature summary above, majority of these machine learning algorithms have been developed from a diagnostic perspective of quantifying a tumor or scoring these masses based on the CT scans. There are only two studies that looked at predicting clinical deterioration (cardiac arrest) or cardiotoxicity in cancer patients. These have however looked at a very small sample of a specific cancer type and have not incorporated appropriate algorithms that could handle interactions within predictors. These have also looked at a specific CVD outcome as opposed to the overall CVD risk in cancer patients. The aim of our study was thus to build a predictive risk model using machine learning algorithms that could incorporate a large number of predictors and interactions within these. This model would then be able to predict future CVD risk among cancer patients.

Specific Aim 1:

To develop and validate predictive risk models to assess the risk of cardiovascular diseases (CVD) among cancer patients using machine learning algorithms

a. To compare the sociodemographic characteristics of cancer patients with CVD to those without CVD
b. To build predictive risk models using random forest, gradient boosting and deep learning algorithms and compare the accuracy to standard regression models
c. To validate the predictive risk models using cross validation techniques and evaluate the model fit on a varied sample
d. To create an interactive web-based application using R-shiny to predict the risk of CVD among cancer patients using the most accurate model identified

CONCEPTUAL FRAMEWORK:

Our study aim 1 was based on the social determinants of health (SDOH) conceptual framework by the World Health Organization (WHO).[11] This framework demonstrates how social, economic, and political factors influence a person's socioeconomic position which, in turn, plays a role in determining health outcomes. Our study used these socioeconomic indicators along with certain cancer related and overall health related characteristics. We used age, gender, race, marital status, employment status, census region, income, education level, health insurance coverage and access to care characteristics as the socioeconomic determinants of health. Patient-related lifestyle factors, cancer related characteristics, overall related health factors, CVD screening services and medication adherence were the intermediary behavioral and biological determinants of health. We identified all the predictors that we used in the model based on literature evidence and social determinants of health conceptual framework.

METHODS AIM 1: TO DEVELOP AND VALIDATE PREDICTIVE RISK MODELS TO ASSESS THE RISK OF CARDIOVASCULAR DISEASES (CVD) AMONG CANCER PATIENTS USING THE MACHINE LEARNING ALGORITHMS

Data Source:

Medical Expenditure Panel Survey (MEPS) 2005-2015 files were used for the study. MEPS is a nationally representative survey that collects data on the health services that Americans use, frequency of utilization, costs and sources of payments associated with these, as well as data on the cost, scope, and breadth of health insurance held by and available to U.S. workers.[12] The dataset was used to obtain complete information needed to identify risk of CVD amongst cancer patients and build a predictive risk model using machine-learning algorithms. The Full Year Consolidated, Medical Conditions, Prescription and Emergency/Inpatient/Outpatient/Office-based visits files from 2005-2015 were used from MEPS to input data into the model.

Proposed study design and sample:

Cross sectional study design was used for this study aim. The study sample consisted of patients diagnosed with cancer over the age of 18 years in the US from 2005-2015.The Medical Conditions files of the Household Component were used to obtain the clinical classification codes to identify cancer patients. Clinical Classification Software (CCS) collapses categories based on ICD-9 codes and generates more meaningful codes which can be used to look at broader categories like 'cancer' and not a specific type of cancer. Hence, clinical classification codes of 11-44 were used to identify cancer patients. The study sample was then restricted to adult respondents who were diagnosed with or had cancer after the age of 18 years as identified by the CCS. Adults who died during the process of reporting were excluded. There were 48,829 patients identified with cancer

and complete information on their CVD status. Out of these, 4,612 patients were excluded since their CVD diagnosis was before their cancer diagnosis. Those with a chronic CVD diagnosis within 1 year of cancer diagnosis were also excluded since in this case it was possible that the CVD condition was unrelated to the cancer condition. The final study sample thus consisted of 44,217 cancer patients, out of which 12,339 (weighted percentage = 28.7%) patients had a chronic CVD diagnosis whereas 31,878 patients (weighted percentage = 71.3%) did not have CVD. Majority of these patients with cancer and CVD had at least a 2 years' time period between their cancer and CVD diagnoses extending up to 9-12 years. Dates of diagnosis of cancer and CVD were used to estimate the mean time between cancer and CVD diagnoses. A categorical time variable was created to group patients into categories of time intervals between cancer and CVD such as 1-3 years, 3-5 years, 5-10years and > 10years. Table 3 summarizes the sample distribution based on the time interval between cancer and chronic CVD diagnoses among those who were diagnosed with cancer and CVD both. In addition to chronic CVD, there were 3,837 inpatient CVD procedures performed, 7,069 hospital stays due to CVD condition and 3,423 emergency visits due to CVD identified after the cancer diagnosis which were sub grouped as acute CVD events. These chronic and acute CVD events were analyzed separately.

Table 3: Time Interval Between Cancer and Chronic CVD Diagnoses (N = 12,339)

Time Interval	Frequency N(%)
2 - 3 years	3,265 (26.46)
3 -5 years	1,384 (11.22)
5 – 10 years	2,033 16.48)
>10 years	3,916 (31.74)
Missing	1,741 (14.1)

CVD diagnoses:

CVD diagnoses was identified using MIDX, ANGIDX, CHDDX AND OTHRDX from Full Year Consolidated files for myocardial infarction, angina, coronary heart disease and other heart diseases respectively for chronic CVD conditions. Acute CVD events were identified using the procedural or clinical classification codes from Inpatient, Outpatient and Emergency Room visits files. The dates of diagnoses of these were used to identify if the CVD was diagnosed before or after cancer and exclude those who had an acute CVD diagnosis before cancer. All these chronic and acute CVD events were analyzed at a patient level by transposing the diagnoses per patient. The predictive models were thus built to predict any one of the chronic/ acute event per patient.

Other Study Variables:

Full year consolidated files were used to identify patient demographics, lifestyle factors, access to care, cancer and overall health related characteristics and certain cardiovascular risk factors such as smoking, obesity, BMI, etc. Prescription medicines files were used to assess patient's adherence to medications and the combinations of drugs that were prescribed. These medications were mainly cardio protective agents or preventative medications such as β-blockers, calcium channel blockers or statins that are prescribed to those with and without CVD both as a preventative measure. Medication adherence was calculated using a proportion of days covered (PDC) measure. A PDC is calculated by dividing the number of days in period covered by the total number of days in the period. We calculated the number of days in period covered by summing the total number of days supplied of medication and the number of refills in a particular year obtained from MEPS whereas the total number of days in period were 365 since we used

cross sectional files. We calculated separate PDCs for separate classes of drugs per patient and then calculated an average PDC per patient as a behavioral measure. Outpatient and Office-based visits files were used to assess if any visits were scheduled to discuss patient's cardiovascular health. Discretization of some continuous variables like age and income was carried out to club them into meaningful categories, Table 4 below provides all the variables that were used to build the predictive risk model using the machine learning algorithms.

Table 4: Study Variables for Aim 1b

Variable	Variable description
Sociodemographic factors: -AGE05X – AGE15X* -SEX -RACEX -MARRY05X – MARRY15X* -EMPST31 -REGION31 -TTLP05X – TTLP15X* -EDUCYR -INSCOV05 – INSCOV15* -PRVEV05-15 (private insurance), MCREV05-15 (medicare), MCDEV05-15 (medicaid), OPAEV05-15 (other public insurance)*	 -Age as of the most recent round -Gender -Racial background -Marital status -Employment status -Census region -Income level -Years of education received -Health insurance coverage indicator -Type of insurance coverage (these will be clubbed into a single variable with multiple categories, Tricare and Employer's insurance would be coded as a separate 'others' category)
Access to care: -MDUNAB42 -PMUNAB42 -HAVEUS42 -DFTOUS42 -LOCATN42	 -Unable to get necessary medical care -Unable to get necessary medications -Does the person have a USC[a] provider -Difficulty in getting to the USC provider -Where is the USC provider located
Cancer and overall health related factors: -ADGENH42 -ASTHDX, DIABDX, ARTHDX -CANCERT -CNCRREMS	 -Overall health status -Comorbidities (asthma, diabetes and arthritis respectively) -Type of cancer (using the binary cancer type indicator variables) -Remission stage

-CHEMOTH -RXNAME -TC1 classification	-Radiotherapy, chemotherapy or surgery -Specific chemotherapy drug -Chemotherapy drug classification
Patient-related lifestyle factors: -ADSMOK42 -EXRCIS53 -NOFAT53 -BMINDX53	 -Do You currently smoke -Advised to exercise more -Restrict high fat/ cholesterol food -Adult BMI index
CVD screening services and risk factors: -CHECK53 -ADDRBP42 -BPCHEK53 -HIBPDX -CHOLCK53 -CHOLDX	 -Time since last routine check up -Did the doctor check blood pressure -Time since last blood pressure check -High blood pressure diagnosis -Time since last cholesterol check up -High cholesterol diagnosis
Medication adherence (To calculate PDC): -PURCHRD -RXBEGMM -RXDAYSUP -RXQUANTY -RXTOT05 – RXTOT15*	 -Round medication purchased in -Month person started taking medicine -Days supplied of the prescribed medicine -Number of tablets prescribed -Number of prescribed medicines including refills
Complex survey measures: -PERWT05X-PERWT15X* -VARPSU -VARSTR	 -Weight variable for the person's weight -Cluster variable -Strata variable

*All of these variables would be recoded/ renamed appropriately to a common variable applicable to all the years
a – Usual Source Care

Statistical Analyses:

Aim 1a: To compare the sociodemographic characteristics of cancer patients with CVD to those without CVD

The sociodemographic characteristics from table 2 were used to characterize the sample of patients with cancer and CVD both compared to those only with cancer. A binomial logistic regression was used to build this model to characterize the sample. Means and frequencies were used to summarize the descriptive statistics for continuous

and categorical variables respectively. SAS v9.4 was used for this aim. The logistic model built to characterize the sample also controlled for survey weights. The analyses thus conducted was weighted.

Aim 1b: To build predictive risk models using random forest, gradient boosting and deep learning algorithms and compare the accuracy to standard regression models

Models were trained using all the predictors mentioned in table 4, to predict CVD diagnoses after cancer. Missing values were imputed using multiple imputations techniques based on the missing data patterns. These models were built using different machine learning algorithms such as random forest (RF), gradient boosting and deep learning.[12] The accuracy of these models in predicting the risk of CVD diagnoses was compared against each other and against standard regression technique. An 80:20 data split was used for the study where, 80% of random data was used to train the model whereas the remaining 20% of the sample was used for validation. The data was shuffled multiple times before making the split to ensure the randomness of observations. Before running any of the ML algorithms, the data was also centered and scaled as a standard data pre-processing practice.

Random Forest (RF): RF analysis is based on an ensemble of classification trees. Literature also suggests that RF algorithms are the most highly applicable when it comes to predicting patients with high risks.[13] It has also been suggested in the literature that tree-based algorithms such as RF, are able to handle missing values through the modeling process alleviating the need for imputation.[14] This model was built using the R caret package. Gini impurity is the loss function that was used to evaluate the predictors that go

in the model to correctly classify patients as having CVD. Higher the gini coefficient, better the split and thus higher the accuracy. [13-15]

Gradient Boosting: Gradient boosting is a technique where in the models built in each iteration learn sequentially from the errors or false negatives (wrongly classified as having CVD) of the previous iteration and tends to minimize this error. R XGBOOST package was used to build gradient boosting models.[15,16] Prior to building these models, the predictors to be used in the model were normalized or one hot encoded (created indicator variables for each category of the categorical variables) to fit the data preprocessing standards of a gradient boosting model.[15,16] Variable importance plot (VIP) was used to evaluate the predictors that most significantly classify patients into getting diagnosed with CVD or not.

Deep Learning: Deep learning along with accounting for the predictors individually also accounts for the interactions within these more accurately. All the predictors mentioned above formed the input layer of the model. Based on the model architecture specified, multiple hidden layers were built to interact these predictors and assign a weight which then predicted the risk of CVD diagnoses. Python KERAS package was used to build this model.[17,18] The model architecture was defined by one input layer, two dense middle layers and an output layer. The activation function used was 'sigmoid'. The models were compiled using stochastic gradient descent (SGD) as an optimizer, 'binary crossentropy' as a loss function and 'accuracy' as an evaluation metric with a learning rate of 0.01. The SGD optimizes the model to find the global loss minimum with the lowest cost function that can then predict the outcome with highest accuracy. The same model architecture was used to build acute and chronic CVD prediction models both. Batch normalization was also carried out by adding another batch normalization layer to the model which standardizes

the inputs to every layer and stabilizes the learning process.[19,20] This layer was added for chronic and acute CVD prediction models both.

The prediction accuracies for the models were calculated using confusion matrix which determines the proportion of individuals correctly predicted by the model. The machine learning methods mentioned above, could not account for and were not compatible with sampling survey weights. The models built were thus unweighted. Adding the weight variables as general predictors into the model would not have been accurate since the weights by themselves should not predict the outcome in any way. They were still tested as predictors in the model, although this reduced the accuracy of the models.

Standard regression techniques: A binomial logistic regression model was used which predicted the likelihood of getting diagnosed with a CVD based on all the predictors mentioned above. Interaction terms were tested for their significance too to be included in the model. Stepwise regression technique was used to assess the predictors to be included in the regression model. SAS v9.4 was used for regression. The regression model also controlled for complex survey methods. The person, strata and cluster weight variables as mentioned in Table 4 were the complex survey methods controlled for in the study. In addition to the standard regression model using stepwise regression technique, a GLMNET model was also built. GLMNET model is one of the intermediate models between standard regression technique and machine learning algorithms. GLMNET is a package that fits a generalized linear model via penalized maximum likelihood function. This penalty function is determined by the regularization path that is computed using all the variables involved in the model. The regularization pathway determines the alpha and the lambda values which are the tuning parameters. The regularization pathway obtained

determines these values for tuning parameters to maximize the likelihood function and thus improve accuracy of the model. [15,21]

Model performance and accuracy:

The remaining 20% of the data was used to test the model. The trained model by using different algorithms was used to make predictions for CVD diagnoses on the test sample. The number of true positives and false positives were calculated using the predicted and the observed sample. These class predictions and predicted probabilities were used to calculate the Receiver Operating Characteristic Area Under the Curve (ROC AUC) or a c-statistic using R and Python both. Higher the c-statistic, higher is the number of true positives and lower is the number of false negatives, thus making the model more accurate. A c-statistic was calculated for each of the models built to assess model performance. Sensitivity and specificity of these models were reported too. Categories of c-statistic as identified in the literature were used to interpret the results with respect to their accuracy. A c-statistic lower than 0.5 indicates a very poor model, 0.5-0.69 indicates a model that is no better than predicting an outcome than random chance, 0.7-0.79 indicates a good model, 0.8-0.99 indicates a strong model whereas a value of 1 means that the model perfectly predicts a certain outcome.[22]

Handling missing data:

Missing data were imputed using the KNN-imputation method. This method imputes values for the missing variables using k-nearest neighbor averaging.[23] In this method, an estimate for the missing value can be approximated by the values of the points that are closest to it, based on other variables. Variables that had about 10-25% missing values

were imputed using KNN imputations. Listwise deletion were used for other variables with very low proportion of random missing values (<5%). Some variables had more than 25% missing observations, for these variables another category was created within the variable to indicate that the value was missing.

Aim 1c: To validate the predictive risk models using cross validation techniques and evaluate the model fit on a varied sample

The model was then validated using internal and external validation techniques.[15] For internal validation, the model was validated using a 10-fold cross validation technique on the entire sample. In order to further validate the model, the data was split based on the census regions of the population. Certain specific census regions were used to train the model whereas it was validated on the other census regions. Based on the prevalence of CVD across census regions, south and west regions were used to train the model whereas north and mid regions were used to test the model.[24] We clubbed a region with high prevalence of CVD (south) with another with a relatively low prevalence (west) as a training sample to ensure uniformity and reduce bias. Similarly, for the testing sample we clubbed a high prevalence region (north) with a relatively low prevalence region (mid).[24] For external validation, the entire dataset from 2005-2015 was used to train the sample whereas the model was tested on 2016-2017 dataset. Thus, chronological splitting was used for external validation of the model. This tested the robustness and generalizability of the model.

Aim 1d: To create an interactive web-based application using the R-shiny to predict the risk of CVD among cancer patients using the most accurate model identified

Using the most accurate machine learning/ regression model identified in aim 1b predicting the risk of CVD in cancer patients, an interactive web-based application was created which would dynamically predict probability of future CVD events in cancer patients. The application would be a user-friendly tool for the physicians to input patient and other characteristics of cancer patients to get an estimated probability of potential CVD risk in the future. The characteristics to be entered were based on the model results from aim1b and the variable importance plot which estimated the most important predictors of CVD risk in cancer patients. This web-based application was designed on R-shiny using the user interface (UI) and the server functions to define the shiny object (web application).[25,26] The R built model predicting the risk of CVD most accurately was fed into the server function which then feeds into the user interface of the application to make predictions. Based on this model, the application predicts a probability of any future CVD event given the characteristics.

RESULTS:

The study sample consisted of 48,829 patients diagnosed with cancer in the United States (US) from 2005-2015. Of these, 16,951 (34.72%) patients also had a chronic CVD diagnosis whereas the remaining 31,878 (65.28%) patient had no chronic CVD diagnoses. Of the 16,951 patients, 4,612 patients had a CVD diagnosis before their cancer diagnosis and were thus excluded from the study. The final analyses were thus conducted on 12,339 cancer patients with CVD (25.27%) as compared to 31,878 cancer patients without CVD (65.28%). There were fewer patients with CVD that reported their cancer type. Table 5 below summarizes the distribution of patients with cancer and CVD both by their cancer type. Breast and prostate were the most frequent types of cancers among patients with CVD.

Table 5: Patients With Cancer and CVD Both By Cancer Type In The US From 2005-2015

Cancer type	Unweighted Frequency
Prostate	1,824
Breast	1,021
Lung	834
Colon	514
Melanoma	417
Cervix	103
Pancreas	20
Other	2,162

Table 6: Acute CVD Events In Patients With Cancer In The US From 2005-2015

CVD events in cancer patients	Unweighted Frequency
Hospital stay related to the condition	7,069
Inpatient procedure	3,837
Emergency room (ER) visits	3,423

Acute CVD events were also analyzed in our study. Table 6 above summarizes the number of acute CVD events that occurred in cancer patients post their cancer diagnoses. As observed from the table above, majority of acute CVD events were due to a hospital stay related to the CVD condition without any inpatient procedure or an ER visit required. These were the number of CVD events that occurred in cancer patients after their cancer diagnosis.

Aim 1A: To compare the sociodemographic characteristics of cancer patients with CVD to those without CVD

Sociodemographic characteristics such as age, gender, race, marital status, education, employment status and income were compared across cancer patients with and without CVD. These results are summarized below in Table 7.

Table 7: Sociodemographic Characteristics of Cancer Patients With And Without CVD In The US From 2005-2015

Sociodemographic characteristis	**Cancer patients with CVD N(row%, column%) 12,339(28.7)**	**Cancer patients without CVD N(row%, column%) 31,878(71.3)**	**p-value**
Age, years (range)			
Mean	71.2	62.1	**<.0001***
Age groups			
18-44 years	302 (5.81, 2.0)	4,211 (94.18, 13.2)	
45-65 years	3,791 (22.53, 26.3)	11,922 (77.47, 36.4)	**<.0001***
>65 years	7,501 (37.64, 66.6)	13,590 (62.35, 44.4)	
Missing	745 (24.94, 5.1)	2,155 (75.05, 6.1)	
Gender			
Males	5876 (31.36, 47.0)	13,139 (68.63, 41.3)	
Females	6,463 (26.63, 53.0)	18,739 (73.37, 58.7)	0.2805
Missing	0 (0)	0 (0)	
Race			
Whites	9,437 (29.42, 89.8)	24,504 (70.57, 86.6)	
Blacks	2,376 (27.85, 8.12)	5,207 (72.14, 8.5)	0.1363
Asians	449 (16.05, 1.6)	1,518 (83.94, 3.3)	

Multiple Races	77 (9.64, 0.4)	649 (90.35, 1.6)	
Missing	0 (0)	0 (0)	
Marital Status			
Married	5,916 (26.5, 52.1)	17,518 (73.48, 58)	
Widowed	2,394 (41.59, 24)	4,684 (58.40, 13.6)	
Divorced	2,802 (29.03, 15.8)	4,535 (70.96, 15.5)	**<.0001***
Separated	602 (71.11, 4.31)	427 (28.88, 0.7)	
Never Married	625 (12.06, 3.73)	4,266 (87.94, 11)	
Missing	0 (0)	448 (100, 1.2)	
Education			
No education	6 (3.75, 0.04)	227 (96.24, 0.4)	
Elementary/Middle School	1,105 (30.60, 5.2)	2,167 (69.39, 4.8)	
High School	2,357 (21.83, 16.7)	8,389 (78.16, 24.1)	**0.0479***
≤4 Years College	2,131 (22.82, 17.9)	7,558 (77.17, 24.4)	
5+ Years College	713 (25.25, 8.0)	2,175 (74.74, 9.5)	
Missing	6,027 (36.21, 52)	11,362 (63.78, 36.7)	
Employment Status			
Employed	2,555 (19.73, 23.8)	11,259 (80.26, 39)	
Unemployed	9,780 (34.14, 76.1)	19,931 (65.85, 59)	**<.0001***
Missing	4 (0.68, 0.03)	688 (99.31, 2)	
Income, per year			
Mean ($)	34,117	34,615	0.3459
Income groups[a]			
Low	4,328 (24.38, 22)	10,887 (75.61, 27.4)	
Middle Class	5,783 (31.63, 54.5)	14,865 (68.36, 47.4)	0.3041
High	2,228 (27.18, 23.4)	6,126 (72.81, 25.2)	
Missing	0 (0)	0 (0)	

The frequencies reported are unweighted and percentages reported are weighted using the complex survey design
*Significantly different characteristics across the two groups
a – Low defined as income <100% FPL (<$12,060), middle class defined as income between the range of 100-400% FPL ($12,060 - $48,240) and high defined as income over 400% FPL (>$48,240)

As seen from the table above, age, marital status, education and employment status differed significantly across cancer patients with and without CVD on conducting bivariate analyses. Cancer patients with CVD were older (mean age: 71.2years vs 62.1years), had lower proportion of those married (52.1% vs 58%) and higher proportion of unemployed (76.1% vs 59%). These factors were included in a logistic model to calculate adjusted odds ratios. Results from the adjusted logistic regression are summarized in table 8 below. Odds ratios (OR) adjusted for all the sociodemographic

factors are reported in table 8. The probability modeled was getting diagnosed with CVD. In presence of all the other variables, age, race, marital status, education and employment status had significant OR associated with these. The significance of these variables was estimated based on the overall p-value as observed from the type 3 effect. However, on assessing the confidence intervals associated with these odds ratio estimates it was observed that the significance could be because of a wide range of interval associated with some categories. This could have been a result of low sample size for those particular categories as compared to the sample size of the 'Reference' category. This suggests that the significance of certain variables such as race, marital status and education could have been attributable to low sample size for significant goups (Race- Multiple races, Marital Status – Separated and Education – No education) withing these variables. The odds ratio estimates suggested that as compared to those who were >65years, those who were younger were less likely to be diagnosed with CVD. As compared to whites, all the other races were less likely to be diagnosed with CVD. As compared to those who were married, those who were widowed, divorced or separated were more likely whereas those never married were less likely to be diagnosed with CVD. As compared to those who had ≤4 years college, those who had no education or studied till highschool were less likely where as those who studied till elementary/middle school and 5+ years college were more likely to get diagnosed with CVD. Those employed were less likely to be diagnosed with CVD as compared to those who were unemployed.

Table 8: Survey logistic Results for the Likelihood of Being Diagnosed With CVD Along with Cancer Based on the Sociodemographic Characteristics Associated With Cancer Patients In the US From 2005-2015

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
Age*			
18-44 years	0.169	0.086	0.331
45-65 years	0.643	0.394	1.047
>65 years	Reference	Reference	Reference
Gender			
Males	1.262	0.804	1.980
Females	Reference	Reference	Reference
Race*			
Blacks	0.984	0.498	1.945
Asians	0.594	0.247	1.432
Multiple Races	0.222	0.067	0.730
Whites	Reference	Reference	Reference
Marital Status*			
Widowed	1.559	0.850	2.764
Divorced	1.073	0.564	2.047
Separated	13.017	3.527	48.049
Never Married	0.712	0.335	1.505
Married	Reference	Reference	Reference
Education*			
No education	0.100	0.011	0.877
Elementary/ Middle School	1.215	0.359	4.106
Highschool	0.997	0.489	2.032
5+ Years College	1.239	0.504	3.045
≤4 Years College	Reference	Reference	Reference
Employment Status*			
Employed	0.568	0.349	0.924

Unemployed	Reference	Reference	Reference
Income level			
Low	0.774	0.493	1.215
High	1.268	0.754	2.132
Middle class	Reference	Reference	Reference

*Significantly associated characteristics with the likelihood of being diagnosed with CVD alongwith cancer
The probability modeled above is cancer patients being diagnosed with CVD

Aim 1B: To build predictive risk models using random forest, gradient boosting and deep learning algorithms and compare the accuracy to standard regression models

In order to build the predictive risk model, the entire dataset was split into training and test validation splits. The baseline characteristics of training and test datasets are summarized below in table 9. This table also shows the proportion of missing values within each predictor. For the categorical predictors that had less than 5% missing data, listwise deletion was used to perform further analyses on complete cases. Those that had more than 10% but less than 25% missing data were imputed using KNN imputation method. This KNN imputation method used 6 nearby neighbors to impute the missing observation with a mean of 6 nearby neighbors. Those categorical variables that had more than 25% missing data used another level of category to indicate that data was missing. The missing continuous variables with less than 10% missing were imputed using the mean imputation technique by assigning the overall mean value of the variable to the missing observation.

Table 9: Baseline Characteristics of Training and Test Data Splits Used For Prediction

Characteristic	Training dataset N (%) 35,858 (80)	Test dataset N (%) 8,964 (20)
Sociodemographic factors		
Age groups		
18-44 years	3,733 (10.4)	916 (10.2)
45-65 years	12,726 (35.5)	3,194 (35.6)
>65 years	16,986 (47.4)	4,289 (47.8)
Missing	2,413 (6.7)	565 (6.3)
Gender		
Males	15,398 (43)	3,905 (43.6)
Females	20,460 (57)	5,059 (56.4)
Missing	0 (0)	0 (0)
Race		
Whites	27,536 (76.8)	6,829 (76.2)
Blacks	6,147 (17.1)	1,551 (17.3)
Asians	1,548 (4.3)	425 (4.7)
Multiple Races	627 (1.7)	159 (1.8)
Missing	0 (0)	0 (0)
Marital status		
Married	18,879 (52.6)	4,718 (52.6)
Widowed	5,722 (15.9)	1,468 (16.4)
Divorced	5,962 (16.6)	1,505 (16.8)
Separated	831 (2.3)	203 (2.2)
Never Married	4,061 (11.3)	981 (10.9)
Missing	403 (1.1)	89 (0.9)
Education		
No education	197 (0.5)	36 (0.4)
Elementary/Middle School	2,652 (7.4)	673 (7.5)
High School	8,750 (24.4)	2,206 (24.6)
≤4 Years College	7,732 (21.5)	1,990 (22.2)
5+ Years College	2,368 (6.6)	563 (6.3)
Missing	14,159 (39.4)	3,496 (39)
Employment status		
Employed	11,163 (31.1)	2,792 (31.1)
Unemployed	24,052 (67.1)	6,019 (67.1)
Missing	643 (1.8)	153 (1.7)

Income groups[a]		
Low	12,520 (34.9)	3,105 (34.6)
Middle Class	16,689 (46.5)	4,108 (45.8)
High	6,649 (18.5)	1,751 (19.5)
Missing	0 (0)	0 (0)
Census region		
Northeast	5,982 (16.7)	1,460 (16.3)
Midwest	7,792 (21.7)	1,947 (21.7)
South	14,482 (40.4)	3,666 (40.9)
West	7,509 (20.9)	1,873 (20.9)
Missing	93 (0.2)	18 (0.2)
Health insurance coverage		
Any private	19,185 (53.5)	4,929 (55)
Public only	15,091 (42.1)	3,642 (40.6)
Uninsured	1,582 (4.4)	393 (4.3)
Missing	0 (0)	0 (0)
Type of insurance coverage*		
Private insurance	17,907 (49.9)	4,577 (51.1)
Medicare	21,299 (59.1)	5,349 (59.7)
Medicaid	8,144 (22.7)	1,960 (21.9)
Other public insurance	67 (0.2)	16 (0.2)
Missing	0 (0)	0 (0)
Access to care		
Unable to get necessary medical care		
Yes	1,036 (2.9)	253 (2.8)
No	32,604 (90.9)	8,187 (91.3)
Missing	2,218 (6.2)	524 (5.8)
Unable to get necessary medications		
Yes	1,581 (4.4)	399 (4.5)
No	32,089 (89.5)	8,040 (89.7)
Missing	2,188 (6.1)	525 (5.9)
Person has a USC[b] provider		
Yes	31,525 (87.9)	7,900 (88.1)
No	2005 (5.6)	507 (5.7)
Missing	2,328 (6.5)	557 (6.2)
Difficulty to get to the USC provider		
Very difficult	294 (0.8)	66 (0.7)
Somewhat difficult	1,982 (5.5)	479 (5.3)
Not too difficult	4,260 (11.9)	1,081 (12.1)
Not at all difficult	13,158 (36.7)	3,365 (37.5)

Missing	16,164 (45.1)	3,973 (44.3)
Location of the USC provider		
Office	24,464 (68.2)	6,141 (68.5)
Hospital, not ER	6,920 (19.3)	1,726 (19.3)
Hospital, ER	44 (0.12)	11 (0.1)
Missing	4,430 (12.4)	1,086 (12.1)
Cancer and overall health related factors		
Overall health status		
Excellent	2,093 (5.8)	493 (5.5)
Very good	6,385 (17.9)	1,687 (18.8)
Good	10,593 (29.5)	2,649 (29.6)
Fair	8,569 (23.9)	2,149 (24)
Poor	3,134 (8.7)	757 (8.4)
Missing	5,084 (14.2)	1,229 (13.7)
Comorbidities*		
Asthma	3,835 (10.7)	932 (10.4)
Diabetes	7,035 (19.6)	1,826 (20.4)
Arthritis	15,703 (43.8)	4,004 (44.7)
Missing	7,175 (20.01)	1,900 (21.2)
Cancer type		
Breast	2,601 (7.2)	628 (7)
Lung	939 (2.6)	234 (2.6)
Cervix	305 (0.8)	71 (0.7)
Prostate	2,448 (6.8)	582 (6.4)
Pancreas	7 (0.02)	2 (0.02)
Colon	905 (2.5)	211 (2.4)
Melanoma	734 (2.1)	193 (2.1)
Other	3,719 (10.4)	916 (10.2)
Missing	18,630 (51.9)	4,657 (51.9)
Cancer treatment received*†		
Radiotherapy	429 (1.2)	99 (1.1)
Chemotherapy	708 (2)	173 (1.9)
Surgery	2,960 (8.2)	673 (7.5)
Missing	31,761 (88.5)	8,019 (89.4)
Cancer in remission		
Yes	2,546 (7.1)	627 (7.0)
No	430 (1.2)	99 (1.1)
Missing	32,882 (91.7)	8,238 (91.9)
Medication adherence		
PDC (>80%)	17,136 (47.79)	4,265 (47.57)

PDC (<80%) Missing	6,942 (19.36) 11,780 (32.85)	1,766 (19.70) 2,933 (32.73)
Patient-related lifestyle factors		
Current smoking status Smoker Nonsmoker Missing	 4,031 (11.2) 26,914 (75.1) 4,913 (13.6)	 1,024 (11.4) 6,746 (75.3) 1,194 (13.3)
Advised to exercise more Yes No Missing	 15,997 (44.6) 15,241 (42.5) 4,620 (12.8)	 4,015 (44.8) 3,829 (42.7) 1,120 (12.5)
Advised to restrict high fat/ cholesterol food Yes No Missing	 16,030 (44.7) 15,585 (43.46) 4,243 (11.8)	 4,024 (44.9) 3,887 (43.4) 1,053 (11.7)
Adult BMI index[c] Underweight (< 18.5 kg/m^2) Normal weight (18.5 – 24.9 kg/m^2) Overweight (25 – 29.9 kg/m^2) Obese (>30 kg/m^2) Missing	 5,338 (14.8) 10,070 (28.1) 11,002 (30.6) 8,949 (24.9) 499 (1.4)	 1,300 (14.5) 2,506 (27.9) 2,813 (31.4) 2,228 (24.8) 117 (1.3)
CVD screening factors and risk factors		
Time since last routine check up < 1 year 1-3 years 4-5 years >5 years Never Missing	 26,954 (75.2) 2,554 (7.1) 414 (1.2) 727 (2) 572 (1.6) 4,637 (12.9)	 6,721 (75) 652 (7.3) 107 (1.2) 199 (2.2) 152 (1.7) 1,133 (12.6)
Blood pressure checked by doctor Yes No Missing	 29,724 (82.9) 765 (2.1) 5,369 (15)	 7,445 (83.1) 204 (2.3) 1,315 (14.7)
Time since last blood pressure check < 1 year 1-3 years 4-5 years >5 years Never	 31,149 (86.9) 664 (1.9) 16 (0.04) 91 (0.3) 21 (0.05)	 7,818 (87.2) 167 (1.9) 3 (0.03) 18 (0.2) 4 (0.04)

Missing	3,917 (10.9)	954 (10.6)
High blood pressure diagnosis		
Yes	19,520 (54.4)	4,936 (55.1)
No	8,751 (24.4)	2,217 (24.7)
Missing	7,587 (21.2)	1,811 (20.2)
Time since last cholesterol check		
< 1 year	27,460 (76.6)	6,873 (76.7)
1-3 years	2,307 (6.4)	575 (6.4)
4-5 years	200 (0.6)	50 (0.6)
>5 years	592 (1.6)	159 (1.8)
Never	708 (1.9)	185 (2.1)
Missing	4,591 (12.8)	1,122 (12.5)
High cholesterol diagnosis		
Yes	17,192 (47.9)	4,390 (49)
No	11,068 (30.9)	2,759 (30.8)
Missing	7,598 (21.2)	1,815 (20.2)
CVD outcome		
CVD diagnosis (Chronic)		
Yes	9,908 (27.6)	2,518 (28.1)
No	25,950 (72.4)	6,446 (71.9)
CVD (Acute)		
Yes	7,949 (22.17)	2,097 (23.4)
No	27,909 (77.83)	6,867 (76.6)

The frequencies and percentages reported are unweighted
*People could have more than one of the listed alternatives
† Specific cancer therapy for those available was also controlled
¥ Medication adherence calculated using PDC (>80% - adherent, <80% - adherent) from round/ month medication obtained in, days suppleid of medication and number of refills
a – Low defined as income <100% FPL (<$12,060), middle class defined as income between the range of 100-400% FPL ($12,060 - $48,240) and high defined as income over 400% FPL (>$48,240)
b – Usual Source Care
c – BMI ranges obtained from the American Cancer Society definitions

On imputing the missing data using KNN imputation and using listwise deletion to retain the complete cases, the training and test datasets were centered and scaled as a part of data pre-processing that is required for machine learning algorithm. Once the data was ready, it was used to build conventional regression, random forest, gradient boosting

and deep learning models. These models were built to predict the risk of an acute cardiovascular condition among patients with cancer. Similar models were built to predict the risk of chronic CVD as well.

Regression model:

Figure 3 below depicts the ROC curve obtained using a conventional regression model to predict the risk of an acute CVD event using all the predictors mentioned in table 9 above. As observed from the ROC curve, the c-statistic obtained was 0.7534. A c-statistic in the range of 0.7 – 0.79 indicates a good model. This shows that the model built using standard regression was good in making predictions. A similar model was built for predicting chronic CVD events using the same predictors. The regression model for chronic CVD events generated a c-statistic of 0.7641 which was slightly higher than that for acute CVD events. The ROC obtained on building this model is summarized in figure 4 below. This suggests that the conventional regression model built for predicting chronic CVD events was more accurate than that built for acute CVD events. A c-statistic of 0.7641 indicates a good model in predicting the risk of chronic CVD in cancer patients.

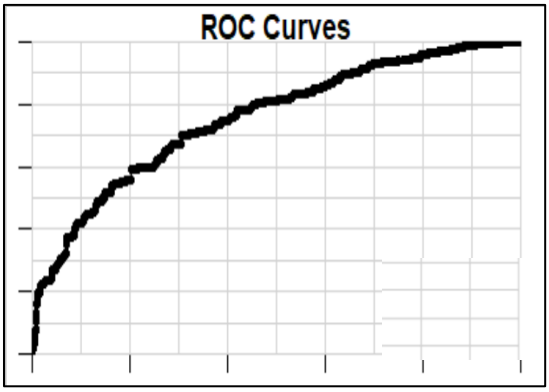


Figure 3: ROC for Acute CVD Prediction Using a Regression Model

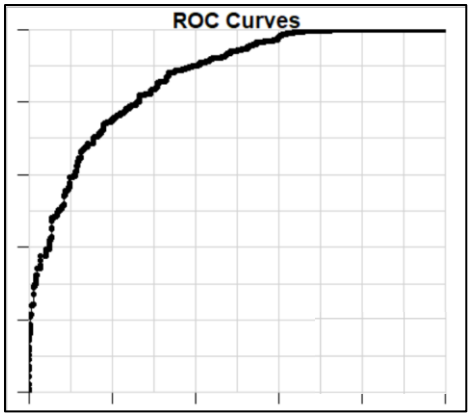


Figure 4: ROC for Chronic CVD Prediction Using a Regression Model

GLMNET package was used to build the GLMNET model. The regularization paths obtained for the acute and chronic CVD models are summarized in appendix figures 1 and 2. The regularization paths suggested the use of ridge regression for acute and chronic CVD models both with $\alpha=0$ as the penalty function. The c-statistic obtained by conducting ridge regression for acute and chronic CVD was 0.7853 (Sensitivity = 0.9846) and 0.7349 (Sensitivity = 0.9473) respectively. Thus, it was observed that for acute CVD, the GLMNET (ridge regression) model performed better than the conventional regression model, whereas for chronic CVD prediction model the traditional regression model performed better than the GLMNET model.

Random forest model:

Random forest model was built using all the predictors mentioned in table 9 above to predict the risk of an acute CVD event. The model was built using 5 cross validation sets which use five bootstraps of samples on the same set of variables. On trying different number and combinations of predictors to predict the outcome by building different trees, a model with 9 variables was chosen as the best model. This model with 9 variables would predict the risk of acute CVD event most accurately and with the lowest validation error.

Figure 5 below depicts the learning curves of a random forest model built to predict acute CVD event. It can be seen from the figure that the model has the highest cross validation accuracy and highest gini gain when the number of predictors are 9 after which the model does not improve much. The accuracies based on the gini gains with each set of predictor combination are summarized in table 10 below.

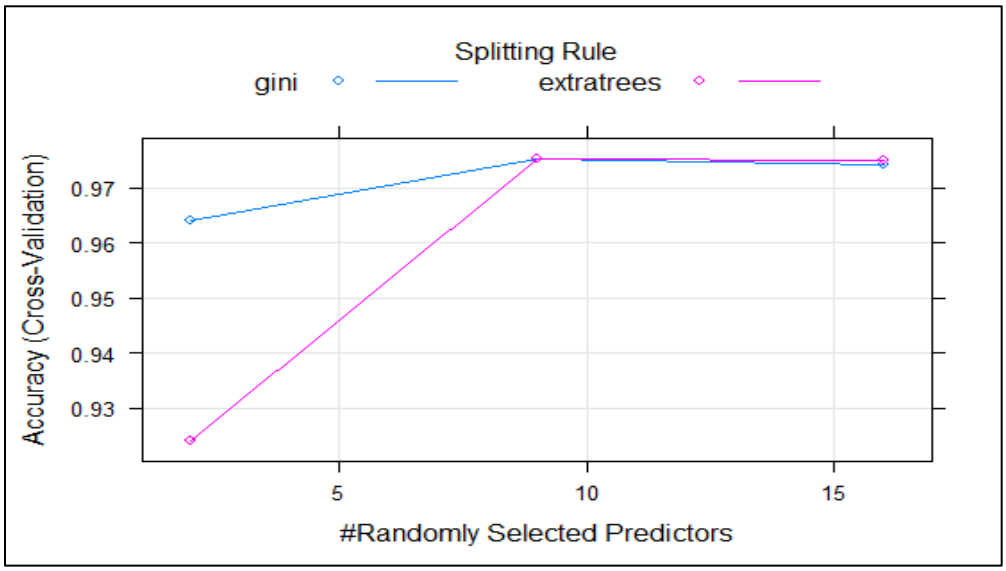


Figure 5: Learning Curves for a Random Forest Model Predicting an Acute CVD Event

Table 10: Learning Accuracies for Random Forest Model Predicting an Acute CVD Event

Number of predictors	Accuracy
2	0.9641
9	0.9754
16	0.9741

Similar model results were obtained for building predictive models for chronic CVD events. In this case, the most accurate model in predicting the risk of chronic CVD events in cancer patients was the one with 7 predictors. The learning curves obtained on training the random forest model are as observed in figure 6 below.

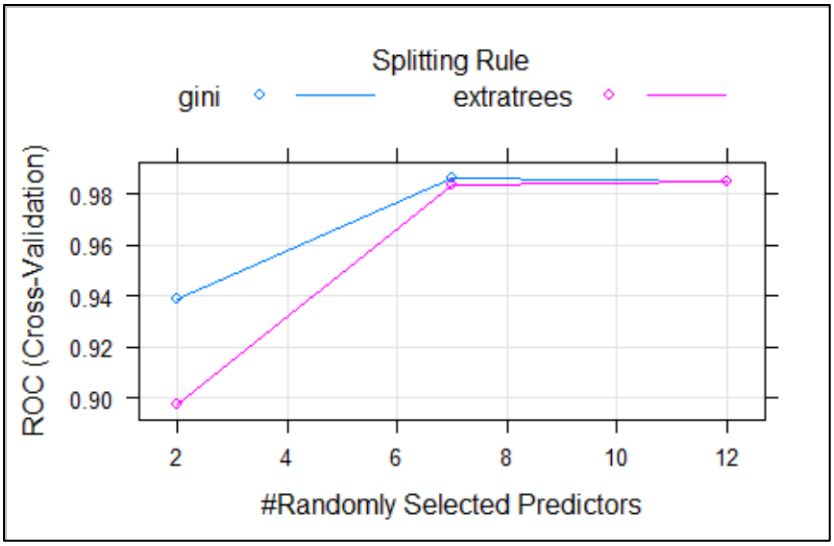


Figure 6: Learning Curves for a Random Forest Model Predicting Chronic CVD

It can be seen from the figure that the model had the highest cross validation accuracy and highest gini gain when the number of predictors are 7 over which the model does not improve much. The accuracies based on the gini gains with each set of predictor combination are summarized in table 11 below. The model with 7 predictors had the highest accuracy.

Table 11: Learning Accuracies for Random Forest Model Predicting Chronic CVD

Number of predictors	Accuracy
2	0.9414
7	0.9888
12	0.9874

The model results from the random forest models built for chronic and acute CVD prediction in cancer patients were used to build the variable importance values. These values assign a relative importance value starting with 100 for the most important variable and ranking the variables in a descending order. Following table 12 and table 13

summarizes the variable importance values for the most important predictors identified using the random forest model for chronic and acute CVD events respectively.

Table 12: Variable Important Values for the Random Forest Model Built for Predicting Chronic CVD

Predictor	Overall value
Medication adherence	100
Cholesterol diagnosis	98.20
High blood pressure diagnosis	77.39
Diabetes diagnosis	69.03
Arthritis diagnosis	68.89
Overall health	59.52
Difficulty to get to the provider	45.23
Census region	41.55
Income	36.71
Marital status	35.86

Table 13: Variable Important Values for the Random Forest Model Built for Predicting Acute CVD Events

Predictor	Overall value
Overall health	100
Medication adherence	80.96
Marital status	63.10
Census region	62.47
Income	47.93
Provider location	44.26
Gender	41.75
Difficulty to get to the provider	40.51
Unable to get necessary prescribed medicines	38.20
Cholesterol diagnosis	36.87

Gradient Boosting model:

Before building a gradient boosting model, the data matrix was converted using one hot encoding into a matrix with just binary responses by flagging each indicator. A gradient boosting model was trained using 100 iterations with the validation error reducing with every iteration. The evaluation metric used for the model was ‘error’. The same modeling techniques were used to build predictive models for acute as well as chronic CVD events.

Every 10 rounds of iteration were printed and are summarized in table 14 below for acute CVD events and table 15 for chronic CVD events. For acute CVD events and chronic CVD, tables 14 and 15 show that the train and the test error reduce with every iteration with iteration 91 being the point where the models are at their best since the error does not change much beyond that point. The final validation error for acute CVD prediction model obtained was 0.1087 and that obtained for chronic CVD prediction model was 0.1483. The ROC obtained for acute CVD prediction model is as seen in figure 7 below with a c-statistic of 0.7833 and that obtained for chronic CVD is as observed in figure 8 below with a c-statistic of 0.7608. The gradient boosting model was thus less accurate in predicting chronic CVD as compared to acute CVD event.

Table 14: Errors Associated With Each Iteration of Gradient Boosting Algorithm For Predicting Acute CVD Event

Iteration	Train error	Test error
1	0.1576	0.1589
11	0.1303	0.1325
21	0.1229	0.1255
31	0.1169	0.1200
41	0.1143	0.1173
51	0.1112	0.1149
61	0.1092	0.1136
71	0.1065	0.1111
81	0.1054	0.1101
91	0.1039	0.1092
100	0.1032	0.1087

Table 15: Errors Associated With Each Iteration of Gradient Boosting Algorithm For Predicting Chronic CVD Event

Iteration	Train error	Test error
1	0.2149	0.2160
11	0.1831	0.1850
21	0.1677	0.1698
31	0.1600	0.1630

41	0.1570	0.1607
51	0.1522	0.1563
61	0.1486	0.1528
71	0.1469	0.1512
81	0.1451	0.1502
91	0.1433	0.1484
100	0.1427	0.1483

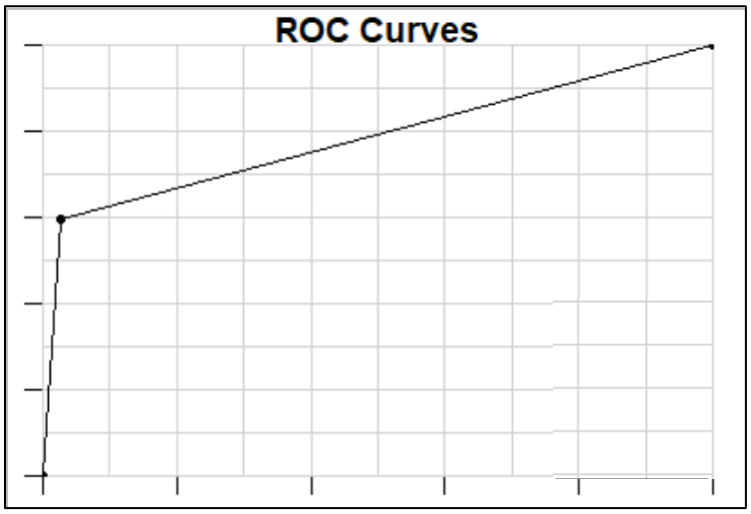


Figure 7: ROC Obtained with a Gradient Boosting Model for Acute CVD Event

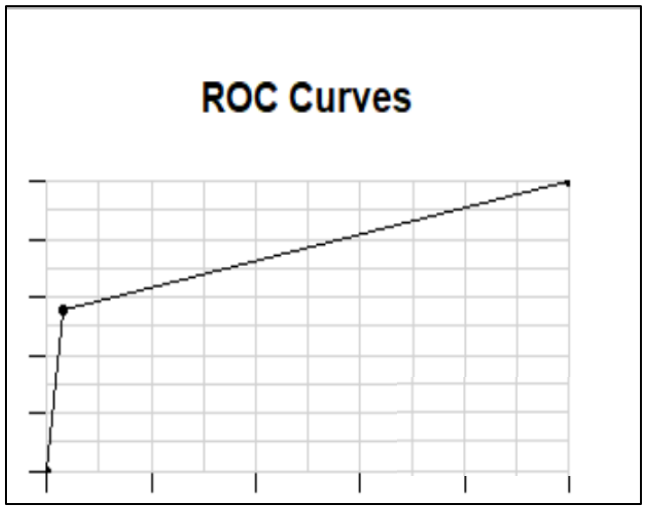


Figure 8: ROC Obtained with a Gradient Boosting Model for Chronic CVD

The model results were also used to plot a variable importance plot (VIP) to evaluate the predictors that have the most significant effect on predicting the acute CVD

outcome. The VIP is depicted in figure 9 below which suggests that medication adherence (identified from the number of refills), overall health, BMI, marital status and education were identified as the top 5 predictors of an acute CVD event.

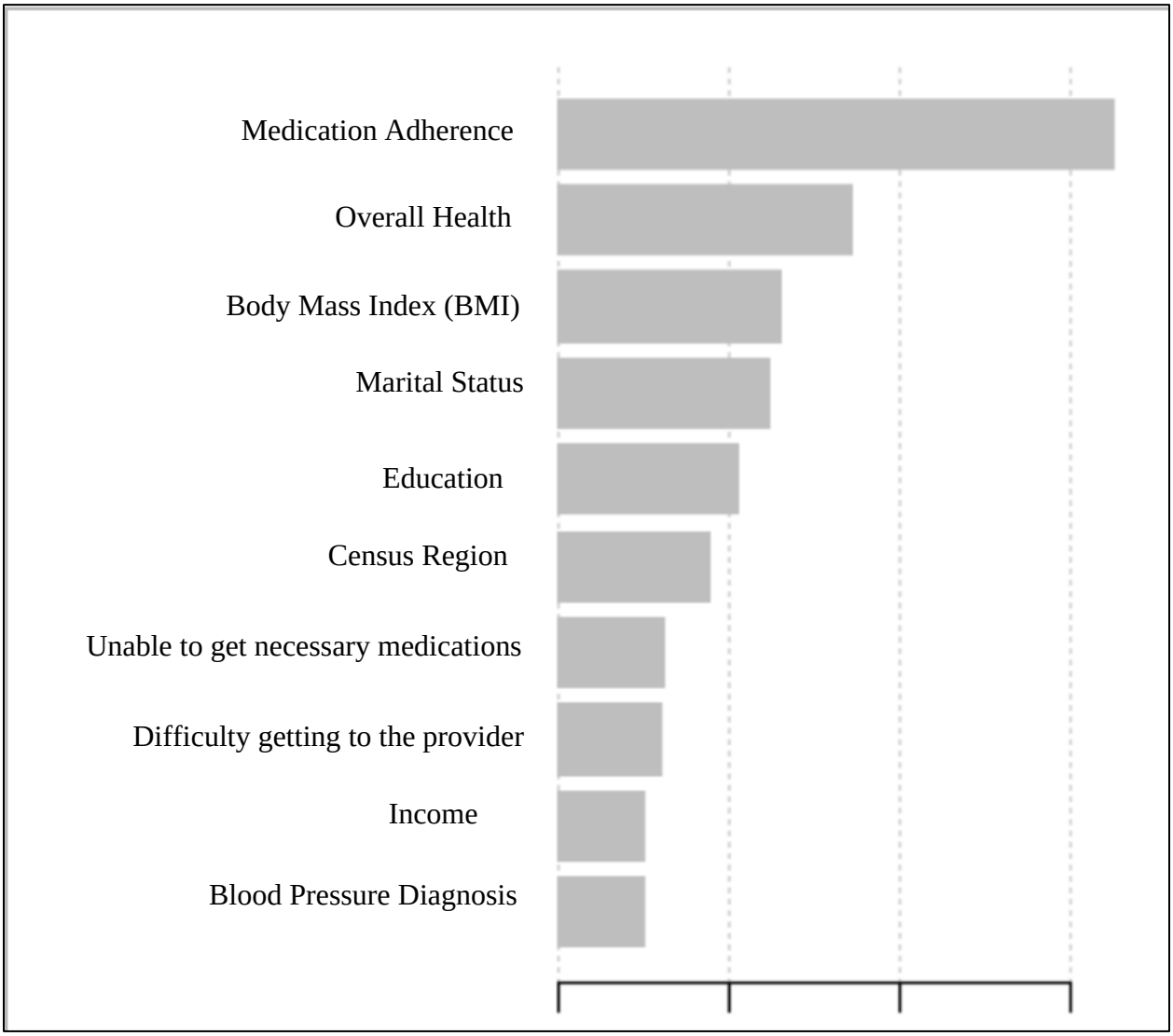


Figure 9: Variable Importance Plot Based on The Gradient Boosting Model for Acute CVD Events

Similar VIP was also used to identify the most important predictors responsible in predicting a chronic CVD outcome among cancer patients. Figure 10 below summarizes the most important predictors which have identified medication adherence (identified using

the number of refills), cholesterol, arthritis, blood pressure diagnosis and overall health to be the most important in predicting the risk of chronic CVD in cancer patients.

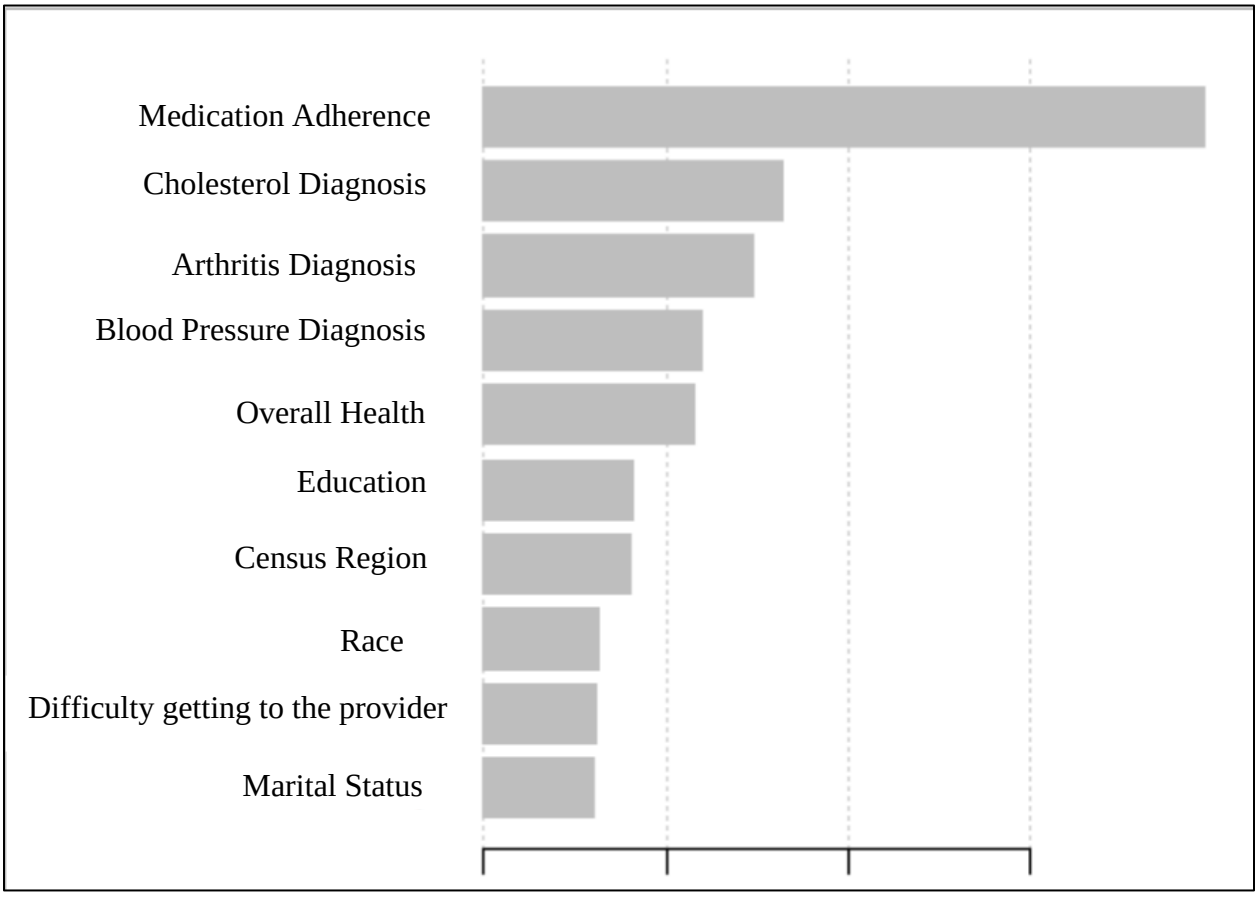


Figure 10: Variable Importance Plot Based on The Gradient Boosting Model for Chronic CVD Events

On comparing the most important predictors that were suggested by using the gradient boosting model (figure 9) to the most important predictors suggested by the random forest model (table 13) in predicting acute CVD event, most of the significant predictors identified were the same. Both the algorithms thus, predicted similar factors to be the most significant in predicting the risk of acute CVD events in cancer patients. Other than the predictors that were the same in both the algorithms, random forest model

suggested provider's location, gender and cholesterol diagnosis instead of BMI, education and high blood pressure diagnosis in the gradient boosting model for predicting acute CVD events among cancer patients. Similar results were obtained for the chronic CVD prediction model with both the algorithms (figure 10 and table 12) predicting similar factors to be the most significant in predicting chronic CVD risk in cancer patients. In addition to the predictors that were the same in both the algorithms, random forest model predicted diabetes diagnosis and income instead of education and race in the gradient boosting model in predicting chronic CVD in cancer patients.

Deep learning model:

Deep learning models were trained using 50 epochs and a batch size of 16 to build predictive models to predict the risk of acute CVD events and chronic CVD in cancer patients. Two separate models each predicting an acute CVD event and chronic CVD respectively were built using the same architecture. The models were then fit on the training and test samples to obtain learning curves as depicted in figure 11 and 12 below for acute CVD and chronic CVD events respectively. The learning curves obtained after conducting batch normalization for acute and chronic CVD prediction models are depicted in appendix figures 3 and 4 respectively. For an acute CVD event prediction model, the model accuracy was the highest with around 30 epochs with the highest accuracy of around 0.83. The c-statistic obtained for this model was 0.8267 which was higher compared to a standard regression, GLM net and gradient boosting models for acute CVD event prediction model. For a chronic CVD prediction model, the model accuracy was the highest with 10 epochs with the highest accuracy of around 0.77. The c-statistic obtained for this model was 0.7662 which was higher compared to standard regression, GLM net

models and slightly higher than the gradient boosting model for chronic CVD prediction model.

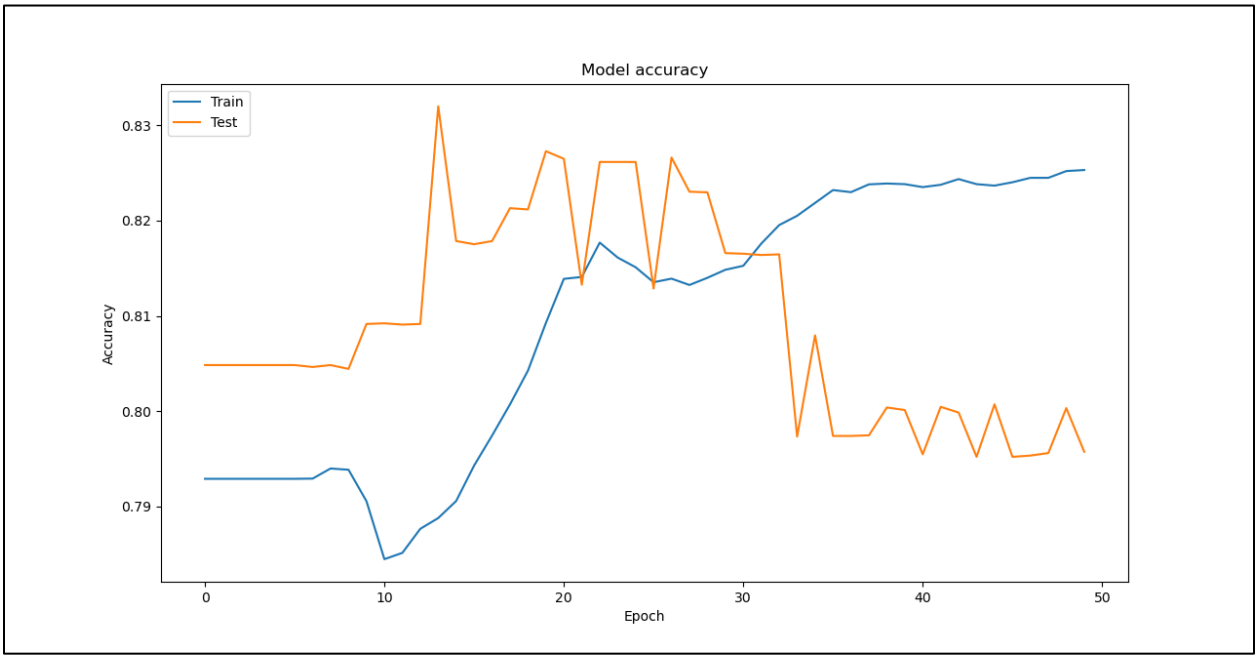


Figure 11: Learning Curves for the Deep Learning Acute CVD Prediction Model

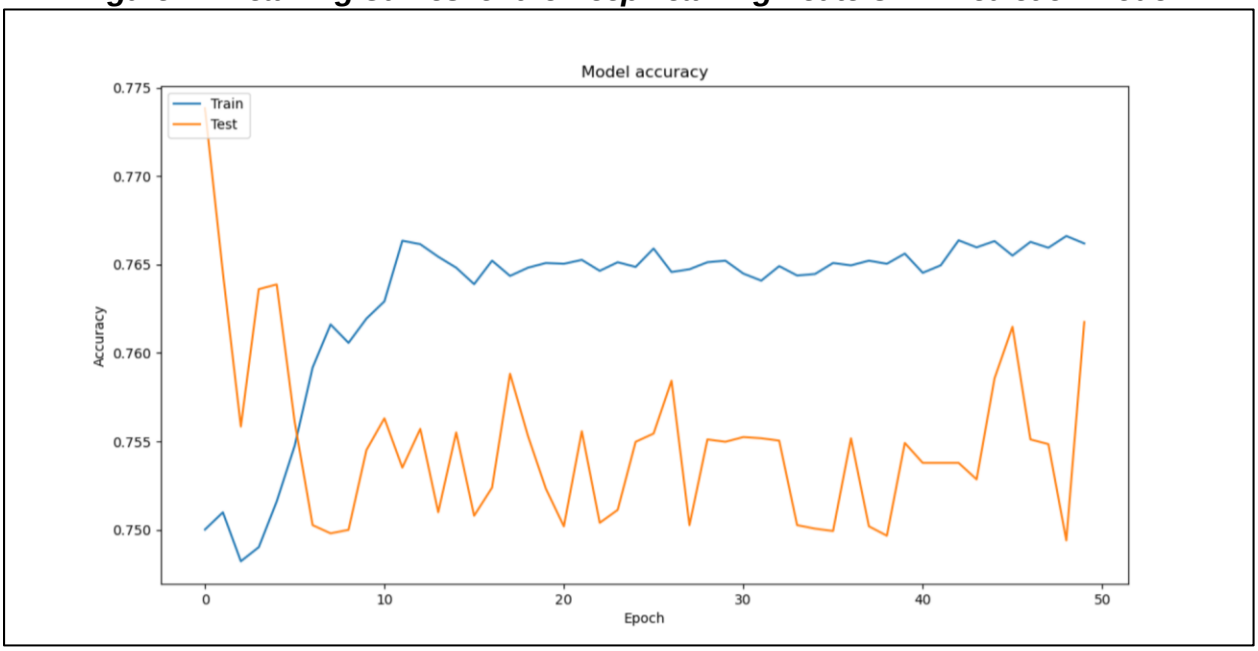


Figure 12: Learning Curves for the Deep Learning Chronic CVD Prediction Model

Comparison of all the predictive risk models:

A comparison of all the models built to predict the risk of an acute CVD event has been summarized below in Table 16. This table shows that a standard regression model had the lowest predictive power with a c-statistic of 0.7534 whereas a random forest model had the highest predictive power with a c-statistic of 0.9738 followed by a deep learning model. The gradient boosting and GLM net models had almost the same predictive power. This suggests that with a large number of predictors the machine learning algorithms perform better than the conventional regression technique.

Table 16: Comparison of Acute CVD Prediction Models

Model	c-statistic
Standard regression	0.7534
GLM net	0.7853
Random Forest	0.9738
Gradient Boosting	0.7833
Deep Learning	0.8267

A comparison of all the models built to predict the risk of chronic CVD has been summarized below in Table 17. This table shows that a GLMNet model had the lowest predictive power with a c-statistic of around 0.73 whereas a random forest model had the highest predictive power with a c-statistic of 0.9872. Standard regression, gradient boosting and deep learning models had a c-statistic of around 0.76. This suggests that the conventional regression model in this case performed as well as some of the machine learning algorithms such as gradient boosting and deep learning in making predictions. Random forest model was the most accurate in making predictions in acute and chronic CVD conditions both.

Table 17: Comparison of Chronic CVD Prediction Models

Model	c-statistic
Standard regression	0.7641
GLM net	0.7349
Random Forest	0.9872
Gradient Boosting	0.7608
Deep Learning	0.7662

Aim 1c: To validate the predictive risk models using cross validation techniques and evaluating the model fit on a varied sample

As random forest model was the one identified to be the most accurate on comparison with other models for chronic as well as acute CVD risk prediction, it was further validated. The chronic and acute CVD event models were validated internally and externally using 10-fold cross validation techniques. For internal validation, the data for both the models was split based on the census regions. South and west regions were used to train the models whereas north and mid regions were used to test the model. These regions were split based on the prevalence of CVD in these regions so they can be compared without any bias involved. A sample of 27,530 records identified in the south and the west regions were used to train the models whereas a sample of 17,181 records in the north and the mid regions were used to test the model. On training the acute CVD events model using south and west regions, following learning curve in figure 13 was observed where a model with 7 predictors like the one obtained above was identified as the best model. This model was then tested on north and mid regions to generate the following accuracy metrics reported in table 18.

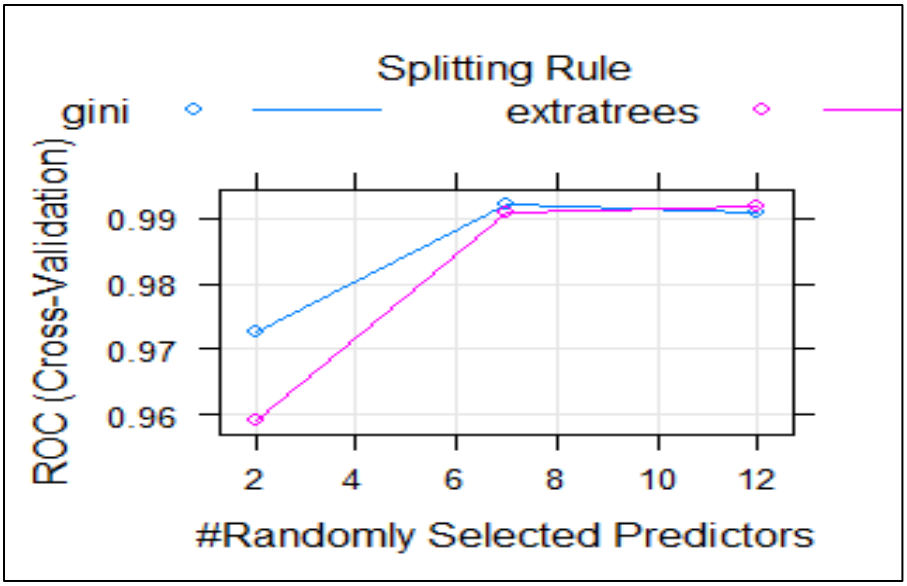


Figure 13: Learning Curves for Internal Validation of Random Forest Model using South and West Census Regions for Acute CVD Events Prediction

Table 18: Accuracy Metrics for Internal Validation of Acute CVD Events Prediction Model

Metric	Value
c-statistic	0.6808
Accuracy	0.7715
Sensitivity	0.8597
Specificity	0.4850

A c-statistic of 0.6808 suggests that the model trained using certain census regions does not perform extremely well on other census regions. This could be due to the innate differences in the demographics of these regions, food patterns etc. However, the accuracy rate as high as 77.15% and the c-statistic close to 0.7, suggests that the model outcomes are not completely random.

Similarly, for chronic CVD prediction model, the following learning curves as obtained in figure 14 were obtained on training the model on south and west regions. A

model with 7 predictors was identified as the best model. This model was then tested on north and mid regions to generate the following accuracy metrics reported in table 19.

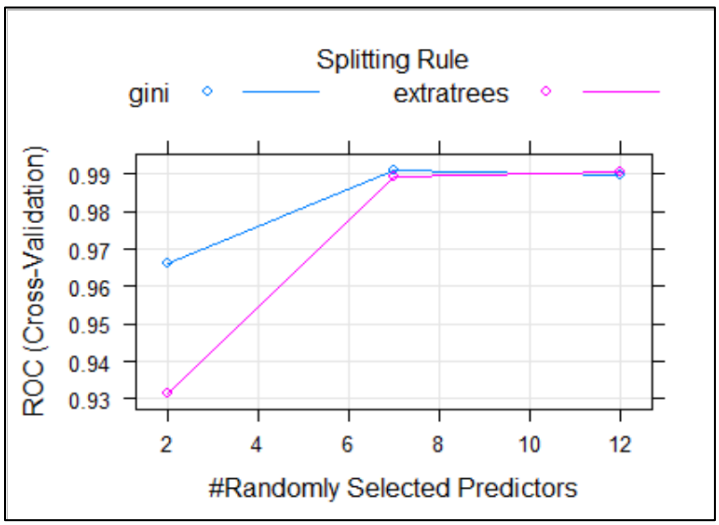


Figure 14: Learning Curves for Internal Validation of Random Forest Model using South and West Census Regions for Chronic CVD Prediction

Table 19: Accuracy Metrics for Internal Validation of Chronic CVD Events Prediction Model

Metric	Value
c-statistic	0.5488
Accuracy	0.6585
Sensitivity	0.8235
Specificity	0.2742

A c-statistic of 0.5488 suggests that the model trained using certain census regions does not perform well on other census regions. Thus on internal validation it was observed that the c-statistic obtained with acute and chronic CVD models both was low. The chronic CVD prediction model was further less accurate than the one built for predicting acute CVD events. This could be due to the innate differences in the demographics of these regions, food patterns etc. that could not be completely controlled for in our model due to data limitations. Medical practice patterns across regions also differ which could also lead

to increased screening and thus increased diagnosis in certain regions due to increased access.[27,28] This suggests that if data was available accounting for some of these factors in our predictive models, it could increase the generalizability and validity of the model. The accuracy rate was also 0.6585, which was lower than that obtained for acute CVD events and suggested that the outcome was predicted accurately 65% of the time based on the current data.

For external validation, the originally trained chronic and acute CVD prediction models using 2005-2015 training dataset were used to test on 2016-2017 data. These originally trained models were then used to predict the risk of acute and chronic CVD events for a sample of 21,674 events observed in 2016-2017. On performing the validation on 2016-2017 following test metrics were obtained as mentioned in table 20 for acute whereas table 21 for chronic CVD event prediction respectively.

Table 20: Accuracy Metrics for External Validation of Acute CVD Event Prediction Model

Metric	Value
c-statistic	0.7114
Accuracy	0.7830
Sensitivity	0.9373
Specificity	0.2856

As observed in table 20 above, a c-statistic of 0.7114 suggests that the original model trained using 2005—2015 training data performs well in predicting a future external 2016-2017 data. The accuracy of 78.30% suggests that the model predicts the new outcomes based on the currently trained model well. This suggests that our predictive model is quite robust in predicting future outcomes.

Table 21: Accuracy Metrics for External Validation of Chronic CVD Event Prediction Model

Metric	Value
c-statistic	0.5795
Accuracy	0.8526
Sensitivity	0.8540
Specificity	0.3050

As observed in table 21 above, a c-statistic of 0.5795 suggests that the original model trained using 2005—2015 training data does not perform well in predicting future external 2016-2017 data. The accuracy was 85.26% which suggests that the model predicted 85.26% of the outcomes accurately based on the current data . This suggests that our predictive model is somewhat robust in predicting future outcomes. This also suggests that there could have been some overfitting involved with the random forest model since it performs extremely well on the training sample although not quite accurate on validating externally. The chronic CVD random forest model was thus more prone to overfitting as compared to the acute CVD event random forest model.

The heat maps depicted in figures 15 and 16 below summarize all the prediction metrics that we compared for the machine learning and regression models. Figure 15 summarizes the metrics for acute CVD events whereas figure 16 summarizes the same for chronic CVD. The figures are color coded as per the value for the metric with green indicating a very high value which indicates a better model, followed by yellow for a moderate value indicating a moderate metric and red indicating a relatively poor value. As seen from these values, most of these models performed poor on their specificity values. This suggests that these models along with identifying those with a positive result accurately would also give a lot of false positives. Literature suggests that usually tests with a high sensitivity also have a high specificity value associated.[29] From a clinical

standpoint identifying more false positives would however be better than identifying someone with a disease as false negative. This would help in taking the necessary precaution irrespective rather than neglecting care.

Predictive Models (Acute)	c-statistic	Accuracy	Sensitivity	Specificity
Standard regression	0.7534	0.7928	0.9251	0.3357
GLM Net	0.7853	0.8028	0.9846	0.2447
Random Forest	0.9738	0.9754	0.9996	0.9771
Gradient Boosting	0.7833	0.8065	0.928	0.3448
Deep Learning	0.8267	0.8345	0.9345	0.3444
RF* Internal validation	0.6808	0.7715	0.8597	0.485
RF External validation	0.7114	0.783	0.9373	0.2856

*RF – Random Forest

Figure 15: Heat Map Depicting Prediction Metrics For Acute CVD Events Models

Predictive Models (Chronic)	c-statistic	Accuracy	Sensitivity	Specificity
Standard regression	0.7641	0.8098	0.9089	0.556
GLM Net	0.7349	0.7787	0.9473	0.5552
Random Forest	0.9872	0.9888	0.9914	0.9452
Gradient Boosting	0.7608	0.7713	0.9808	0.3075
Deep Learning	0.7662	0.7753	0.9878	0.3122
RF* Internal validation	0.5488	0.6585	0.8235	0.2742
RF External validation	0.5795	0.8526	0.854	0.305

*RF – Random Forest

Figure 16: Heat Map Depicting Prediction Metrics For Chronic CVD Models

Aim 1d: To create an interactive web-based application using the R-shiny to predict the risk of CVD among cancer patients using the most accurate model identified

The most accurate model as identified above was the random forest model. This model was then used to build a dynamic web-based application using the most significant predictors summarized in tables 12 and 13 by the random forest model for acute and chronic CVD models. This application can then be used by the physicians to evaluate the risk of CVD in cancer patients given the information about the predictors that are in the application. The most important predictors like those identified above for a chronic CVD condition and those identified for an acute CVD condition were used to build the web-based application. Two separate web-based applications were built for predicting the probability of a chronic and acute CVD event among cancer patients. One for predicting the risk of acute CVD whereas another to predict the risk of chronic CVD events. In addition to those identified in tables 12 and 13, the web-based application also included some basic sociodemographic characteristics that the physician might want to ask their patients in any case.

Following are the web-based applications built using the above-mentioned predictors on the R Shiny app as shown in figures. These are dynamic and predict the probability of acute/ chronic CVD given the combination of input values for the predictors.

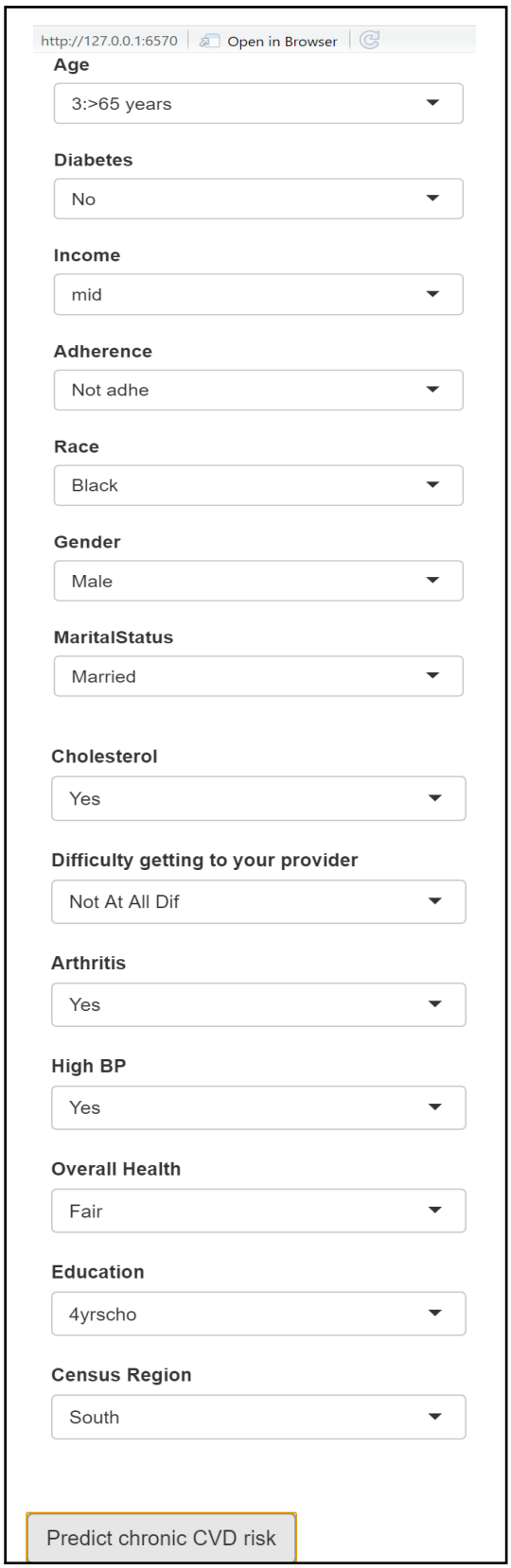


Figure 17: Web-Based Application to Predict the Risk of Chronic CVD In Cancer Patients

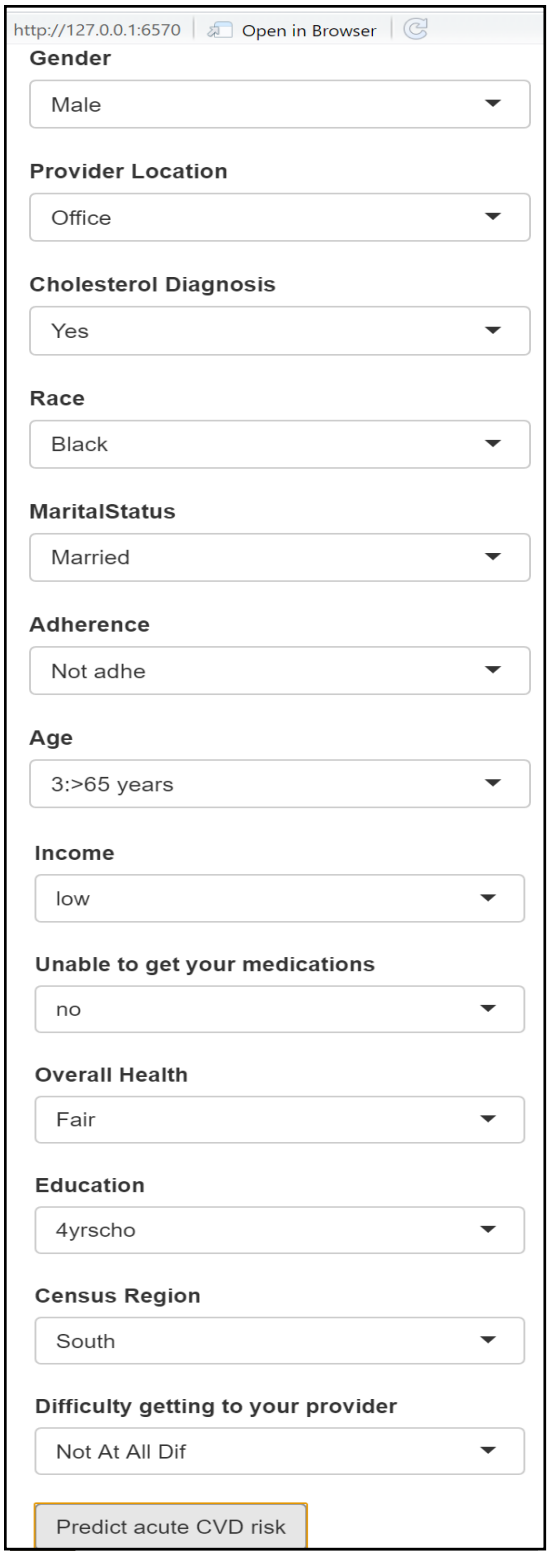


Figure 18: Web-Based Application to Predict the Risk of Acute CVD In Cancer Patients

DISCUSSION:

We characterized our study population that were cancer patients with and without CVD based on sociodemographic characteristics. We found that age, marital status, education and employment status differed significantly across cancer patients with and without CVD. However, we suspected that given the wide range of confidence intervals, this could have been due to a low sample size in specific groups as compared to reference groups. As observed from our study, the machine learning algorithms were more accurate in predicting the CVD risk in cancer patients as compared to the conventional stepwise regression method. It was observed that the c-statistic obtained was lower for the conventional regression methods for acute and chronic CVD prediction models both. The model built for predicting acute CVD events was more accurate as compared to chronic CVD prediction model. On conducting internal and external validation as well, it was observed that the chronic CVD prediction model was less accurate as compared to acute. Our validation techniques suggest that future work could be done on both the acute and chronic models to incorporate more predictors to make the model more generalizable and valid that can then increase the usefulness of prediction models. Adding some more predictors that can account for the regional differences might help in increasing the validity and usefulness of these models. Overall prediction was more accurate with the models built using machine learning algorithms as compared to those built using conventional stepwise regression approach.

Some previous studies have looked at application of machine learning algorithms to prediction of CVD risk as compared to the conventional approaches. A study conducted by Weng et al., found similar results to our study where the machine learning algorithms

improved prediction in comparison to already existing approaches.[4] Neural networks were the highest achieving algorithm in terms of accuracy. This study was however conducted in general population and was not restricted to cancer patients. In addition, it was also restricted to routine clinical data from family practices which limits the generalizability.

A study conducted by Goldstein et al. also looked at cardiovascular risk prediction using machine learning algorithms. This study also suggested that machine learning algorithms are more advantageous when it comes to generating a predictive model as compared to traditional regression approaches.[30] This was similar to the findings in our study. The prior study was however not specific to cancer patients.

A study conducted by Dranitsaris et al., similar to our study, developed a predictive model to estimate cardiotoxic risk in cancer patients. It was however restricted to breast cancer patients receiving anthracyclines thus limiting the generalizability. They used generalized estimating equations (GEE) model along with nonparametric bootstrapping to develop the predictive models. The c-statistic obtained was 0.84 which suggests that it was a good model. Similar to our results, this study suggested that machine learning models have good predictive capacity.[9] However this study did not compare the accuracy to multiple other machine learning algorithms or conventional regression models. Thus, they lacked a comparison to conventional approaches which used standard regression techniques.

However, there have been studies in the literature conducted by Christodoulou et al. and Gravesteijn et al. that have compared machine learning algorithms to regression models for clinical prediction.[31,32] These studies suggested that machine learning algorithms performed no better than regression models. However in our study, as

observed from the results for cancer population, the c-statistics obtained for ML models were higher compared to the standard regression models. Although, all other ML models except for random forest had a c-statistic close to each other and to the standard regression model. However, external validation for a random forest model suggested that specifically for a chronic CVD model, the model was not quite robust in predicting future outcomes. This suggests that there could be some overfitting involved with the chronic CVD random forest model which could be the reason for a very high c-statistic. This was also suggested in the study conducted by Gravesteijn et al. where the random forest model was more prone to overfitting. The authors also implied that prediction models need continuous updating and validation because their performance is often worse in newer cohorts which was also observed in our study.[32]

The major strength of our study is that it is one of the first studies to explore the application of machine learning approaches to predict the risk of CVD in cancer patients comparing that to the traditional regression approach. There have been some studies in the literature looking at CVD outcomes in cancer patients. However, these studies have mainly been from diagnostic purposes quantifying calcifications and other clinical outcomes obtained from X-rays and CT scans. These have used algorithms such as deep learning, neural networks and noisy-threshold classifier to predict outcomes and suggest that they are superior in comparison to standard regression approaches. Our study used a nationally representative sample which increases the generalizability of aim 1A results where adjusted analyses was conducted using complex survey weights. Since the models developed are predictive, they would help in managing the cardiovascular outcomes of cancer patients more efficiently. The web-based application built to dynamically predict

the risk of CVD in cancer patients given all the predictors would help the physicians in tailoring the interventions based on the results.

However, the study also has certain limitations. Firstly, the models were trained on cross-sectional data. Certain time dependent variables, that would have been available if it were a longitudinal dataset could not be controlled. Secondly, since the dataset was cross-sectional, certain lifestyle characteristics were assumed to be the same at the time of cancer diagnosis even if the patients were now in remission. Thirdly, the risk factors for certain CVD conditions differ innately which was why separate models were built for chronic and acute CVD events. There could also be some differences in the risk factors involved in myocardial infarction, angina, coronary artery disease and other heart diseases which were all clubbed as 'Chronic' CVD events due to sample size limitations and ease of interpretability. In addition, certain variables such as specific cancer therapy used, and cancer type had high missing values. Cancer therapy could be broadly classified into chemotherapy, surgery and radiotherapy used which was then used in the study along with the specific cancer therapy variable. The missing values for cancer type were recoded using KNN imputation, however the percentage of missing values could lead to some bias in the study. Finally, MEPS data being self-reported could lead to some recall and selection bias. However, this would be minimized since AHRQ also verifies the patient medication and medical conditions reports from their physicians and their pharmacists.[12] There could also be some coding errors involved if the providers did not code the ICD-9 codes for patient diagnoses accurately. In addition, we also made an assumption that all the patients with comorbidities and certain CVD risk factors such as high blood pressure and cholesterol have accurate diagnosis in the datasets. It is possible that some might have

these diagnoses without having an actual code in the file, however these would be misrepresented in our analyses as 'No' or 'Missing'.

Despite the limitations stated above, the strengths and the novelty involved in the approach makes it a strong study and adds a lot to the literature. Our study results suggest that the use of novel techniques such as machine learning might be more beneficial than the conventional approaches in predicting outcomes. The model accuracy results might help in guiding certain real-world evidence analysis approaches especially in this era when the field is moving more towards data innovation. Our internal and external validation models suggested that there might be overfitting involved in the random forest chronic and acute CVD models, although future model calibration can increase the usefulness and validity of these models. Even though our models did not perform the best on conducting validation, they did suggest that they had more predictive power than the standard regression techniques (higher c-statistic). Our study is a good starting point to suggest the use of machine learning algorithms in evaluating healthcare outcomes as more data becomes available. With more data and long-term outcomes available, the validity of our models can be increased to be made more useful for real-world application. The predictive power of these models can be used to a great extent in planning future treatment plans for cancer patients given their cardiovascular risks. The web-based application that we created might help the physicians in giving a real time estimate of the cardiovascular risk cancer patients. Based on the predicted cardiovascular risk, the physicians can monitor the cardiovascular health of cancer patients more efficiently. It can also guide certain developers to create more web-based and mobile applications to make healthcare more efficient. Down the line, it might help in managing the CVD condition more efficiently and

have better health and cost implications for cancer patients. With more data and calibration these models can be made more robust for use in the real world.

APPENDIX:

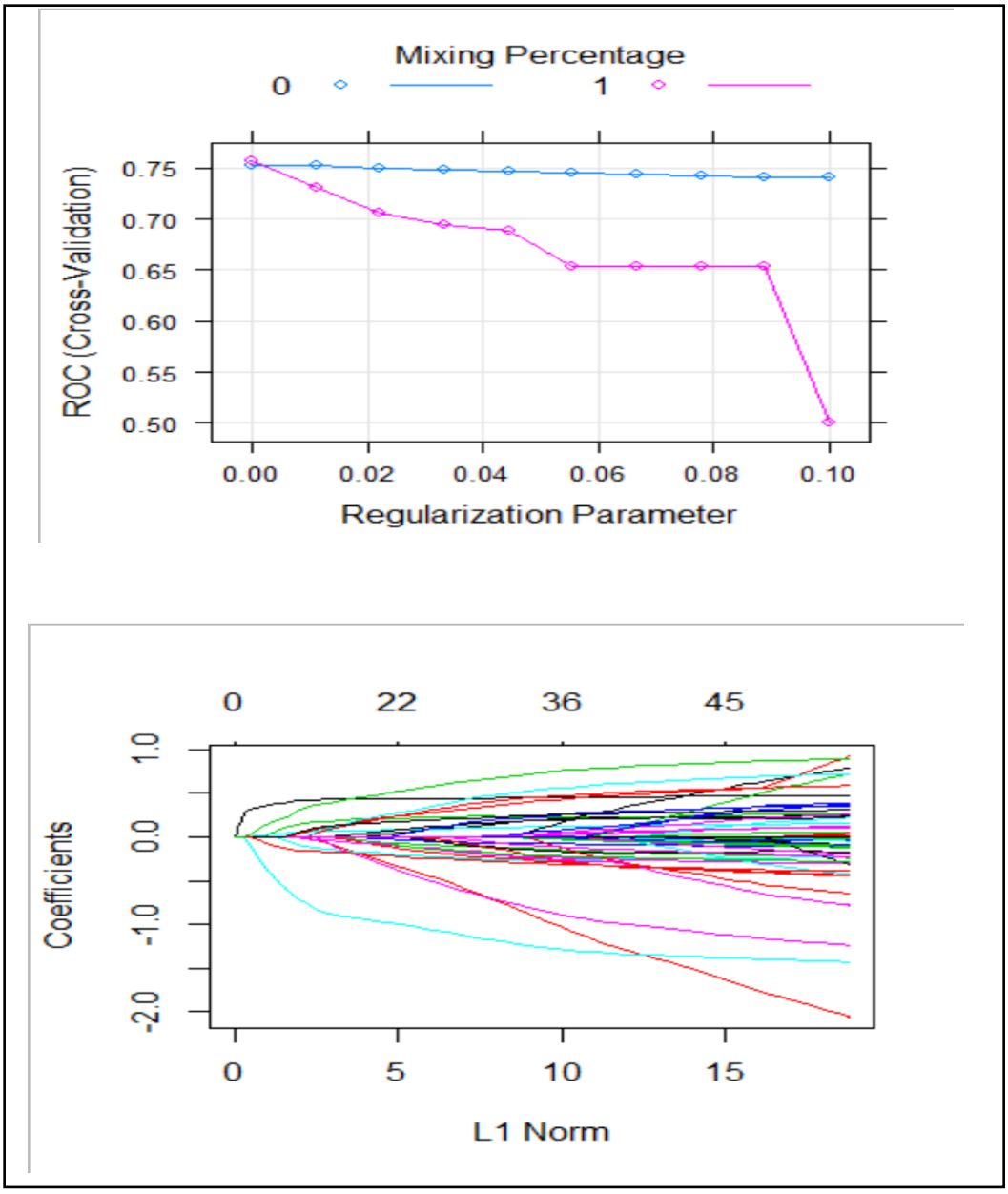


Figure 1: Regularization pathway for acute CVD GLMNET model

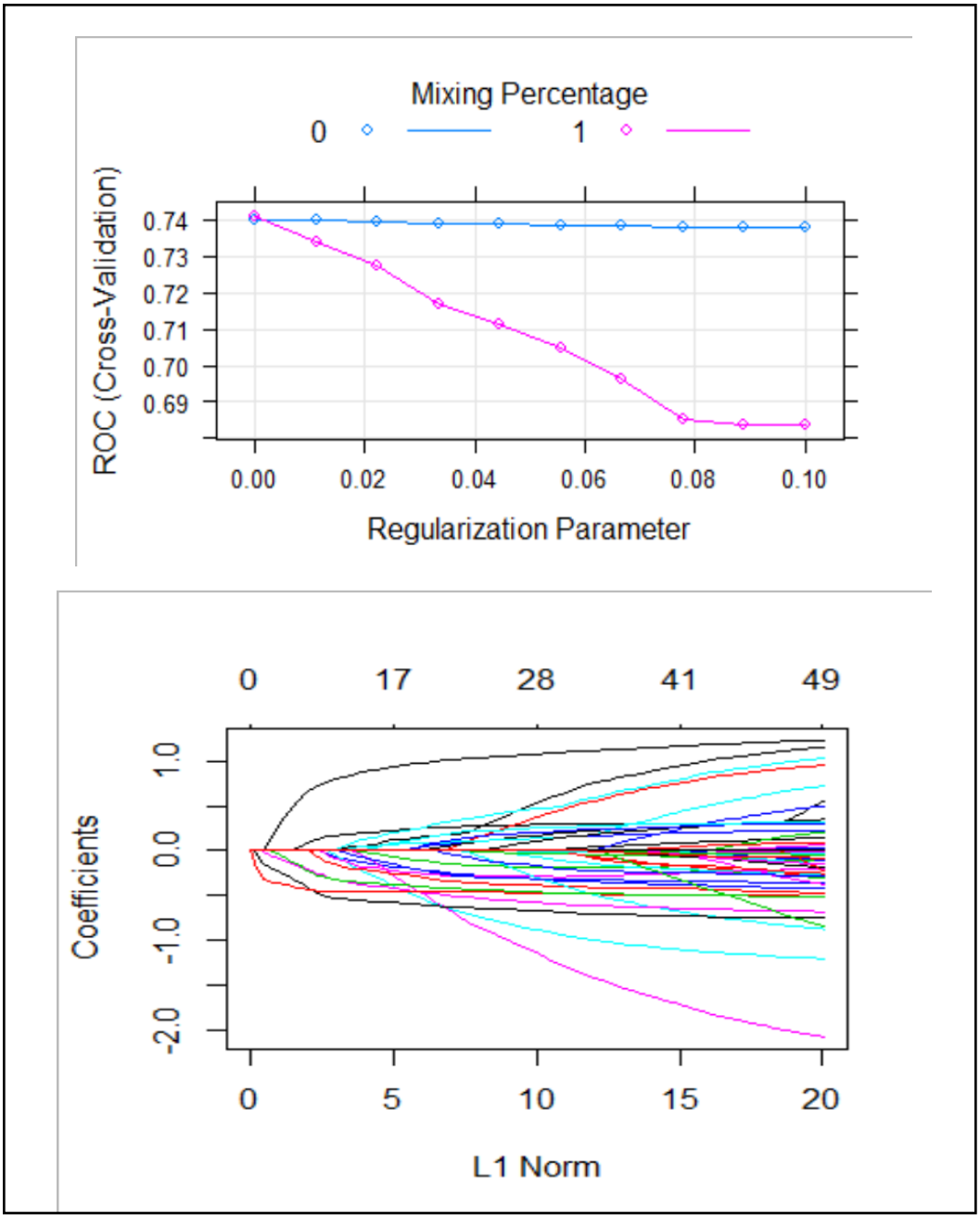


Figure 2: Regularization pathway for chronic CVD GLMNET model

Thus, as seen from the figures above, the mixing percentages of 0 and 1 indicate the alpha values whereas the X-axis (regularization parameter) indicates the lambda values. The figures above suggest that for both the acute and chronic CVD models α = 0 performs better than α = 1 since the ROC values obtained for the former are constantly higher than those obtained for the later. This indicates that a ridge regression would be

better for both acute and chronic CVD prediction models as compared to a lasso regression. The regularization path obtained for both the models suggest that, as we increase the penalty on the model (going from right to left) and decrease the complexity of the model, most of the regression coefficients tend to move towards 0 and being nonsignificant. Thus, a simpler model would probably be more accurate in making predictions as compared to a more complex model.

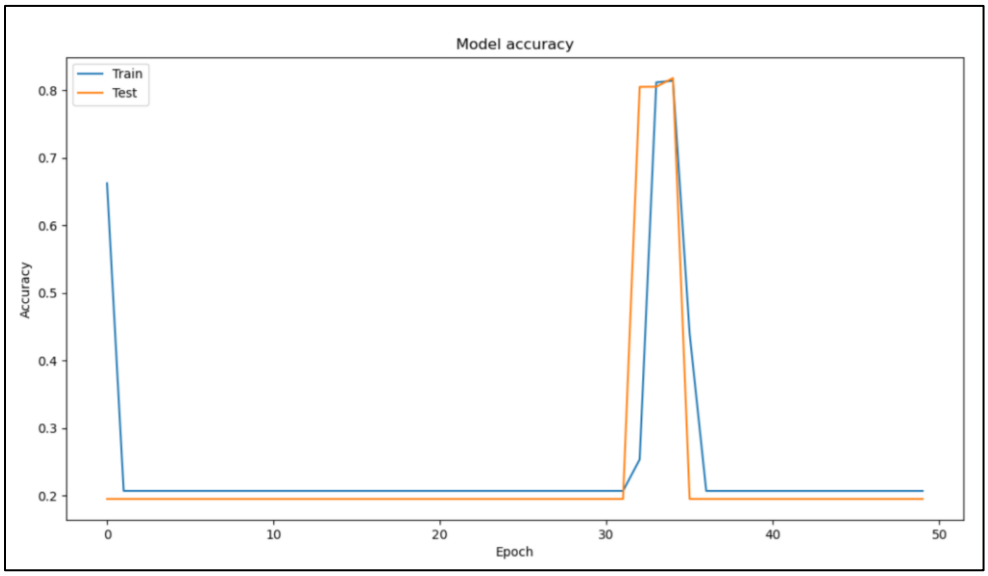


Figure 3: Batch normalization deep learning curves for acute CVD prediction model

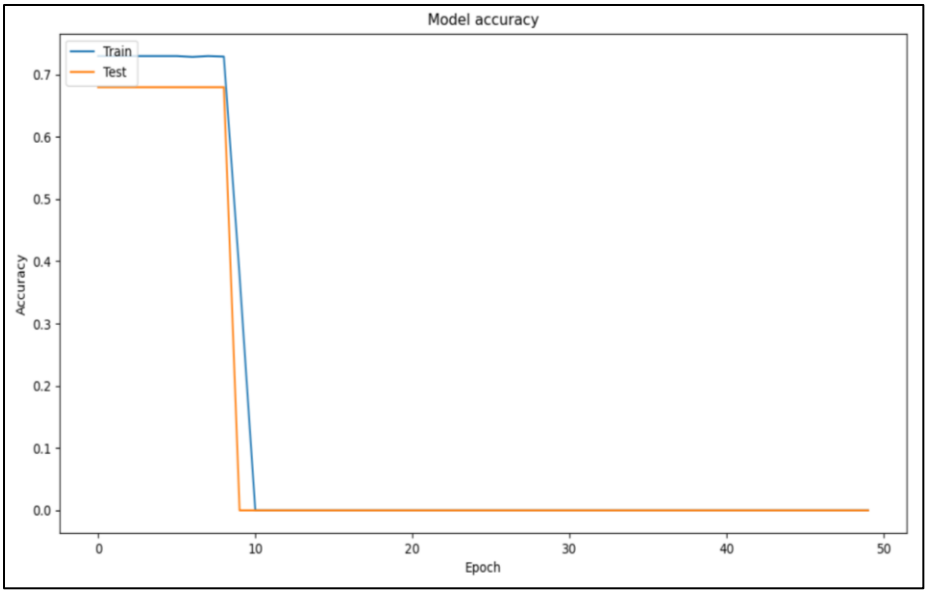


Figure 4: Batch normalization deep learning curves for chronic CVD prediction model

REFERENCES:

1. Royer, J. et al. PNS196 Use Of Machine Learning In Health Economics And Outcomes Research (Heor): Optimizing Predictive Models. 2019 *Value in Health*, Volume 22, S319

2. Rahimian F, Salimi-khorshidi G, Payberah AH, et al. Predicting the risk of emergency admission with machine learning : Development and validation using linked electronic health records. 2018;779:1-18.

3. Chen J, Hoogh K De, Gulliver J, et al. A comparison of linear regression , regularization , and machine learning algorithms to develop Europe-wide spatial models of fi ne particles and nitrogen dioxide. 2019;130(February). doi:10.1016/j.envint.2019.104934

4. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data ? 2017;355:1-14.

5. Gernaat SAM, Velzen SGM Van, Koh V, et al. Automatic quantification of calcifications in the coronary arteries and thoracic aorta on radiotherapy planning CT scans of Western and Asian breast cancer patients. *Radiother Oncol*. 2020;127(3):487-492. doi:10.1016/j.radonc.2018.04.011

6. Lessmann N, Ginneken B Van, Zreik M, et al. Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks With Dilated Convolutions. 2018;37(2):615-625. doi:10.1109/TMI.2017.2769839

7. Hu SB, Wong DJL, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural

Network Model. 2016:1-12. doi:10.1371/journal.pone.0161401

8. Takx RAP, Jong PA De, Leiner T, Oudkerk M, Koning HJ De. Automated Coronary Artery Calcification Scoring in Non- Gated Chest CT : Agreement and Reliability. 2014;9(3). doi:10.1371/journal.pone.0091239

9. Dranitsaris G, Rayson ÆD, Vincent ÆM, et al. The development of a predictive model to estimate cardiotoxic risk for patients with metastatic breast cancer receiving anthracyclines. 2008:443-450. doi:10.1007/s10549-007-9803-5

10. Gerven MAJ Van, Jurgelenaite R, Taal BG, Heskes T, Lucas PJF. Predicting carcinoid heart disease with the noisy-threshold classifier. 2007. doi:10.1016/j.artmed.2006.09.003

11. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. Social Determinants of Health Discussion Paper 2 (Policy and Practice)

12. Medical Expenditure Panel Survey (MEPS) – AHRQ https://www.meps.ahrq.gov/mepsweb/ accessed on: 08/23/2020

13. DataCamp: Machine Learning Toolkit course

14. Stromberg KA, Agyemang AA, Graham KM, et al. Using Decision Tree Methodology to Predict Employment After Moderate to Severe Traumatic Brain Injury [published correction appears in J Head Trauma Rehabil. 2019 Nov/Dec;34(6):401]. *J Head Trauma Rehabil*. 2019;34(3):E64-E74. doi:10.1097/HTR.0000000000000438

15. DataCamp: Machine Learning with caret in R

16. DataCamp: XGBoost Package

17. DataCamp: Introduction to Deep Learning with KERAS

18. DataCamp: Advanced Deep Learning with KERAS

19. Machine Learning Mastery: A gentle introduction to batch normalization for deep neural networks https://machinelearningmastery.com/batch-normalization-for-training-of-deep-neural-networks/ accessed on: 08/31/2020

20. Machine Learning Mastery: How to accelerate learning of deep neural networks with batch normalization https://machinelearningmastery.com/how-to-accelerate-learning-of-deep-neural-networks-with-batch-normalization/ accessed on: 08/31/2020

21. GLMNET Vignette https://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html accessed on : 08/30/2020

22. Glen S. C-Statistic: Definition, Examples, Weighting and Significance. https://www.statisticshowto.datasciencecentral.com/c-statistic/. Published 2016.

23. The use of KNN for missing values https://towardsdatascience.com/the-use-of-knn-for-missing-values-cf33d935c637 accessed on: 07/23/2020

24. Centers for Disease Control and Prevention (CDC). Prevalence of coronary heart disease--United States, 2006-2010. MMWR Morb Mortal Wkly Rep. 2011 Oct 14;60(40):1377-81. PMID: 21993341.

25. How to build a Shiny App – Shiny from Rstudio https://shiny.rstudio.com/articles/build.html accessed on : 08/24/2020

26. Codementor - Building web data products with R and Shiny https://www.codementor.io/@jadianes/building-a-web-based-sentiment-classifier-using-r-shiny-du107oagb accessed on: 08/24/2020

27. Song Y et al. Regional variations in diagnostic practices. *N Engl J Med* 2010 Jul 1; 363:45. http://dx.doi.org/10.1056/NEJMsa0910881

28. Zuckerman S et al. Clarifying sources of geographic differences in Medicare spending. *N Engl J Med* 2010 Jul 1; 363:54. http://dx.doi.org/10.1056/NEJMsa0909253

29. Understanding Medical Tests: Sensitivity, Specificity and Positive Predictive Value https://www.healthnewsreview.org/toolkit/tips-for-understanding-studies/understanding-medical-tests-sensitivity-specificity-and-positive-predictive-value/ accessed on: 11/04/2020

30. BA Goldstein, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *European Heart Journal*, Volume 38, Issue 23, 14 June 2017, Pages 1805-1814, https://doi.org/10.1093/eurheartj/ehw302

31. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12-22. doi:10.1016/j.jclinepi.2019.02.004

32. Gravesteijn BY, Nieboer D, Ercole A, et al. Machine learning algorithms performed no better than regression models for prognostication in traumatic brain injury. *J Clin Epidemiol*. 2020;122:95-107. doi:10.1016/j.jclinepi.2020.03.005

CHAPTER 3: BENEFIT-RISK ASSESSMENT OF BREAST CANCER THERAPIES USING MCDA MODEL

BACKGROUND:

Patients with cancer have an increased burden and a reduced quality of life majorly due to the associated comorbidities.[1] Cardiovascular diseases are one of the major comorbidities associated with cancer patients. The prevalence of cardiovascular diseases (CVD) in cancer patients is growing making CVD one of the most associated comorbidity in cancer patients.[2] In cancer patients, the CVD mortality rate has increased by 20-30% in recent years, whereas the cancer mortality rate has decreased by 20-30%.[3] With growing concerns of CVD in cancer patients, it has become more important to consider cancer treatment related factors that put patients at an even higher risk. There are multiple factors that put cancer patients at an even higher risk for developing CVD that include, patient demographics, lifestyle factors and cancer treatment related factors.[4,5] With increasing availability of cancer therapies, it would be important to focus on cardiovascular implications of therapies for planning the patient's treatment plan more efficiently. Certain cancer treatments have been established to be more cardiotoxic than the others. Cardiac complications are specifically higher if patients are receiving anthracyclines, radiotherapy or certain targeted therapies.[6,7] These targeted therapies usually form the first line treatment regimens for cancer patients. These cardiotoxic effects of therapies can be seen even after years of being diagnosed with cancer or after remission.[2] These targeted therapies are however also associated with an higher overall survival in cancer patients.[8] It is thus necessary to assess the cardiotoxic and other adverse events profiles of such therapy regimens including targeted therapies to quantify the trade-off between adverse

events and survival. Assessing these outcomes might help the physicians in planning the treatment better.

Evaluating specific characteristics associated with cardiotoxic adverse events would help in managing/ preventing the condition and tailoring cancer therapies appropriately. There have been some studies looking at cardiotoxic potential of specific drugs,[9-11] however, characteristics associated with these therapies have not been described yet. Most of these drugs that have a high cardiotoxic potential are used in breast cancer patients.[9-11] Focusing on breast cancer patients would thus be more informative and useful. Identifying combinations of drugs and therapy characteristics used in breast cancer treatment could help in understanding if there is a synergistic effect of multiple factors involved. If certain therapy related characteristics such as dosage or route of administration turn out to be associated with the increased potential of cardiotoxicity, the results can inform the physicians to be cautious respectively. The targeted therapies used in breast cancer treatment are also associated with a higher survival among cancer patients.[10] Efficacy and tolerability which could be measured in terms of survival and adverse events are the two important considerations by physicians for treatment choices.[12] Given that these therapies have higher risks and higher survival both there would be an uncertainty in decision-making while prescribing. Consolidating these benefits and risks together in a model might be beneficial to make the decision-making easier. Conducting a benefit-risk assessment of these therapy regimens used in HER 2 positive patients would help in consolidating and quantifying the benefits (survival) and risks (cardiac complications along with other adverse events) outcomes in a single model. Breast cancer being one of the most prevalent cancer with cardiotoxic potential, a benefit-risk

assessment of therapies involved in breast cancer patients would immensely help physicians in decision-making. This can be achieved using a Multiple Criteria Decision Analysis (MCDA) model to attempt to reduce the uncertainty involved in decision-making. This model can be used to evaluate the trade-off between the therapy regimens that are usually used in HER 2 positive breast cancer patients by consolidating the benefits and the risks criteria.

A MCDA model can be built using multiple perspectives namely, benefit-risk assessment, health technology assessment (HTA), portfolio decision analysis (PDA), commissioning decisions, shared decision making (SDM), and prioritizing patient's access to health care. [13] HTA bodies use MCDA to make coverage decisions, PDA is conducted by scientific companies to choose the criteria where best to direct R&D efforts, commissioning decisions are mainly used to assess resource allocation whereas SDM and patient's access models incorporate criteria that would be more important and subjective to patients such as quality of life, treatment satisfaction, etc.[13] However, given that there is a survival-adverse events trade-off involved in the therapy alternatives, it would be the most appropriate to build a MCDA model from a benefit-risk assessment perspective. MCDA modeling technique from this perspective is a way of incorporating benefits and risks and evaluating alternative treatment options at once by including the therapy regimens. The decision-making process can be made more transparent by describing the risks and benefits trade-off in a formal manner. MCDA provides a framework for systematic and replicable analyses of complex decision problems involving value trade-offs. [13] Survival and cardiac implications of targeted therapies have not been studied as a value trade-off in combination with other conventional therapies as a part of standard

regimens. Most of the studies in the current literature have only looked at one targeted therapy at a time and its effects on population outcomes.[6-11] These outcomes were also studied separately in separate studies and these have not been comparative across various therapy regimens. [6-11] Just looking at one targeted therapy at a time from a clinical utility/ decision-making standpoint would not be sufficient since in the real world these are usually given in combination with other therapies. A comparison between therapy regimens that include these specific drugs would thus be a fair comparison for MCDA model. Targeted therapies such as trastuzumab and pertuzumab although improve survival in HER 2 positive breast cancer patients, they also increase the risk of adverse events. [6,9] For HER 2 positive breast cancer patients NCCN guidelines enlist certain first-line therapy regimens such as trastuzumab in combination with a taxane, trastuzumab in combination with a pertuzumab and a taxane and trastuzumab in combination with cyclophosphamide/ carboplatin and a taxane.[14] These therapy regimens can be used as model alternatives while building a MCDA model. There is also enough evidence on these regimens like summarized in Table 25 below to conduct a benefit-risk assessment using MCDA model. A MCDA model incorporating the benefits and risks of therapy regimens used in HER 2 positive breast cancer patients might help in providing evidence to make the decision-making process more transparent. There is a high demand for transparency in healthcare decision making with the availability of growing and emerging options and the fields becoming more multidisciplinary. This demand for transparent decision processes can be fulfilled by a systematic construct of benefit-risk assessment. [15] A MCDA model helps in structuring various outcomes/ preferences together and systematically integrating these into a decision-making process. By assigning scores and weights to

different measures considered in the MCDA model and reporting these systematically in the model structure it can formalize the decision-making process. In the field of oncology which is highly multidisciplinary use of an MCDA model can help immensely in decision-making. [16] It would help in choosing from the three standard therapy regimens in HER2 positive breast cancer patients like mentioned above. The objective of our study was to compare these therapy regimens with respect to their benefits and risks criteria that would be mentioned in the MCDA model below.

Literature Review:

We conducted a literature review to identify studies that have looked at the application of a decision-making framework like multiple criteria decision analysis to conduct a benefit risk assessment of cancer therapy regimes. We also looked at studies that evaluated patient preferences in choosing a cancer treatment. The following search strategy using a combination of MeSH terms was used: (((((("Molecular Targeted Therapy"[Majr])) OR "Drug Therapy"[Mesh:NoExp])) AND (((benefit risk) OR risk assessment) OR multiple criteria decision analysis)) AND "Neoplasms"[Majr])).The titles and abstracts were then screened for their eligibility using the following inclusion/exclusion criteria.

We included studies that looked at a risk-benefit trade off associated with a cancer therapy and considered cardiotoxicity as at least one of the risk factors under consideration. We also restricted our search to studies conducted in humans and published in English. We excluded studies that focused just on risks or just on benefits. We also excluded studies that did not look at cardiotoxic risk. We excluded studies that

were conducted in pediatric population or were narrative reviews and did not look at any outcomes.

The search strategy resulted into 97 studies. On applying the inclusion/exclusion criteria, 94 articles were excluded. The final literature review included three studies.[17 - 19] Figure 19 below includes a PRISMA flowchart of studies that were included in the literature review.

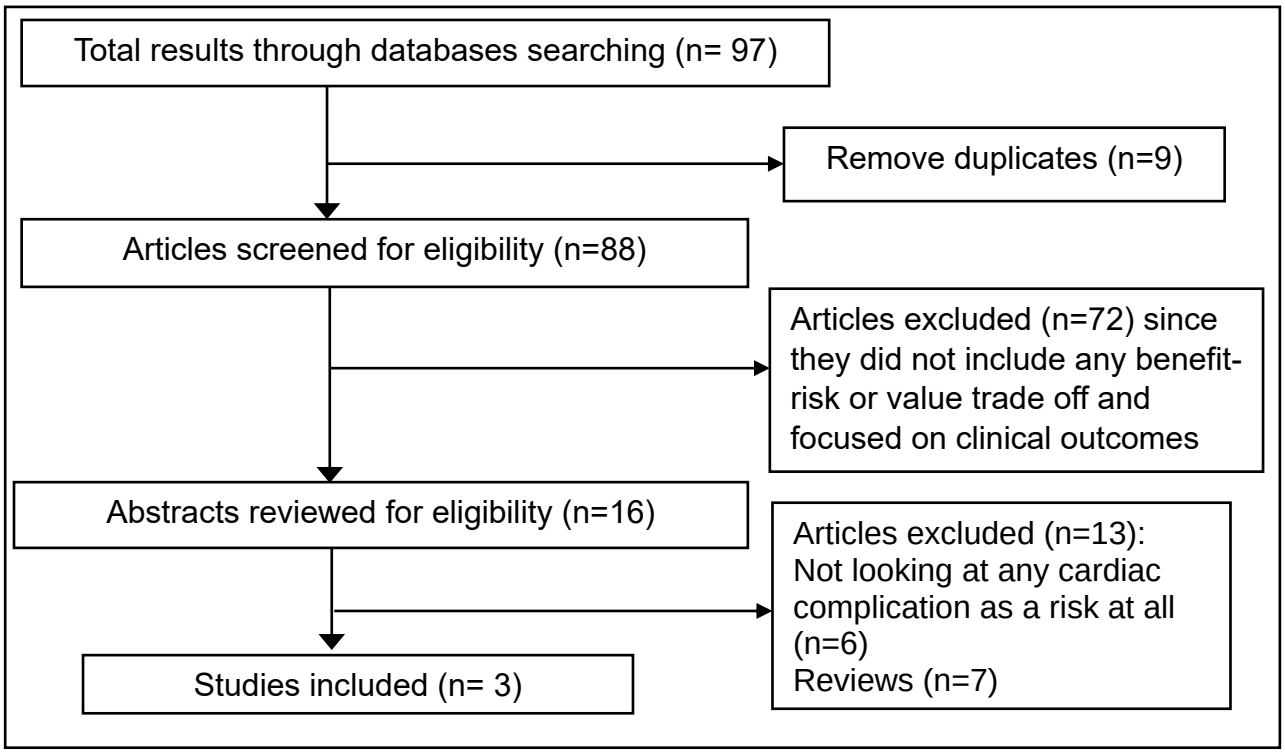


Figure 19: PRISMA Flowchart of Literature Review (Aim 2)

Table 22 below summarizes the studies that were included in the literature review. There were only 3 studies that looked at a benefit-risk trade off associated with cancer therapies. Majority of the studies in literature are narrative reviews that define the steps involved in MCDA. There are very few studies that have conducted a structured decision-making process to quantify the benefits and risks associated with cancer therapies.

Table 22: Summary Of Literature Review (Aim 2)

Study	Study Objective
Postmus *et al.,* 2018[17]	To elicit the preferences of patients with multiple myeloma regarding the possible benefits and risks of cancer treatments
Wagner et al., 2018[18]	To apply MCDA shared-decision framework to explore what matters to patients in considering the treatment options
Lifford et al., 2015[19]	To understand older women's decision making and coping in context of breast cancer treatment

There are only three studies that have looked at application of decision making models to explore patient preferences in choosing the treatment alternatives in cancer patients. In a study conducted by Postmus et al. it was suggested that progression-free survival (PFS) was weighted higher (0.54) than severe life-threatening toxicities (0.32) for patients with multiple myeloma. This preference for PFS was irrespective of other factors included in the model. This study was however restricted to a small sample size and did not compare different cancer treatment options. The study focused on patient preferences for PFS over adverse events and did not compare these risks and benefits across treatment alternatives.[17]

Wagner et al. also applied the MCDA framework in eliciting patient preferences. The EVIDEM-derived MCDA framework was used in the study where five patients and six physicians assigned criteria weights. This study was thus conducted from a shared decision-making perspective. The participants individually weighted the relative importance of the criteria on the basis of what mattered the most and least to them when making a decision on the cancer management options. Similar results were obtained in

this study where patients preferred treatment over watchful waiting (weight 0.32 vs 0.24) with the largest contribution from PFS (weight = 0.11) over fatal adverse events (weight = 0.06) and impact on health-related quality of life (weight = 0.04). [18] This study also focused on patient preferences for criteria rather than ranking therapy alternatives.

A study conducted by Lifford et al. looked at decision-making process of older breast cancer women in coping with cancer treatment.[19] Semi-structured interviews were carried out with older women to assess women's information and support needs, their breast cancer diagnosis and treatment decisions. The authors found that past experience of cancer and its treatment, benefits and the risks associated with these treatments were all ranked important from a patient's perspective. Women also described various strategies to cope with breast cancer and their treatment decisions. These included seeking information, obtaining practical and emotional support from healthcare professionals, friends and relatives, and relying on personal faith.[19] Like the previous studies, this study was also mainly conducted to elicit patient preferences in making a treatment decision.

As seen from these studies, they were from a patient's decision-making perspective and there is no clear consensus on choosing a specific therapy regimen given the benefit-risk trade off. These were conducted mainly to elicit the factors that patients might consider important while making a decision about a therapy alternative.

Gaps in the literature:

As seen from the above literature, there is a lack of evidence of studies that incorporate the benefits and risks of cancer therapies together in a model to assign values and make the decision process more transparent. In cancer care, multidisciplinary teams have to work together to create patient's treatment plan which makes decision-making

even more difficult, especially if it is carried out in an informal non-transparent manner. This makes the need for a more transparent decision-making process like MCDA even greater in oncology.[16] Currently for HER 2 positive breast cancer patients, the physicians usually prescribe trastuzumab with an additional chemotherapeutic agent that is tailored as per the patient needs with the aim of prolonging survival. [20] It mainly depends on clinical characteristics such as patient's tumor type, size and stage of cancer. [21] With reports of adverse events growing, taking survival and adverse events both into consideration while treatment planning becomes equally important. However, there are no standardized decision-making guidelines that can formalize this process which may increase conflicts that are inherent to clinical decision making. With multidisciplinary teams involved, a guided approach to decision-making might help decision-makers, providers and patients in deliberation and communication.[16] Stating the criteria considered while making the decision explicitly and scoring and weighting these, might increase the transparency of the process.

Current studies that have used MCDA modeling techniques have only looked at patient preferences for the criteria (Table 22) and there are no studies that look at value assessment of cancer therapies from a benefit-risk perspective. Current studies are mainly conducted using a decompositional approach where criteria weights were derived later based on the therapy alternative that the patients preferred. [17-19] They were from a patient perspective to evaluate the criteria that the patients considered most important while making a decision and hence mainly only include patient reported criteria such as satisfaction with the treatment, quality of life, etc. This helps in ranking the criteria rather than ranking the therapy alternatives. We wanted to develop a model that can rank the

therapy alternatives given the inputs on the criteria that would be identified from the literature mainly from a physician decision-making perspective. Developing a MCDA model using a compositional approach would help in assigning a value to each breast cancer therapy regimen using the evidence in the literature for criteria inputs. This might then help the physicians and multidisciplinary teams in making a decision. Targeted therapies (trastuzumab and pertuzumab in HER 2 positive breast cancer patients) as suggested are associated with a higher survival as well as a higher cardiotoxic potential which makes choosing the therapy alternative difficult. Most of these targeted therapies associated with a higher survival and cardiotoxic profile are used in breast cancer patients.[6,9] Most of the clinical studies conducted looking at these outcomes have also been in breast cancer patients. Thus more evidence is available to conduct a benefit-risk assessment of breast cancer therapy regimens using a MCDA model. Our goal of study was thus to develop a Multi Criteria Decision Analysis Model (MCDA) to evaluate benefits and risks associated with breast cancer therapy regimens. Literature inputs were used to assign a value with each regimen and rank these to make decision-making easier.

Specific Aim 2:

To assess cardiotoxicity associated with targeted therapies as compared to non-targeted therapies and develop a model to conduct benefit-risk assessment of therapy regimens in breast cancer patients

d. To describe the cancer therapy characteristics associated with the cardiotoxic adverse events in breast cancer patients receiving targeted therapies as compared to those receiving non-targeted therapies and evaluate the drug-event association using a disproportionality analysis

e. To develop a Multi Criteria Decision Analysis (MCDA) model to conduct benefit-risk assessment of breast cancer therapy regimens
f. To conduct sensitivity analyses to assess the MCDA model performance and uncertainty in the model

METHODS AIM 2: TO ASSESS CARDIOTOXICITY ASSOCIATED WITH TARGETED THERAPIES AS COMPARED TO NON-TARGETED THERAPIES AND DEVELOP A MODEL TO CONDUCT BENEFIT-RISK ASSESSMENT OF THERAPY REGIMENS IN BREAST CANCER PATIENTS

Aim 2a: To describe the cancer therapy characteristics associated with the cardiotoxic adverse events in breast cancer patients receiving targeted therapies as compared to those receiving non-targeted therapies and evaluate the drug-event association using a disproportionality analysis

Data Source:

FDA Adverse Event Reporting System (FAERS) quarterly files from 2005 to 2015 were used for this study aim. These files were downloaded from the public dashboard as a zip file which were then extracted for analyses. FAERS is a voluntary, spontaneous reporting database that provides information on adverse event and medication error reports submitted to the U.S. FDA by healthcare professionals, consumers, and manufacturers worldwide.[22] FAERS quarterly files contain information on demographics and administrative information along with the initial report for the patients, drug and reaction information from the reported adverse events case reports and patient outcomes.[22] Files DRUGyyQq.txt and THERyyQq.txt were used to identify targeted and

non-targeted therapies and to identify the duration of these therapies. REACyyQq.txt files were used to identify cardiovascular adverse events. OUTCyyQq.txt files were used to identify the severity of the cardiovascular adverse event. Patient characteristics were obtained from the DEMOyyQq.txt files. These files were event-level files that were used to identify cardiotoxic and non-cardiotoxic reported adverse events in breast cancer patients.

Proposed study design and sample:

Cross sectional study design was used for this study aim. The study sample consisted of adverse events identified in breast cancer patients over the age of 18 years in the US from 2005-2015. INDIyyQq.txt files contain patients' disease information, which were used in our study to identify breast cancer patients. Medical Dictionary for Regulatory Activities (MedDRA) preferred terms (variable PT) were used to identify breast cancer patients. These MedDRA terms were coded as strings of words to identify any particular cancer type. The string 'breast cancer' was used to identify the study sample. The string search identified breast cancer patients and excluded those who just reported cancer pain without any specific diagnoses. Using the breast cancer string resulted into a sample size of 35,630,544 events identified in breast cancer patients from FAERS 2005-2015 files.

Study Variables:

Certain patient-related characteristics such as patient's age, gender, death date and the time of patient's visit the reaction was reported were identified from FAERS files. Cardiovascular adverse event related factors that were identified were the reporter's type of occupation, the adverse event date, preferred terms to identify the specific cardiotoxic events, severity of the adverse event outcome and if the adverse event reaction stopped

on discontinuing the drug or if it recurred on initiating the drug again. The drug related factors identified from different FAERS files were the specific drug/ targeted therapy name, the role of the drug in the adverse event, route of administration, therapy start/ end date, therapy duration and certain dosage related characteristics such as dosage amount, unit, form and frequency. The targeted therapies used in HER2 positive breast cancer patients were identified using terms such as 'Trastuzumab', 'Ado-trastuzumab', 'Herceptin', 'Kadcyla', 'Pertuzumab' ,'Perjeta', 'Lapatinib', 'Tykerb', 'Neratinib', 'Nerlynx', ' Tucatinib' and ' Tukysa'. These targeted therapies were chosen based on the guidelines recommended by American Cancer Society specific to targeted therapies used in HER 2 positive breast cancer patients.[23] All the variables identified from FAERS to be included in the study (patient characteristics, cardiovascular adverse event related factors and drug related factors) are summarized in Table 23 below.

Table 23: Study Variables for Aim 2A identified from the FAERS 2005-2015 database

Variable	**Variable description**
Patient characteristics: -AGE_COD -GENDR_COD -DEATH_DT -I_F_COD	 -Age -Gender -Patient's death date -Initial/ follow up code
Cardiovascular adverse event related factors: -OCCR_COD -EVENT_DT -PT (string searches for 'Myocardial Infarction', 'Arrhythmia', 'Cardiac Failure', 'Cardiac signs and symptoms' and 'Myocardial Disorders' -OUTC_COD -DECHAL	 -Reporter's type of occupation (physician, pharmacist, other health professional, etc.) -Adverse event date -Preferred term to identify cardiovascular adverse event -Patient's adverse event outcome (death, life threatening, hospitalization, disability, required intervention, other serious event)

-RECHAL	-Reaction stopped on stopping the drug therapy -Reaction recurred on restarting the drug therapy
Drug related factors: -DRUGNAME -ROLE_COD -ROUTE -DOSE_AMT/ DOSE_UNIT/ DOSE_FORM/ DOSE_FREQ -START_DT/ END_DT -DUR_COD	 -Chemotherapeutic drug name and also to identify other drugs that were given along with it -Role of the drug (primary/ secondary suspect drug, concomitant or interacting) -Route of administration -Dosage amount, unit, form and frequency -Therapy start and end date -Therapy duration

Statistical Analyses:

Descriptive analyses were used to summarize the cancer therapy characteristics associated with cardiovascular adverse events. Frequencies and means were used to summarize categorical and numerical variables respectively. Disproportionality analyses was conducted to confirm a potential association between a specific cancer drug and a cardiovascular adverse event.[24] This analysis was conducted using a logistic regression model for bivariate and multivariate analysis. The odds ratios that are derived using these logistic regression models (bivariate and multivariate) are referred to as reporting odds ratios when looking at specific drug-adverse event pairs. Disproportionality analyses with reporting odds ratios (ROR) was used to evaluate the magnitude of event signals in the FAERS. Cases were reports of cardiotoxicity events and non-cases were all reports of adverse events other than cardiotoxicity. The ROR for disproportionality analysis was calculated using a case/non-case method using the formula stated below.

$$ROR = \frac{a / b}{c / d}$$

where, a = Cardiotoxic adverse events reports within those receiving targeted therapy

b = Non-cardiotoxic adverse events reports within those receiving targeted therapy

c = Cardiotoxic adverse events reports within those not receiving targeted therapy

d =Non-cardiotoxic adverse events reports within those not receiving targeted therapy

The ROR is estimated as the odds of cardiotoxicity in those exposed to each cancer drug divided by the odds of cardiotoxicity in those not exposed to the drug of interest (all other drugs in the database). A significant disproportionality, or in other words a possible signal was defined as the lower bound of the 95% confidence interval (95% CI) exceeding 1. Once the bivariate association was established using the disproportionality approach, a logistic regression model was built accounting for other covariates. These cancer therapy related characteristics were defined by using all the other variables mentioned in Table 23.

Aim 2b: To develop a MCDA model to conduct benefit-risk assessment of breast cancer therapies

Multiple criteria decision analysis (MCDA) model was used for comparative benefit-risk assessment of breast cancer therapy regimens that cause cardiotoxicity. The indication considered in our study to define the decision problem was breast cancer therapy regimens. Based on the cardiotoxicities involved, trastuzumab-based therapy regimens were chosen to be included in the model since these were the most cardiotoxic.[25] Most of the literature available on benefits and risks associated with therapies was surrounding breast cancer therapy regimens including trastuzumab. [26-36] There was thus

enough evidence to build a MCDA model evaluating breast cancer therapy regimens. The data needed for MCDA model or the model inputs were thus entirely obtained from the literature. [26-36] The alternatives under consideration were breast cancer therapy regimens that are associated with highest cardiotoxic potential but also improve cancer outcomes/survival. The three breast cancer trastuzumab-based therapy regimens considered were namely, trastuzumab with a taxane, trastuzumab and pertuzumab with a taxane and trastuzumab with cyclophosphamide/ carboplatin along with a taxane. These specific regimens were chosen based on the NCCN treatment guidelines and available evidence in the literature for breast cancer patients.[14] These are the first line breast cancer therapies that are used in HER 2 positive patients and had enough literature evidence for the MCDA model with respect to the benefits and the risks criteria. We also met and discussed these regimens with a pharmacist at VCU, Dr. Erin Hickey who specializes in oncology treatment. On consulting with her and scanning the literature for available evidence on similar groups of HER 2 positive breast cancer patients, we finalized our therapy alternatives for the MCDA model. Specific drugs under each therapeutic category are summarized in Table 24 below.

Table 24: Treatment Alternatives Considered in the MCDA Model

Treatment alternative	Specific therapy
Targeted therapy	Trastuzumab Pertuzumab
Taxanes	Paclitaxel Docetaxel

The focus of the MCDA was on the benefit-risk assessment of breast cancer therapy regimens and the difficulty in choosing a first-line treatment. We focused on the

first-line treatment options since those would form the basis of your treatment plan and future decisions might depend on the outcomes of the first-line treatment. On further consulting with Dr.Hickey, we also decided to expand the criteria of our model to include other adverse events and not restrict to cardiovascular implications of these therapies. These other adverse events such as diarrhea, peripheral neuropathy and febrile neutropenia might also have a major impact on decision-making. Our final model thus focused on survival and quality of life outcomes as benefits criteria and diarrhea, peripheral neuropathy and febrile neutropenia along with cardiovascular adverse effects as risks criteria. A set of nonoverlapping evaluation criteria were chosen to assess the risk/benefit outcomes. Figure 20 below represents the effects tree that summarized the criteria that were evaluated in the MCDA model. Effects tree is a technique used to organize and visualize the MCDA model that summarizes all the favorable/benefits criteria under one branch and all the unfavorable/risks criteria under another branch for the same therapy alternatives.[13] Each of the criteria mentioned in the effects tree were scored and weighted to quantify the benefit-risk score associated with each of the alternatives in the model.

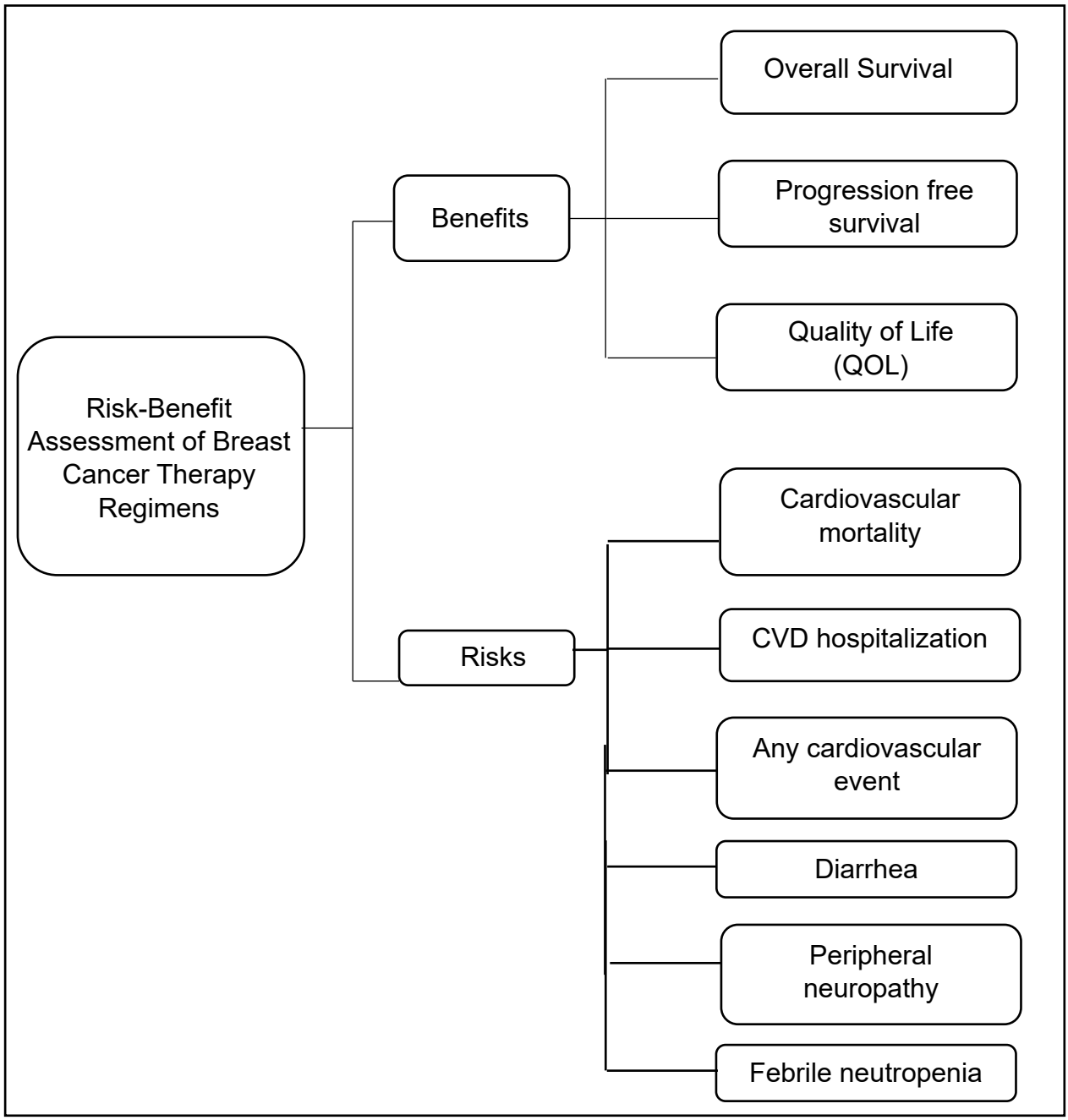


Figure 20: Effects Tree for the MCDA Model

The model was thus built using favorable (overall survival, progression-free survival and quality of life {expressed as quality adjusted life years – QALY}) and unfavorable effects (number of cardiovascular adverse events, cardiovascular mortality, CVD related

hospitalizations, diarrhea, peripheral neuropathy and febrile neutropenia) criteria. A benefit-risk assessment was carried out using these favorable and unfavorable effects criteria. The performance of each of the therapy regimen alternatives (trastuzumab + taxane, trastuzumab + pertuzumab + taxane, trastuzumab + cyclophosphamide/ carboplatin + taxane) on the criteria mentioned above was evaluated using data from the literature. Similar studies from the literature were used to evaluate performance of therapy alternatives on each of these criteria by providing model inputs.[26-36] A literature review was conducted to evaluate studies that have looked at benefits and risks associated with breast cancer therapy regimens. Our search strategy used a combination of terms such as "Molecular Targeted Therapy" OR "Drug Therapy" AND “Survival” OR “Adverse Events” OR “Risks” OR “Benefits” AND "HER 2 positive”. We then evaluated all the studies by screening the titles and abstracts to include only those studies that have looked at the breast cancer therapy regimens that we were interested in as therapy alternatives for the MCDA model. We excluded studies that had a significantly different study population, specifically if they were looking at the specified therapy regimens adjuvantly with other treatment options such as anthracyclines, radiotherapy, or surgery. We also excluded studies that only reported clinical and laboratory outcomes. We included studies that reported results on criteria that were included in our MCDA model. After finalizing studies to provide model inputs, we also identified some more studies from the references of the pre finalized studies. The criteria for evaluating the cancer therapies using an MCDA model were further tuned and defined based on Dr.Hickey’s suggestions and the available literature evidence (Febrile neutropenia was added after evaluating literature since it was well reported) . Studies conducted by Swain et al., Li et al., Hajjar et al. and Garrison et al.

were used to input benefits criteria.[26-29] Studies conducted by Swain et al., Advani et al., Tolaney et al., Schneeweiss et al., Woodward et al., Tanaka et al. and Hussain et al. were used to input the risks criteria.[30-36] These are summarized in Table 25 below. The literature sources used to provide inputs for all of these criteria for each therapy alternative are summarized in Table 25 below.

Table 25: Literature Sources for Model Inputs

Model criteria	Therapy alternatives	Sources
Benefits		
Overall Survival	Trastuzumab + Docetaxel Trastuzumab + Pertuzumab + Docetaxel Docetaxel + Carboplatin + Trastuzumab	Swain et al., 2015[26] Swain et al., 2015[26] Li et al., 2018[27]
Progression free survival	Trastuzumab + Docetaxel Trastuzumab + Pertuzumab + Docetaxel Docetaxel + Carboplatin + Trastuzumab	Swain et al., 2015[26] Swain et al., 2015[26] Li et al., 2018[27]
Quality of life (expressed as QALY)	Trastuzumab + taxane Trastuzumab + Pertuzumab + chemotherapy Trastuzumab + Carboplatin + Docetaxel	Hajjar et al., 2019[28] Garrison et al., 2019[29] Hajjar et al., 2019[28]
Risks		
CVD adverse events		
Cardiovascular mortality	Trastuzumab + Paclitaxel Trastuzumab + Pertuzumab + Docetaxel Trastuzumab + Cyclophosphamide + Paclitaxel	Swain et al., 2015[26] Swain et al., 2015[26] Advani et al., 2020[30]
CVD hospitalization	Trastuzumab + Paclitaxel Trastuzumab + Pertuzumab + Docetaxel Trastuzumab + Cyclophosphamide + Paclitaxel	Tolaney et al., 2015[31] Schneeweiss et al., 2018[32] Advani et al., 2020[30]
Any cardiovascular event*	Trastuzumab + Paclitaxel Trastuzumab + Pertuzumab + Docetaxel Trastuzumab + Cyclophosphamide + Paclitaxel	Tolaney et al., 2015[31] Swain et al., 2015[26] Woodward et al.,2019[33] Advani et al., 2020[30]
Any other adverse event		
Diarrhea	Trastuzumab + Paclitaxel	Swain et al., 2017[34]

	Trastuzumab + Pertuzumab + Docetaxel Trastuzumab + Cyclophosphamide + Paclitaxel	Swain et al., 2017[34] Hussain et al., 2018[35]
Peripheral neuropathy	Trastuzumab + Paclitaxel Trastuzumab + Pertuzumab + Docetaxel Trastuzumab + Cyclophosphamide + Paclitaxel	Tolaney et al., 2015[31] Woodward et al.,2019[33] Tanaka et al., 2015[36]
Febrile neutropenia	Trastuzumab + Paclitaxel Trastuzumab + Pertuzumab + Docetaxel Trastuzumab + Cyclophosphamide + Paclitaxel	Swain et al., 2017[34] Swain et al., 2017[34] Hussain et al., 2018[35]

*The cardiovascular events evaluated were heart failure, dysrhythmia, ischemia or cardiomyopathy

The effects tree was defined based on the outcomes mentioned in the studies above. A performance matrix (effects table) was built using the model inputs from the literature to describe the performance of each of the therapy alternatives on the criteria mentioned. Performance matrix summarizes the extracted information from the literature with the outcome values (model inputs) for each therapy alternative under consideration.[13] These model inputs for each of the criteria across therapy regimens are summarized in Table 26 below which was the performance matrix for the MCDA model.

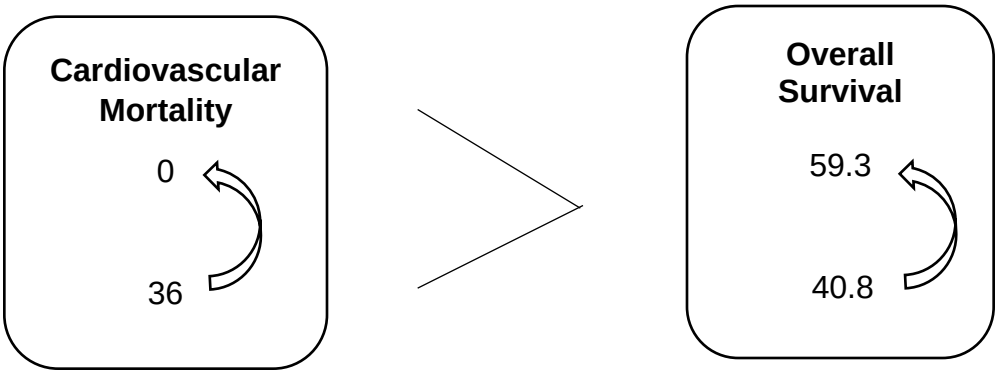
Table 26: Model Inputs for the MCDA Model – Performance Matrix

Criteria	Regimen 1	Regimen 2	Regimen 3
Overall survival (months)[26,27]	40.8	56.5	59.3
Progression-free survival (%)[26,27]	78.8	70.6	84.6
QALY[28,29]	16.17	15.57	15.02
Cardiovascular mortality (%)[26,30]	29	36	0
Cardiovascular hospitalization (%)[30-32]	3.2	16	61
Any other cardiovascular event (%)[26,30,31]	0.5	18	2.8
Diarrhea (%)[34,35]	43	59	54
Peripheral neuropathy (%)[31,33,36]	13.1	56	12
Febrile neutropenia (%)[34,35]	7	13	4.5

Regimen 1: Trastuzumab + taxane
Regimen 2: Trastuzumab + Pertuzumab + taxane
Regimen 3: Trastuzumab + Cyclophosphamide/Carboplatin + taxane

These criteria were scored to enable comparison onto a common scale. Direct rating compositional scoring approach using partial value functions was used to score the criteria. The values from literature for each criterion were used as functions that were then scored giving the highest score to the best value and others relative to it. These scores have been summarized later in Table 32. There are multiple scoring techniques that can be used namely, visual analog scale (VAS), analytical hierarchy process (AHP), MACBETH and point allocation.[37] All the other techniques are mainly used with qualitative data or inputs where the criteria can be compared based on their categories. For VAS and point allocation, points or scores are assigned to alternatives in proportion to their relative importance on a criteria. AHP and MACBETH compare alternatives pairwise on each criterion to assess their importance and assign an average score of those comparisons. With qualitative categorical criteria it is easier to allocate these points and make pairwise comparisons. However with numerical values, it is not possible to make such comparisons since there are no categories within criteria to be compared. With quantitative data that reports mean values and percentages, it is preferable to use direct literature inputs as partial functions and score the criteria. Other than these, if time and money are not an issue and it is feasible to conduct a study, decompositional scoring approaches can also be used to score criteria where overall value of the alternative is assessed to begin with and scores and weights are then derived from these. Discrete choice experiment is a decompositional approach where the recruited study sample rank their alternative first and later specify the most important criteria they considered while making the choice. However, designing a discrete choice experiment requires funding and is not very feasible. We chose the direct rating approach with partial functions for feasibility, ease of interpretability

and the type of model inputs we used in the MCDA model. Once these alternatives were scored, these were weighted using the swing weighting approach. Similar to scoring, AHP and MACBETH can also be used for weighting however due to our type of model inputs we decided to use swing weighting. Swing weighting also allows integrating scores while assigning weights where highest weight is assigned to the criterion that would improve the overall value of the alternative the most on swinging from its worst to best score.[37] Another technique used for weighting is SMARTER ranking technique that weights the criteria irrespective of the score assigned.[37] However, when the scores have been assigned using direct rating by considering the worst and the best value (partial functions), it is preferable to use swing weighting that considers the worst and the best value as well. An example of swing weighting would be as follows which shows that swinging the scale of cardiovascular mortality from 36 (worst value) to 0 (best value) might be more important than swinging the scale for overall survival from 40.8 (worst value) to 59.3 (best value) based on literature inputs:



Aggregate scores were back calculated using the additive model to assign a value for each of the alternatives. Following is the function that was used to assign value using the additive model:

$V_j = \sum S_{ij} \,.\, W_i$ (V = overall value, S = score, W = weight)

Scores, weights and aggregate values per alternatives were estimated using the MS Excel, 1000minds MCDA software and RStudio.[38,39] We used MS Excel to create the performance tables that were then pulled into R for analyses. Minimum and maximum values for each of the criteria were specified (maximum as the best for benefits and minimum as the best for risks) on R. The MCDA package on R then assigns a score of 1 to the maximum value of criterion across the therapy regimens, and others are then scored and normalized with respect to the best value. [39] These assigned scores are summarized below in Table 32. Once normalized and scored, each criterion was then weighted using the relative importance values obtained from 1000minds by assigning a negative weight to risks and positive weight to benefits. To assign these weights 1000minds provides a series of comparisons between criteria. For each comparison, we picked the criteria that we considered would be more valuable with respect to swinging the score from worst to best based on literature inputs. All the comparisons were then aggregated to estimate a criterion preference value (Figure 23 below) by 1000minds software that was then used to assign weight manually on R (Table 33 below). These were then ranked on R by aggregating scores and weights and assigning a quantitative value to each alternative. These overall generated values were then compared for alternatives prespecified in the MCDA model. [13,37-44] This entire process of developing the MCDA model has been summarized in Figure 21 below.

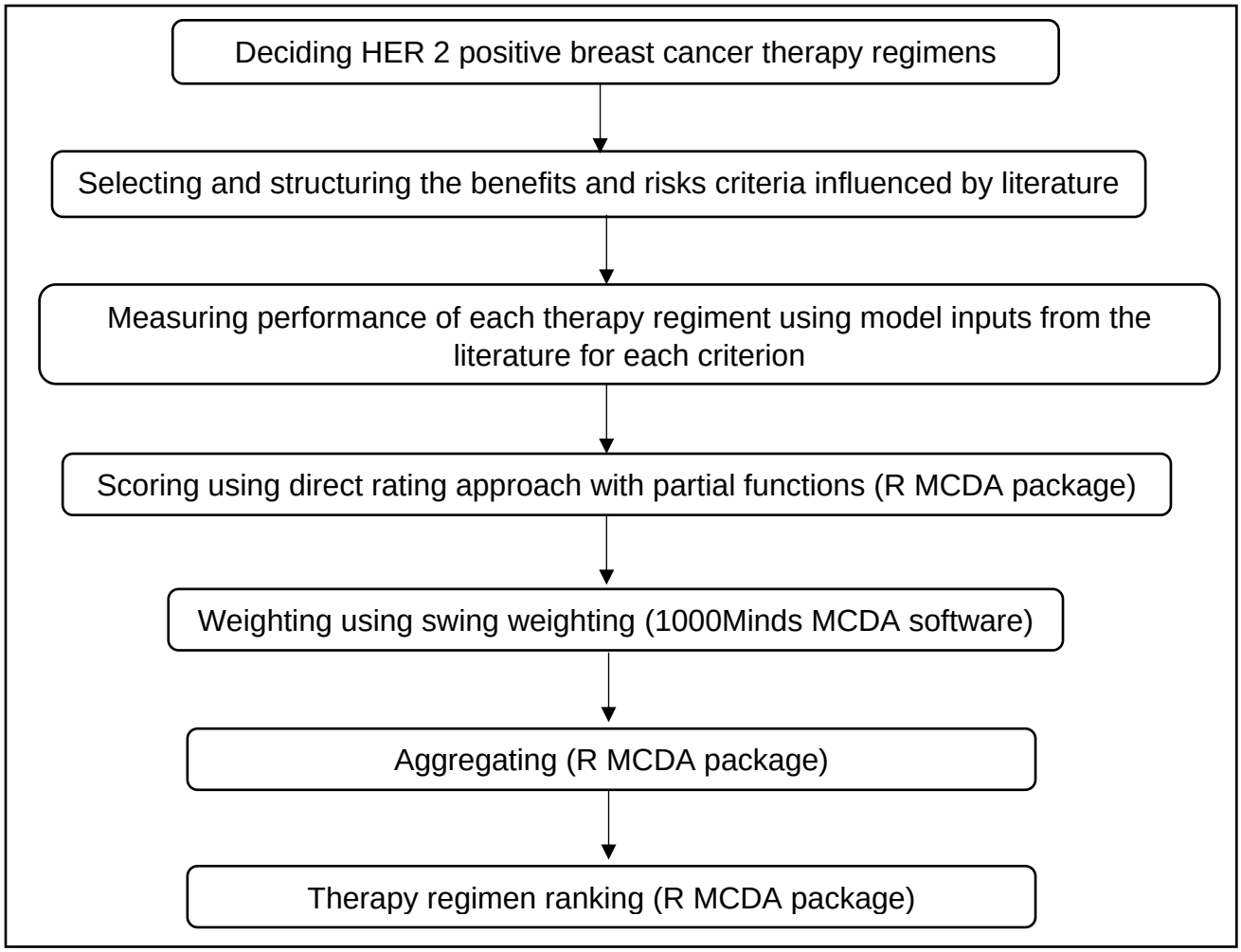


Figure 21: Process Of Developing MCDA Model

Aim 2c: To conduct sensitivity analyses to assess the model performance and uncertainty in the model

Sensitivity analyses were conducted using ranges from literature for the parameter inputs to assess uncertainty in the model. These ranges were confidence intervals that were reported in the studies mentioned above in Table 25 for certain criteria within specific therapy alternative. Ranges for parameter inputs such as overall survival, cardiovascular hospitalization and any other cardiovascular event were used. The inputs from Table 26

were used as the base case values whereas the inputs from Table 27 were used for the sensitivity analyses. One parameter was changed at a time to assess the impact on the outcome. We replaced one base value at a time in the performance Table 26 by values mentioned in Table 27 to assess its effects on preferences for therapy regimens. The results of the sensitivity analyses were used to make conclusions about the uncertainty of the model and the parameters that affect the model the most. Along with using the values from Table 37, we also used certain different sets of weights to weight these criteria differently and assess the effects on therapy alternative ranking. The results for therapy ranking with different sets of weights tried have been summarized in the 'Results' section below. All these values were derived from the literature.

Table 27: Values Used for Sensitivity Analyses of The MCDA Model

Criteria	Sensitivity analyses values
Overall survival (months)[26,27]	
Regimen 1	35.8, 48.3
Regimen 2	49.3, NA*
Regimen 3	NA
Cardiovascular hospitalization (%)[30-32]	
Regimen 1	1.7, 5.4
Regimen 2	NA
Regimen 3	34, 77
Any other cardiovascular event (%)[30,31,33]	
Regimen 1	0.1, 1.8
Regimen 2	NA
Regimen 3	1.6, 4.1

NA* - Not available
Regimen 1: Trastuzumab + taxane
Regimen 2: Trastuzumab + Pertuzumab + taxane
Regimen 3: Trastuzumab + Cyclophosphamide/Carboplatin + taxane

RESULTS:

Aim 2A: To describe the cancer therapy characteristics associated with the cardiotoxic adverse events and evaluate the drug-event association using a disproportionality analysis

There were 35,630,544 adverse events reported in breast cancer patients identified from FAERS 2005-2015 files. In the FAERS data we found that, of those receiving targeted breast cancer therapy, 3.82% reported events were a cardiotoxic adverse event as compared to 3.46% in the non-targeted therapy group. Majority of cardiotoxic adverse events reported on FAERS were at an initial visit. In majority of the cases, the targeted therapy drug was a primary suspect and was reported by the physician. The route of administration in majority of the cases was 'Oral' followed by 'Others' which mainly comprised of subcutaneous and intraperitoneal along with some other categories. These descriptive characteristics of events reported on FAERS have been reported in Table 28 below. We did not report the dosage related characteristics and rechallenge/ dechallenge characteristics since the proportion of missing values in these was more than 90%.

Table 28:Descriptive Characteristics Associated with Cardiotoxic Events in Breast Cancer Patients in the US From FAERS 2005-2015 files

Characteristics	**Cardiotoxic event N(row%, column%) 1,234,823 (3.47)**	**No cardiotoxic event N(row%, column%) 34,395,721 (96.53)**
Gender		
Females	629,926 (3.06, 51.01)	19,920,000 (96.94, 57.93)
Neutral	398 (0.14, 0.03)	280,435 (99.86, 0.85)
Missing	604,499 (4.08, 48.96)	14,195,933 (95.91, 41.22)
Age, in years (mean)	45.8	54.8
Time of visit		
Initial visit	590,384 (2.97, 47.81)	19,290,000 (97.03, 56.08)

Follow up visit	644,419 (4.09, 52.19)	15,100,000 (95.91, 43.91)
Missing	0 (0)	5,017 (0.01)
Role of the drug		
Interacting drug	2,434 (2.05, 0.20)	116,052 (97.95, 0.34)
Primary suspect	417,106 (3.39, 33.78)	11,870,000 (96.61, 34.51)
Secondary suspect	192,389 (2.89, 15.58)	6,475,129 (97.11, 18.83)
Concomitant	622,894 (3.76, 50.44)	15,920,000 (96.24, 46.30)
Missing	0 (0)	10,707 (0.03)
Occupation of the reporter		
Lawyer	130,022 (9.68, 10.53)	1,212,530 (90.32, 3.53)
Physician	430,611 (4.19, 34.87)	9,852,746 (95.81, 28.65)
Other health professional	222,937 (3.46, 18.05)	6,223,498 (96.54, 18.09)
Pharmacist	50,652 (3.08, 4.10)	1,595,841 (96.92, 4.64)
Registered nurse	9 (0.02, 0)	40,546 (99.98, 0.12)
Consumer	253,811 (2.04, 20.55)	12,170,000 (97.96, 35.39)
Missing	146,781 (4.26, 11.89)	3,297,233 (95.74, 9.59)
Route of administration		
Oral	389,933 (4.09, 31.58)	9,145,221 (95.91, 26.59)
Parenteral	1,036 (3.53, 0.08)	28,329 (96.47, 0.08)
Respiratory	8,917 (2.44, 0.72)	356,855 (97.56, 1.04)
Rectal/Vaginal	1,919 (2.68, 0.16)	69,755 (97.32, 0.20)
Topical	37,880 (2.23, 3.07)	1,658,865 (97.77, 4.82)
Others	121,220 (2.96, 9.82)	3,969,386 (97.04, 11.54)
Missing	673,918 (6.5, 54.58)	19,162,313 (93.5, 55.73)
Adverse event outcome		
Death	142,064 (6.26, 11.50)	2,126,695 (93.74, 6.18)
Life-Threatening	72,911 (6.98, 5.90)	971,642 (93.02, 2.82)
Hospitalization	505,101 (5.34, 40.90)	8,947,317 (94.66, 26.01)
Disability	21,229 (2.70, 1.72)	765,364 (97.30, 2.23)
Congenital Anomaly	14,479 (10.61, 1.17)	121,993 (89.39, 0.35)
Required Intervention	9,212 (3.63, 0.75)	244,274 (96.37, 0.71)
Other serious events	427,455 (3.33, 34.62)	12,410,000 (96.67, 36.07)
Missing	42,372 (0.48, 3.43)	8,811,590 (99.52, 25.62)
Therapy duration		
< 1 Day	7,826 (4.66, 0.63)	159,948 (95.34, 0.47)
Days	135,107 (5.73, 10.94)	2,224,555 (94.27, 6.47)
Weeks	5,643 (3.54, 0.46)	153,895 (96.46, 0.45)
Months	18,111 (7.05, 1.47)	238,817 (92.95, 0.69)
Years	18,186 (5.12, 1.47)	337,177 (94.88, 0.98)
Missing	1,049,950 (3.25, 85.03)	31,280,760 (96.75, 90.94)

Disproportionality approach was used to evaluate the drug-event association between targeted therapy and a cardiotoxic adverse event in breast cancer patients identified using FAERS files. A 2x2 table of targeted therapy and CVD outcome (Table 29)

was created to calculate the odds ratios. Since we are looking at a specific targeted drug and adverse event pair, these would be denoted as reporting odds ratio in this case. The disproportionality approach thus uses the reporting odds ratios obtained from binomial logistic regression.

Table 29: 2x2 Table for Those Reporting A Cardiovascular Adverse Event Across The Breast Cancer Therapy Groups in the US From FAERS 2005-2015 files

		Targeted therapy	
		Yes	No
CVD	Yes	7,173	1,227,650
	No	180,375	34,215,346

The following unadjusted odds ratios reported in Table 30 were obtained for the association between targeted therapy and the odds of cardiovascular adverse event. These results suggest that in the FAERS database, the patients who received targeted therapy had higher odds of reporting a cardiovascular adverse event as compared to those with no targeted therapy. This effect was however not

Table 30: Unadjusted Reporting Odds Ratios for the Association Between Targeted Therapy And Cardiovascular Adverse Event In Breast Cancer Patients in the US From FAERS 2005-2015 files

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
Received targeted therapy	1.108	1.082	1.135
No targeted therapy	Reference	Reference	Reference

Those receiving targeted therapies as identified from the FAERS dataset were 1.108 times more likely be diagnosed with a cardiovascular adverse event as compared

to those without a targeted therapy. These results were then adjusted for age, time of visit (initial visit/ follow up), role of the drug, occupation of the reporter, duration of the therapy, route of administration and severity of the adverse event. On accounting for other factors identified using FAERS, it was found that those receiving a targeted therapy were more likely to be diagnosed with a cardiovascular event as compared to those who were not. These results are reported below in Table 31.

Table 31: Reporting Odds Ratios on Adjusting for Other Covariates Identified from the FAERS 2005-2015 Files

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
Targeted therapy			
Received targeted therapy	1.042	1.014	1.071
No targeted therapy	Reference	Reference	Reference
Age			
Age	1.000	1.000	1.000
Time of visit			
Initial visit	0.926	0.922	0.930
Missing	1.764	1.032	3.013
Follow up	Reference	Reference	Reference
Role of the drug			
Interacting drug	0.533	0.509	0.557
Primary suspect	1.122	1.116	1.128
Secondary suspect	0.771	0.767	0.776
Concomitant	Reference	Reference	Reference
Occupation of the reporter			
Lawyer	3.157	3.129	3.185
Physician	1.272	1.265	1.280
Other health professional	1.077	1.069	1.084
Pharmacist	0.964	0.954	0.975

Registered nurse	0.008	0.004	0.016
Missing	1.332	1.322	1.342
Consumer	Reference	Reference	Reference
Duration of the therapy			
<1 Day	0.793	0.773	0.814
Weeks	0.774	0.751	0.798
Months	1.392	1.368	1.417
Years	1.063	1.044	1.081
Missing	<0.001	<0.001	>999.999
Days	Reference	Reference	Reference
Route of administration			
Parenteral	0.829	0.775	0.887
Respiratory	0.902	0.880	0.924
Rectal/Vaginal	0.913	0.868	0.961
Topical	0.700	0.691	0.709
Others	0.802	0.796	0.808
Missing	0.888	0.882	0.895
Oral	Reference	Reference	Reference
Adverse Event Outcome			
Death	1.242	1.233	1.250
Life-Threatening	1.279	1.268	1.290
Disability	0.446	0.439	0.453
Congenital Anomaly	2.479	2.415	2.545
Required Intervention	0.668	0.653	0.684
Other serious events	0.572	0.569	0.575
Hospitalization	Reference	Reference	Reference

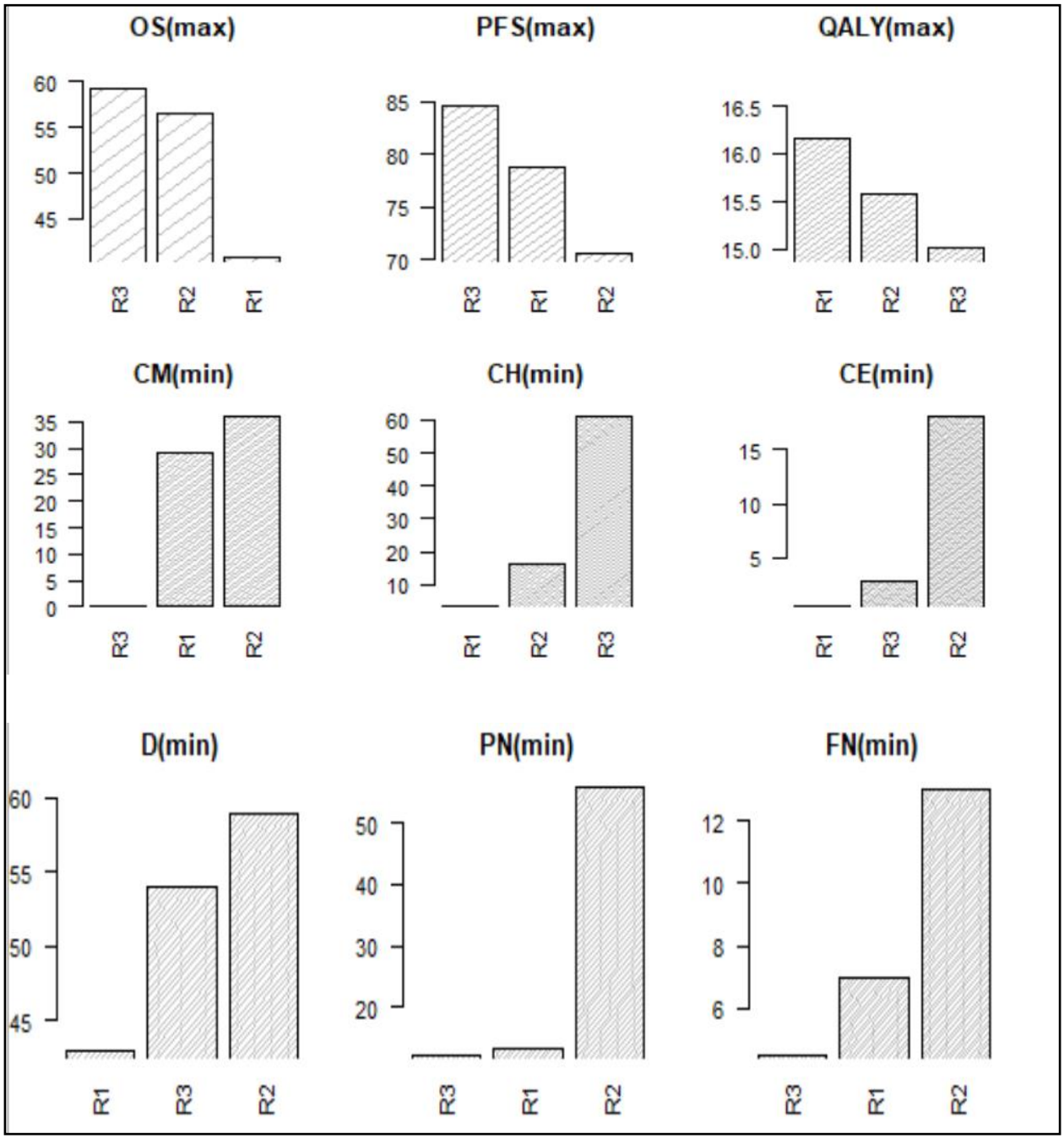
The probability modeled in the logistic regression above was reporting a cardiovascular adverse event

All the above covariates except age were significant predictors as per FAERS data

We can see from the results obtained above using FAERS files that targeted therapy was associated with cardiotoxic adverse events with patients receiving targeted therapy being more likely to be diagnosed with a cardiotoxic adverse event (adjusted OR = 1.042, 95% CI : 1.014, 1.071). However, it can be observed from the confidence interval that the effect size was not that high since the interval and the point estimate were very close to 1. The point estimated obtained on adjusting for other covariates was closer to 1 than unadjusted (1.042 vs 1.108) suggesting that the association might not be significant on accounting for other covariates identified using FAERS 2005-2015 dataset.

Aim 2b: To develop a MCDA model to conduct benefit-risk assessment of breast cancer therapies

An MCDA model was built using breast cancer therapy regimens as the alternatives among breast cancer patients. This model was built from a benefits-risks perspective to assign a quantitative value to each therapy regimen and rank these. There were ten studies identified from the literature (Table 25) that were used to generate model inputs for the MCDA model. These model inputs from the performance table are represented graphically in Figure 22 below to make data visualization easier. This figure suggests that regimens 1 and 3 performed the best on most of the benefits criteria (highest score) and risks criteria (lowest score).



Abbreviations: OS – Overall Survival, PFS – Progression free survival, QALY – Quality Adjusted Life Years, CM – Cardiovascular mortality, CH – Cardiovascular Hospitalizations, CE – Cardiovascular events, D – Diarrhea, PN – Peripheral Neuropathy, FN – Febrile Neutropenia

Regimen 1 (R1): Trastuzumab + taxane

Regimen 2 (R2): Trastuzumab + Pertuzumab + taxane

Regimen 3 (R3): Trastuzumab + Cyclophosphamide/Carboplatin + taxane

Figure 22: Performance of HER 2 positive Therapy Regimen Alternatives on Each Criterion of the MCDA Model Identified From the Literature

Like mentioned above, we used partial function scoring technique to assign scores to the criteria. This uses values identified from the literature that have been entered in the performance table to assign scores. We loaded the performance table onto R, the MCDA package preloaded then normalized these values to convert them into scores. Table 32 below summarizes these scores that were assigned to each criterion for each treatment alternative.

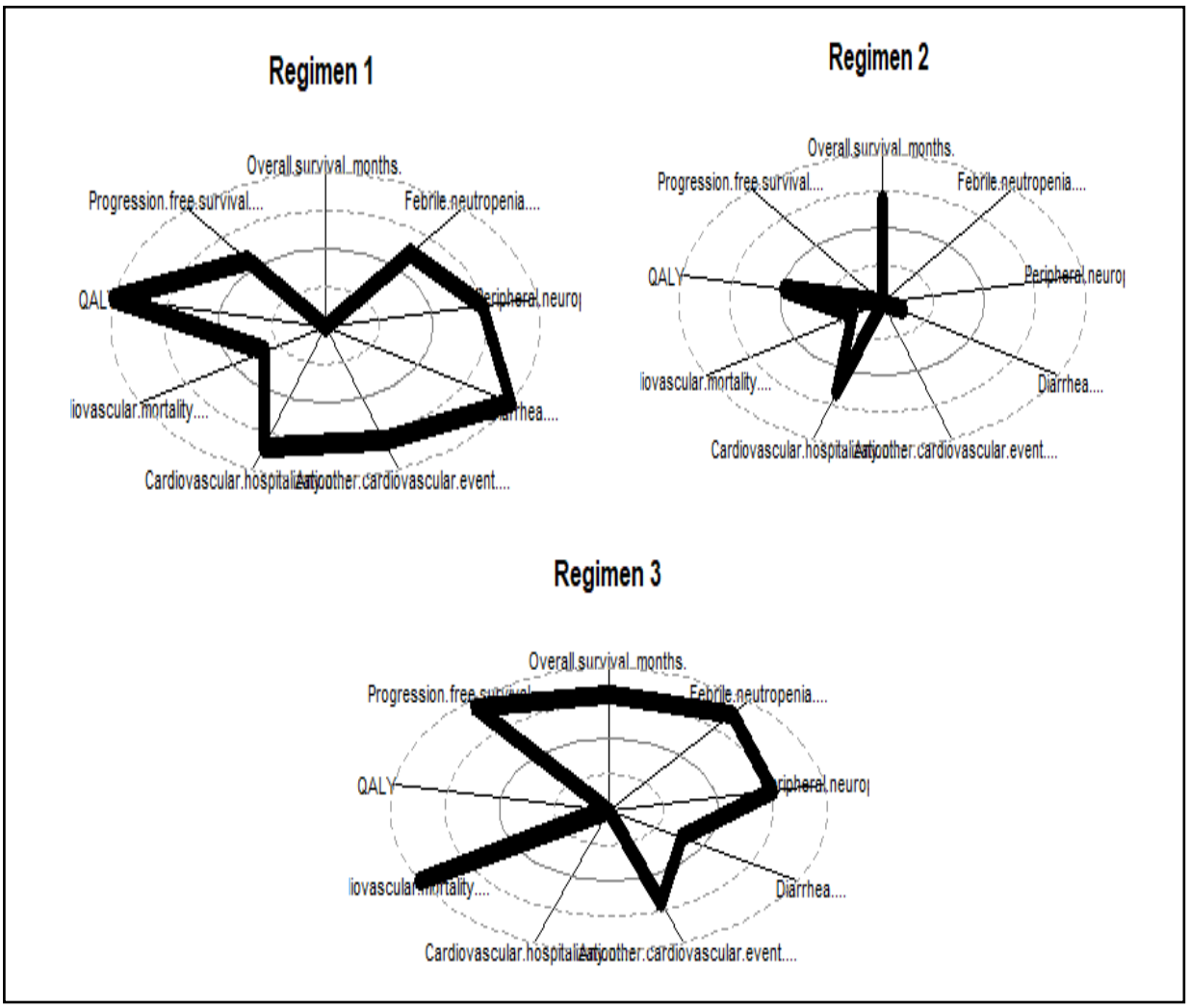
Table 32: Scores Assigned To Criteria Used In The MCDA Model for HER 2 positive Therapy Regimen Alternatives

Criteria	Regimen 1	Regimen 2	Regimen 3
Overall Survival	0.688	0.9527	1
Progression-free survival	0.9314	0.8345	1
QALY	1	0.9628	0.9288
Cardiovascular mortality	0.8055	1	0
Cardiovascular hospitalization	0.0524	0.2622	1
Any other cardiovascular event	0.0277	1	0.1555
Diarrhea	0.7288	1	0.9152
Peripheral neuropathy	0.2339	1	0.2142
Febrile neutropenia	0.5384	1	0.3461

QALY – Quality Adjusted Life Years indicative of quality of life

These assigned scores were then used to plot the following radar plots as seen in Figure 23. We specified for each benefits criterion (overall survival, progression-free survival and QALY) that higher the score better the performance whereas for the remaining risks criteria higher the score lower the performance. The radar plots then take this into consideration to highlight for each treatment alternative it's best performance criteria. The radar plots suggest that regimen 1 thus performs the best on QALY (highest score), diarrhea (lowest score) and cardiovascular hospitalizations (lowest score). Similarly,

regimen 2 performs relatively better on overall survival whereas regimen 3 performs the best on progression free survival, overall survival and cardiovascular mortality.



Regimen 1: Trastuzumab + taxane
Regimen 2: Trastuzumab + Pertuzumab + taxane
Regimen 3: Trastuzumab + Cyclophosphamide/Carboplatin + taxane

Figure 23: Radar Plots For Each HER 2 positive Therapy Regimen Alternative of the MCDA Model Using Literature Inputs And Scores Assigned

These criteria were then run through the 1000minds software to assign relative importance values to each criteria. The software generated criterion preference values as seen in Figure 24 below. Higher criterion preference value suggested that cardiovascular

mortality was ranked the most important criteria while considering the tradeoff between the breast cancer therapy regimens followed by overall and progression-free survival. The 1000minds software provides pairwise comparisons of these criteria to assess which criteria you would consider to be more important to swing the score from worst to best in relation to its comparator. On providing these multiple comparisons the software calculated that cardiovascular mortality was picked 31% times over its comparator to swing the score, which was the highest. The relative importance of cardiovascular mortality was thus the highest. The decision to pick one over the other was just based of physician preferences in the literature.[2,12,45,46] Febrile neutropenia was considered the least important criterion while making trade-offs between the breast cancer therapy regimen.

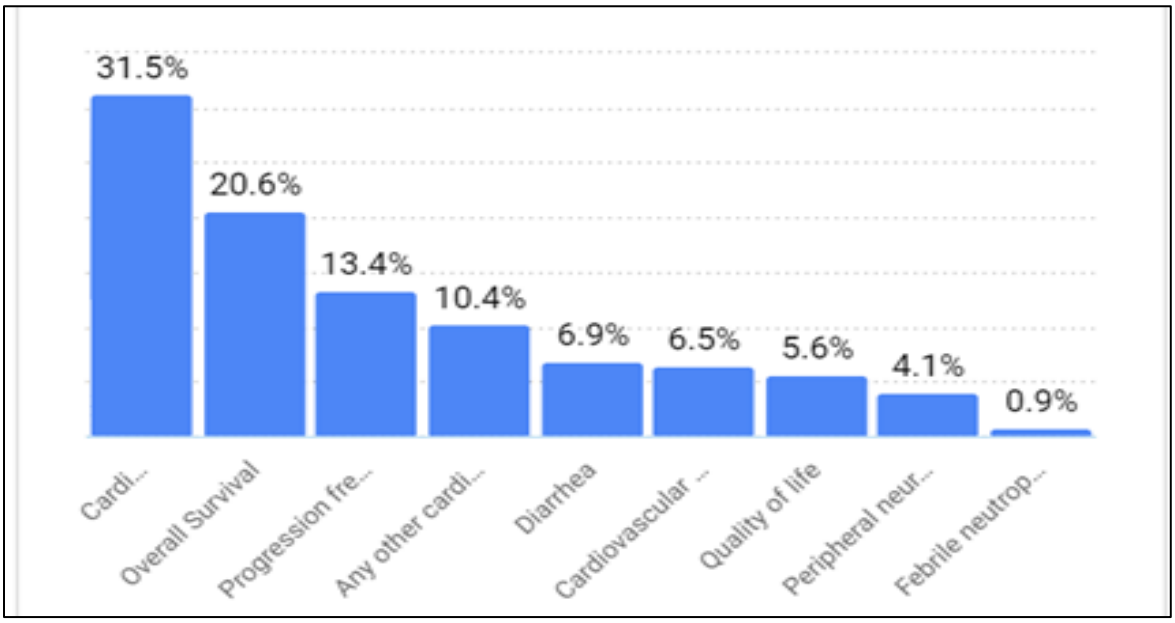


Figure 24: Criterion Preference Values Representing the Relative Importance of Criteria In the MCDA Model

The criterion value functions plot in Figure 25 below was obtained on plotting the preference values of each of the criterion. We had 3 levels for each criterion since we were

comparing 3 treatment alternatives. Each criterion thus had 3 score values for each therapy regimen. The lowest score indicating the lowest level where as the highest score indicating highest level. The criterion value functions/ preference values in Figure 24 indicate that the preference for all the criteria at level 1 is the lowest and it relatively goes on increasing as the criteria performs better (level increases) with that for cardiovascular mortality being the highest. This plot suggests that the marginal importance of cardiovascular mortality was constantly increasing and the highest across all levels of criteria for all the alternatives from its lowest score (lowest level) to its highest score (highest level). Febrile neutropenia had the lowest marginal importance relative to others.

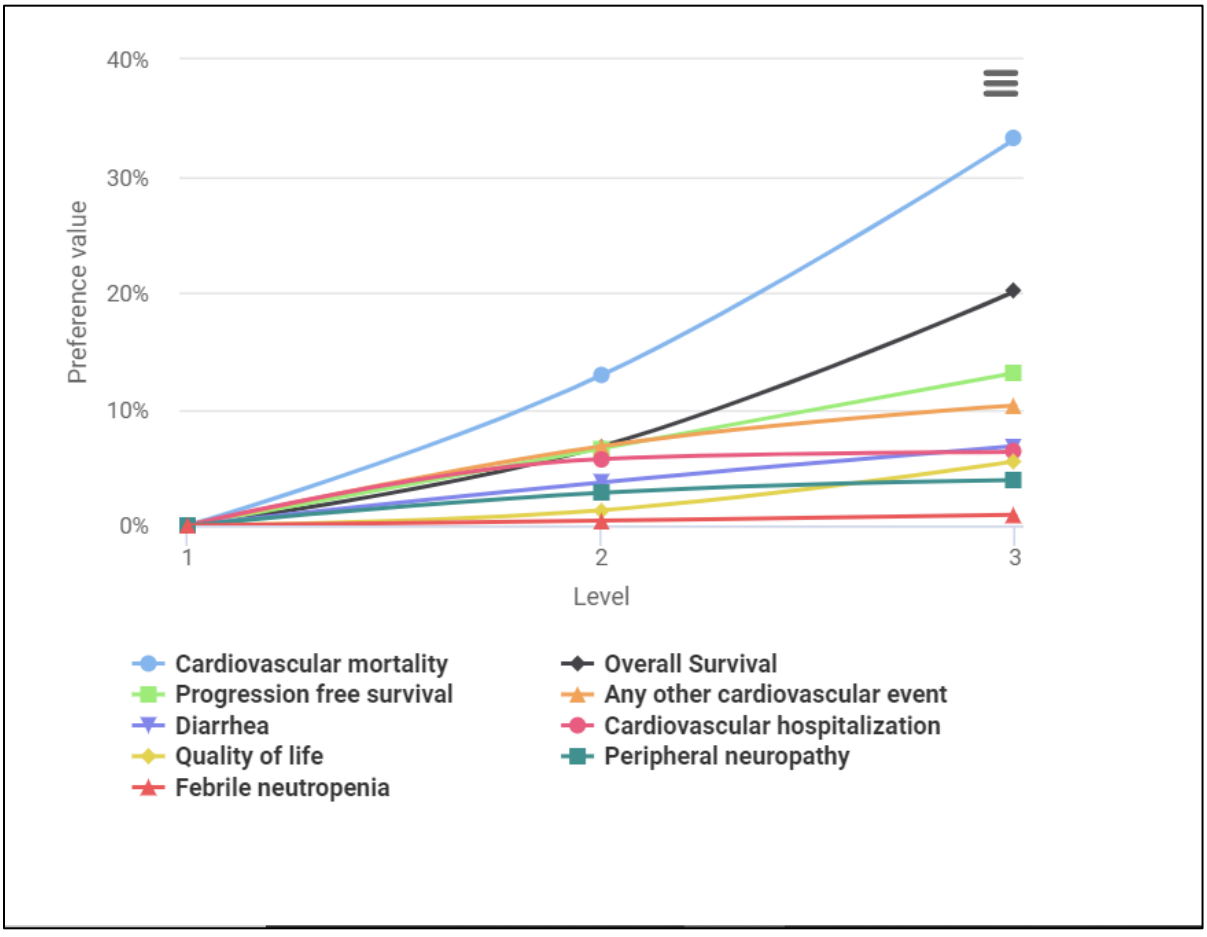
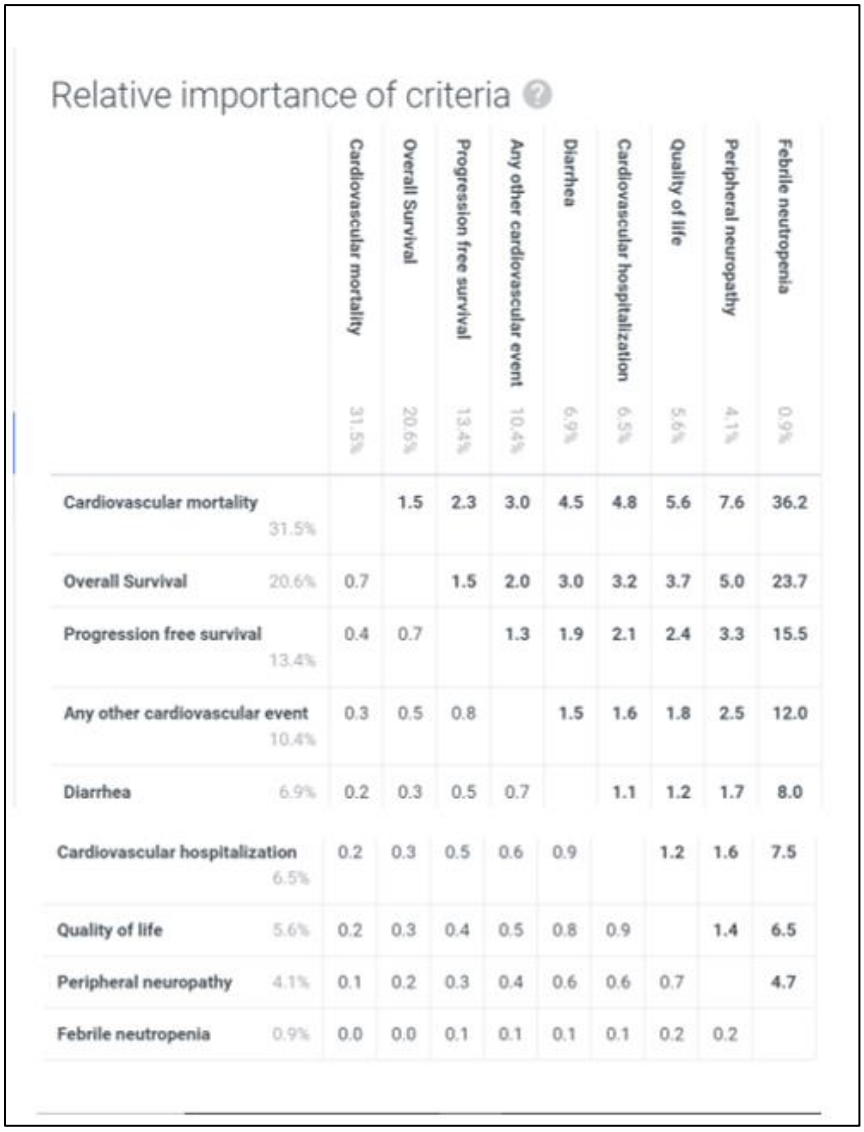


Figure 25: Criterion Value Functions Across the Therapy Regimen Alternatives

The following Figure 26 summarizes the relative importance of each criterion in relation to each other specific criterion. It suggests that cardiovascular mortality is 1.5 times more important than overall survival and so on.



Relative importance of criteria

	Cardiovascular mortality 31.5%	Overall Survival 20.6%	Progression free survival 13.4%	Any other cardiovascular event 10.4%	Diarrhea 6.9%	Cardiovascular hospitalization 6.5%	Quality of life 5.6%	Peripheral neuropathy 4.1%	Febrile neutropenia 0.9%
Cardiovascular mortality 31.5%		1.5	2.3	3.0	4.5	4.8	5.6	7.6	36.2
Overall Survival 20.6%	0.7		1.5	2.0	3.0	3.2	3.7	5.0	23.7
Progression free survival 13.4%	0.4	0.7		1.3	1.9	2.1	2.4	3.3	15.5
Any other cardiovascular event 10.4%	0.3	0.5	0.8		1.5	1.6	1.8	2.5	12.0
Diarrhea 6.9%	0.2	0.3	0.5	0.7		1.1	1.2	1.7	8.0
Cardiovascular hospitalization 6.5%	0.2	0.3	0.5	0.6	0.9		1.2	1.6	7.5
Quality of life 5.6%	0.2	0.3	0.4	0.5	0.8	0.9		1.4	6.5
Peripheral neuropathy 4.1%	0.1	0.2	0.3	0.4	0.6	0.6	0.7		4.7
Febrile neutropenia 0.9%	0.0	0.0	0.1	0.1	0.1	0.1	0.2	0.2	

Figure 26: Criterion Preference Values Representing the Relative Importance of Criteria In the MCDA Model

The criterion preference values like those obtained in Figure 24 above were used to assign weights on R by converting them into probabilities. Cardiovascular mortality having the highest weight of 0.315 with others following in the same order. For the risks criteria, the absolute values of the assigned weights were considered however they were denoted by a negative sign as per the MCDA package requirements to signify unfavorable effects. To account for uncertainty in these preference values, weights were changed and included in the sensitivity analyses as well. The base values used for weights have been summarized in Table 33 below.

Table 33: Weights Assigned To Criteria Used In The MCDA Model

Criteria	Assigned Weights
Overall Survival	0.206
Progression-Free Survival	0.134
QALY*	0.056
Cardiovascular Mortality	-0.315
Cardiovascular Hospitalization	-0.0605
Any Other Cardiovascular Event	-0.104
Diarrhea	-0.069
Peripheral Neuropathy	-0.041
Febrile Neutropenia	-0.009

*QALY – Quality Adjusted Life Years indicative of quality of life

On aggregating the scores and weights assigned to each criterion, the alternatives were ranked in the following order as per the value assigned to these.

Preference 1: Therapy regimen 3 (Trastuzumab + Cyclophosphamide/Carboplatin + taxane)

Preference 2: Therapy regimen 1 (Trastuzumab + taxane)

Preference 3: Therapy regimen 2 (Trastuzumab + Pertuzumab + taxane)

These preference rankings were derived based on the inputs we obtained from the literature. On applying these to real world these might not hold true since the population might have certain differences. We chose studies for model inputs that had similar population to avoid any biases. However, in a real world setting HER 2 positive breast cancer patients might not be as alike and might be on different therapy regimens in addition to those considered in our MCDA model. The decision-making framework might thus need to be revised by adding more inputs and data from the real world to be applicable on a broader scale and result into changes in guidelines. Our preference rankings are however a good starting point to showcase the application of a formal decision-making model in a multidisciplinary field like oncology. Currently there is no clear consensus on the preferences and physician suggest using trastuzumab with any other chemotherapy drug (taxane or cyclophosphamide/ carboplatin) or pertuzumab as the first line treatment.[20,47] Our study might help in narrowing down these broad categories to specific therapy regimens and might add to the current practice guidelines to make it more specific. The treatment alternatives we included in our model were all first-line options and there has not been an attempt in the literature earlier to look at each of these specific regimens and weigh out the benefits and risk associated with each. Earlier guidelines have looked at trastuzumab along with chemotherapy as a whole. Our results help in further narrowing down those chemotherapy options and therapy regimens based on preferences to make decision-making more transparent. The treatment landscape of HER 2 positive breast cancer is constantly growing and emerging, studies like ours can help in navigating the landscape better by assigning rankings and narrowing down the options.

Aim 2c: To conduct sensitivity analyses to assess the model performance and uncertainty in the model

We conducted sensitivity analyses to assess for model robustness and the uncertainty involved. We used values from Tables 27 and 34 to check if the model was sensitive to any of these and if the therapy regimen preference ranking changed on changing any of these values as compared to the base case values (Tables 26 and 33). We changed one performance value at a time from Table 27 by keeping the rest same and assessed the model output. Like mentioned above, these performance values summarized in Table 27 were identified from the literature. The confidence intervals were available in the literature for overall survival, cardiovascular hospitalization, and cardiovascular events. These criteria were thus used in sensitivity analyses. We observed that our model was robust to any of these changes in the criteria performance values. We obtained the same results with the same therapy regimen ranking on changing the base case value to any of the values mentioned in Table 27. The model was also robust to most of the weight changes mentioned in Table 34 below. We kept the same base case values for the performance matrix although changed the criteria weights for this analysis. We used four different sets of weights. In set 1 we kept all the other weights same, although weighted overall survival higher than the cardiovascular mortality given the trade-off between these decisions.[2,12,45,46] The remaining sets of weights were based on a 'trial and error' methodology to assess the effects on therapy ranking. In set 2 we weighted overall survival as high as cardiovascular mortality with others relatively lower. In set 3 we weighted all the benefits criteria higher than risks whereas in set 4 we weighted all the risks criteria higher than benefits. These weights were not driven by literature sources but

based on general understanding. We had similar therapy regimen ranking with sets 1, 2 and 3 of weights. Although on weighting all the unfavorable effects higher than the favorable effects, the ranking of the alternatives changed a little bit. Therapy regimen 1 with a trastuzumab and taxane was now ranked the highest as the most preferred regimen followed by therapy regimen 3 with trastuzumab, cyclophosphamide/ carboplatin and a taxane. The least preferred alternative was therapy regimen 2 with trastuzumab, pertuzumab and a taxane. These results have been summarized in Table 34 below.

Table 34:Different Sets Of Weights Used To Conduct Sensitivity Analyses To Assess The Robustness Of The MCDA Model

Model criteria	Weights 1	Weights 2	Weights 3	Weights 4
Overall survival	0.315	1	1	0.5
Progression-free survival	0.134	0.75	1	0.5
QALY*	0.056	0.5	1	0.5
Cardiovascular mortality	-0.206	-1	-0.5	-1
Cardiovascular hospitalization	-0.0605	-0.75	-0.5	-1
Any other cardiovascular event	-0.104	-0.5	-0.5	-1
Diarrhea	-0.069	-0.25	-0.5	-1
Peripheral Neuropathy	-0.041	-0.25	-0.5	-1
Febrile Neutropenia	-0.009	-0.25	-0.5	-1
Therapy Ranking	**R3[a] - 1, R1[b] - 2, R2[c] - 3 (no change)**	**R3 - 1, R1 - 2, R2 - 3 (no change)**	**R3 - 1, R1 - 2, R2 - 3 (no change)**	**R1 - 1, R3 - 2, R2 - 3 (changed)**

*QALY – Quality Adjusted Life Years indicative of quality of life
[a] R1 – Therapy Regimen 3 (Trastuzumab + Cyclophosphamide/Carboplatin + taxane)
[b] R2 – Therapy Regimen 1 (Trastuzumab + taxane)
[c] R3 – Therapy Regimen 2 (Trastuzumab + Pertuzumab + taxane)

DISCUSSION:

Our findings suggest that based on FAERS data breast cancer patients receiving targeted therapies were more likely to be diagnosed with a cardiovascular event as compared to those who were receiving conventional therapies. These results were adjusted for age, time of visit, occupation of the reporter, therapy duration, route of administration and the severity of the outcomes. However, as mentioned in the results earlier the effect size was small on adjusting for confounding factors and the confidence intervals closer to 1. This suggests that on controlling for other factors observed in FAERS data, this association might not be clinically relevant. Although statistically significant, this could have low clinical utility and more data would be required to make any more conclusions.

On conducting multicriteria decision analysis, we found that the breast cancer therapy with Trastuzumab, cyclophosphamide/ carboplatin and a taxane (paclitaxel/ docetaxel) was the most preferred therapy alternative given the benefits and the risks associated with each of the alternatives. Therapy regimen containing Trastuzumab, Pertuzumab and a taxane (paclitaxel/ docetaxel) was the least preferred alternative. The most important criteria considered in the decision making was cardiovascular mortality followed by overall survival. This is reflective of physicians and oncologists starting to get more worried about the side effects of these therapies along with the cancer outcomes of the patients.[2,12,45,46] This could also be the reason for therapy regimen 2 with two targeted therapies (trastuzumab and pertuzumab) being the least preferred regimen. Since targeted therapies have higher cardiovascular implications, adding two targeted therapies to a regimen can further increase the risk of CVD and down the line can become the least

preferred regimen when compared to others.[8] Currently physicians prescribe treatment based on the clinical factors associated with the tumor type with the goal of prolonging life span.[21,47] With cardiovascular mortality increasing in cancer patients, the goal of prolonging lifespan would not only depend on the tumor type but also on cardiovascular implications. It has been suggested in the literature that clinicians need to be aware about the cardiovascular consequences of certain types of cancers and cancer therapies to have a better coordinated cardiovascular care where the cancer treatment planning can limit the use of targeted therapies.[34] Currently, physicians encourage any trastuzumab based therapy usually coupled with another chemotherapy drug (could be an alkylating agent like carboplatin/ cyclophosphamide or a taxane like paclitaxel/docetaxel).[47] Our study helps in narrowing down these treatment regimen options to specific chemotherapy drugs used. NCCN enlists the treatment guidelines for HER 2 positive breast cancer patients although they do not assess preferences within these.[14] Our study helps in assigning a preferential ranking to these regimens give the inputs on criteria we included in our study. Stating the criteria we considered, and weights used upfront helps in making this decision-making process more transparent by informing the stakeholders (providers in our case) about our approach.[15] The providers can then revise this model as per their needs by adding more criteria and changing the weights to make the model more generalizable. This guided approach would help in treatment planning especially with multidisciplinary teams involved where everyone could weigh in on the model inputs. Our study sets up a basic MCDA model that can further be revised by adding more criteria as evidence gets available to be utilized in the real-world. If the data is available, more alternatives for decision-making

such as surgery, radiotherapy, other treatment regimens can also be added to this model to further make it more applicable.

A study conducted by Scherrer et al. looked at sequential decision making using a multicriteria decision modelling in breast cancer therapy planning.[35] This study suggested that the novel decision-making approach was more efficient in clinical decision making. The authors stated that the model facilitated the establishing of a rule-based system, which encodes medical knowledge of treatment options originating from various sources in a precise and reliable way. This process was time-efficient in making decisions and treatment planning, which addressed the time shortage issue in clinical routine mentioned by the authors. This study was however restricted to case studies of few patients and thus was not generalizable to the entire population. In addition, this study focuses more on breast cancer therapy planning taking into consideration number of physician visits, frequency of medication, etc.[35] Our study however compares therapy regimens used in breast cancer patients from a benefit-risk perspective. The previous study focuses more on problems that are consistent with planning the treatment that involve diagnosis timeline, physician visits, etc. rather than focusing on the therapy alternatives to choose from. Our study is more specific to therapy alternatives that are available for physicians to choose from. We incorporated risks and benefits outcomes associated with each regimen in our model where as the prior study just looked at treatment as one of their criteria and not an alternative. The prior study gives an idea of how efficiently the physicians can plan the treatment with respect to the timeline to be more efficient whereas our study gives an idea of which treatment to choose from given the available options in our model.

Another study conducted by Lin et al. looked at physician experiences and preferences in the treatment of HER 2 negative breast cancer patients. Since this study focused on HER 2 negative patients, their treatment options were different. Treatment preferences were collected by class of endocrine therapy-based regimens versus chemotherapy. This study did not use the MCDA approach, they only looked at survey-based responses by physicians to treatment preferences. The findings suggest that physicians used anastrozole most frequently, followed by everolimus - based therapy followed by fulvestrant-based therapy. Efficacy was the most important consideration for treatment choices followed by tolerability, quality of life and cost of drug in that order. This study however only used a survey - based approach to evaluate current physician preferences to be considered while making a treatment choice.[12] This would help in assigning weights to the criteria, although our study looks at performance of therapy alternatives on prespecified criteria to rank the regimens. We used literature inputs to assess the performance to then assign value to each regimen. In addition, our study was focused on HER 2 positive breast cancer patients.

Another study conducted by DeKoven et al. also looked at treatment patterns for HER 2 positive breast cancer patients. The findings of this study suggested that trastuzumab-based regimens were the preferred option for treating HER 2 positive metastatic breast cancer patients. However, this was also a survey-based study where practicing oncologists were surveyed to identify breast cancer patients and their treatments. The preferred regimen was evaluated based on the biomarker status. This study also did not look at specific regimens but trastuzumab-based regimens as a whole. Our study however looked at specific HER 2 positive breast cancer treatment regimens.[48]

The prior study was also not conducted from a decision-making perspective rather to just evaluate treatment patterns. Our study helps in decision-making by scoring and weighting criteria using literature inputs and rank the therapy regimens.

Our study has many strengths. Firstly, this is one of the first studies that has been conducted in the literature that has applied an MCDA technique to aid formal decision making from a benefit-risk perspective of cancer therapies. The MCDA studies that have been conducted so far in the literature have mainly been from a patient perspective involving shared decision making between the patient and the physician.[7 - 9] These have only looked at patient preferences and satisfaction in order to make a decision. Our study, however, takes into consideration some of the clinical outcomes associated with these therapy regimens that makes comparison of these regimens more formal from a benefit-risk perspective. Secondly, there have been studies in the literature that have individually looked at benefits and risks associated with these therapy regimens or individual targeted drugs. This is one of the first studies that consolidates the beneficial and adverse outcomes of cancer therapy regimens together in a single model to assign a value to each of these regimens. The previous studies that have been conducted with a benefit-risk perspective have been in non-cancer population.[40-43] Thirdly, given the growing concerns associated with the adverse effects of these therapy regimens our findings provide a guided formal decision-making model. They help in choosing one treatment over the other based on the criteria considered in our model and the performance of alternatives on these identified from the literature. Given the growth in research with respect to oncology treatments, the decision-making would get more and more difficult with increasing options. A guided approach like this would thus be necessary in making value trade-offs. Our study

also creates radar plots for each therapy regimen to show the criteria it performs the best on. These findings would inform the physicians about the criteria to consider while choosing between these regimens. Based on the criteria the regimen performed the worst on, physicians can be more cautious about these while prescribing. It also informs physicians about the criteria we considered while making the decision. If any physician has different criteria that they might consider important, based on the inputs of each alternative for these criteria the ranking of the alternatives might change. This model can then be revised and tailored to fit the physician's needs to include more criteria and inputs on these to make it more applicable in the real world.

However, our study also has major limitations. Firstly, FAERS only captures information on adverse events that were reported. Based on literature there is high under reporting of spontaneous adverse events.[49] Given that targeted therapies such as trastuzumab are first line agents, it is possible that there were more events associated with these although were not reported on FAERS. This should be considered while interpreting odds ratios. This underreporting could have also been the reason for smaller effect sizes. This underreporting might limit generalizability and lower validity of the study due to some misclassification and information bias. FAERS also does not capture demographics extensively and thus there could be some confounding effect in the relationship of likelihood of being diagnosed with CVD and receiving targeted therapies. It is possible that certain demographics factors such as race and employment status could affect access to care and hence affect receipt of targeted therapy and diagnosis of cardiotoxic events as well. This however could not be controlled for in our study due to data limitations. Additional data on dosage related characteristics might have also helped in characterizing the cancer

treatment related factors further to help in tailoring these as per patient needs. Secondly, the studies included in the MCDA model were done in slightly different populations which could bias the results to some extent by making the groups not equal in comparison. This would however be reduced since we used a strict inclusion/ exclusion criterion to identify studies that would provide the model inputs to maintain uniformity. We restricted our studies to those that looked at HER 2 positive breast cancer patients. However slight differences in their biomarkers existed which might have an impact on the outcome. In this case the model inputs might have been a result of the type of their biomarker rather than the therapy alternative. Although, we also used sensitivity analyses that would assess the robustness and uncertainty of the model to a range of model inputs to reduce this further. Thirdly, due to lack of evidence across all the therapy regimens, we could only consider criteria that we had evidence on and were consistent across all the alternatives. On gaining additional evidence on some more criteria that the physicians might consider important, the ranking of therapy regimens might be altered. In addition, due to limited availability in the literature, we had to select some criteria such as overall survival and QALY as a substitute for quality of life which might not completely be nonoverlapping. However, QALY was the only measure of quality of life that was reported well in the literature. Progression-free survival might also overlap with overall survival, however in our studies overall survival was measured in months whereas progression-free survival was measured as proportion of patients who were progression-free which might reduce this overlap. Thus, due to limited literature evidence, we could not keep our benefits criteria completely nonoverlapping. Including more data on long term outcomes of these therapies by conducting a longitudinal study in the future might also help in making the model more

generalizable for the real-world. Revising the model can then lead to different preferential rankings between therapies. It is however one of the first studies conducted using an MCDA approach for formal decision making and will inform certain future studies.

Despite the study limitations, the novelty of the methodology and the implications add to the literature. It demonstrated the application of a decision-making methodology in oncology care. With further updates to the model, it can be used in a real-world setting to make decisions and impact the current treatment guidelines. As mentioned earlier, since cancer treatment requires a multidisciplinary approach, a multicriteria decision analysis model would help in making decision-making more formal and transparent. A value assessment of breast cancer therapy regimens can serve as a basis for various payment policies, clinical treatment selection and development on the part of pharmaceutical companies.[50] The use of value-based frameworks in guiding payment policy decisions is increasing.[51] A study like ours with further revisions, can help in making certain policy decisions by reimbursing the therapy regimen that is ranked as the most preferred regimen higher. Treatment ranking can also directly be factored into clinical treatment selection and encouraging more discovery from the pharmaceutical companies for similar regimens and drugs.

REFERENCES:

1. Jones SMW, Chennupati S, Nguyen T, Fedorenko C, Ramsey SD. Comorbidity is associated with higher risk of fi nancial burden in Medicare bene fi ciaries with cancer but not heart disease or diabetes.
2. Liu D, Ma Z, Yang J, et al. Prevalence and prognosis significance of cardiovascular disease in cancer patients : a population-based study. 2019;11(18):7948-7960.
3. Herrmann J. From trends to transformation: where cardio-oncology is to make a difference. *Eur Heart J*. 2019;40(48):3898-3900. doi:10.1093/eurheartj/ehz781
4. Koene RJ, Prizment AE, Blaes A, Konety SH. Shared Risk Factors in Cardiovascular Disease and Cancer. *Circulation*. 2016;133(11):1104-1114. doi:10.1161/CIRCULATIONAHA.115.020406
5. Florescu M, Cinteza M, Dragos V. Chemotherapy-induced Cardiotoxicity. *Maedica - a J Clin Med*. 2013;8(1):59-67.
6. Walls GM, Lyon AR, Harbinson MT, Hanna GG. Cardiotoxicity Following Cancer Treatment. *Ulster Med J*. 2017;86(1):3-9.
7. Mozdzanowska D, Woźniewski M. Radiotherapy and anthracyclines – cardiovascular toxicity. doi:10.5114/wo.2014.40108X
8. Gm W, Ar L, Mt H, Gg H. Cardiotoxicity Following Cancer Treatment. 2017;86(June 2016):3-9.
9. Lenneman CG, Sawyer DB. Review An Update on Cardiotoxicity of Cancer-Related Treatment. 2016:1008-1020. doi:10.1161/CIRCRESAHA.115.303633
10. Chen YY, Wang LINWEI, Chen FF, Wu BIBO, Xiong BIN. Efficacy , safety and administration timing of trastuzumab in human epidermal growth factor receptor 2

positive breast cancer patients : A meta - analysis. 2016:1721-1733. doi:10.3892/etm.2016.3095

11. Abdel-Razaq W, Alzahrani M, Al Yami M, Almugibl F, Almotham M, Alregaibah R. Risk Factors Associated with Trastuzumab-induced Cardiotoxicity in Patients with Human Epidermal Growth Factor Receptor 2-positive Breast Cancer. *J Pharm Bioallied Sci*. 2019;11(4):348-354. doi:10.4103/jpbs.JPBS_276_18
12. Lin PL, Hao Y, Xie J, et al. Physician experiences and preferences in the treatment of HR+/HER2- metastatic breast cancer in the United States: a physician survey. *Cancer Med*. 2016;5(2):209-220. doi:10.1002/cam4.580
13. Thokala P, Devlin N, Marsh K, et al. Multiple Criteria Decision Analysis for Health Care Decision Making — An Introduction : Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Heal*. 2016;19(1):1-13. doi:10.1016/j.jval.2015.12.003
14. NCCN guidelines for cancer patients https://www.nccn.org/patients/guidelines/cancers.aspx accessed on: 11/03/2020
15. Mühlbacher AC, Kaczynski A. Making Good Decisions in Healthcare with Multi-Criteria Decision Analysis: The Use, Current Research and Future Development of MCDA. Appl Health Econ Health Policy. 2016 Feb;14(1):29-40. doi: 10.1007/s40258-015-0203-4. PMID: 26519081.
16. Adunlin G, Diaby V, Montero AJ, Xiao H. Multicriteria decision analysis in oncology. Health Expect. 2015 Dec;18(6):1812-26. doi: 10.1111/hex.12178. Epub 2014 Mar 17. PMID: 24635949; PMCID: PMC4318795.

17. Postmus D, Richard S, Bere N, et al. Individual Trade-Offs Between Possible Benefits and Risks of Cancer Treatments: Results from a Stated Preference Study with Patients with Multiple Myeloma. *Oncologist*. 2018;23(1):44-51. doi:10.1634/theoncologist.2017-0257
18. Wagner M, Samaha D, Cuervo J, et al. Applying Reflective Multicriteria Decision Analysis (MCDA) to Patient – Clinician Shared Decision-Making on the Management of Gastroenteropancreatic Neuroendocrine Tumors (GEP-NET) in the Spanish Context. *Adv Ther*. 2018;35(8):1215-1231. doi:10.1007/s12325-018-0745-6
19. Lifford KJ, Witt J, Burton M, et al. Understanding older women ' s decision making and coping in the context of breast cancer treatment. 2015:1-12.
20. WebMD: What medications can help treat HER2 positive breast cancer? https://www.webmd.com/breast-cancer/qa/what-medications-can-help-treat-her2-positive-breast-cancer accessed on: 11/09/2020
21. American Cancer Society: Understanding Your Options and Making Treatment Decisions https://www.cancer.org/treatment/treatments-and-side-effects/planning-managing/making-decisions.html accessed on: 11/09/2020
22. FDA Advers Event Reporting System (FAERS) Public Dashboard https://www.fda.gov/drugs/questions-and-answers-fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard accessed on : 08/21/2020

23. American Cancer Society: Targeted Therapy for Breast Cancer https://www.cancer.org/cancer/breast-cancer/treatment/targeted-therapy-for-breast-cancer.html accessed on : 11/04/2020

24. Dias P, Penedones A, Alves C, Ribeiro CF, Marques FB. The Role of Disproportionality Analysis of Pharmacovigilance Databases in Safety Regulatory Actions: a Systematic Review. *Curr Drug Saf*. 2015;10(3):234-250. doi:10.2174/1574886310666150729112903

25. Narayan V, Ky B. Common Cardiovascular Complications of Cancer Therapy: Epidemiology, Risk Prediction, and Prevention. *Annu Rev Med*. 2018;69:97-111. doi:10.1146/annurev-med-041316-090622

26. Swain SM, Baselga J, Kim SB, et al. Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer. *N Engl J Med*. 2015;372(8):724-734. doi:10.1056/NEJMoa1413513

27. Li S, Wei W, Jiang Y, et al. Comparision of the efficacy and survival analysis of neoadjuvant chemotherapy for HER-2-positive breast cancer. *Drug Des Devel Ther*. 2018;12:3085-3093. Published 2018 Sep 21. doi:10.2147/DDDT.S171534

28. Hajjar A, Ergun MA, Alagoz O, Rampurwala M. Cost-effectiveness of adjuvant paclitaxel and trastuzumab for early-stage node-negative, HER2-positive breast cancer. *PLoS One*. 2019;14(6):e0217778. Published 2019 Jun 5. doi:10.1371/journal.pone.0217778

29. Garrison LP Jr, Babigumira J, Tournier C, Goertz HP, Lubinga SJ, Perez EA. Cost-Effectiveness Analysis of Pertuzumab With Trastuzumab and Chemotherapy Compared to Trastuzumab and Chemotherapy in the Adjuvant Treatment of

HER2-Positive Breast Cancer in the United States [published correction appears in Value Health. 2019 Jul;22(7):843]. *Value Health*. 2019;22(4):408-415. doi:10.1016/j.jval.2018.11.014

30. Advani PP, Ballman KV, Dockter TJ, Colon-Otero G, Perez EA. Long-Term Cardiac Safety Analysis of NCCTG N9831 (Alliance) Adjuvant Trastuzumab Trial. *J Clin Oncol*. 2016;34(6):581-587. doi:10.1200/JCO.2015.61.8413
31. Tolaney SM, Barry WT, Dang CT, et al. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer [published correction appears in N Engl J Med. 2015 Nov 12;373(20):1989]. *N Engl J Med*. 2015;372(2):134-141. doi:10.1056/NEJMoa1406281
32. Schneeweiss A, Chia S, Hickish T, et al. Long-term efficacy analysis of the randomised, phase II TRYPHAENA cardiac safety study: Evaluating pertuzumab and trastuzumab plus standard neoadjuvant anthracycline-containing and anthracycline-free chemotherapy regimens in patients with HER2-positive early breast cancer. *Eur J Cancer*. 2018;89:27-35. doi:10.1016/j.ejca.2017.10.021
33. Woodward N, De Boer RH, Redfern A, et al. Results From the First Multicenter, Open-label, Phase IIIb Study Investigating the Combination of Pertuzumab With Subcutaneous Trastuzumab and a Taxane in Patients With HER2-positive Metastatic Breast Cancer (SAPPHIRE). *Clin Breast Cancer*. 2019;19(3):216-224. doi:10.1016/j.clbc.2019.02.008
34. Swain SM, Schneeweiss A, Gianni L, et al. Incidence and management of diarrhea in patients with HER2-positive breast cancer treated with pertuzumab [published correction appears in Ann Oncol. 2018 Apr 1;29(4):1075] [published

correction appears in Ann Oncol. 2018 Jul 1;29(7):1607] [published correction appears in Ann Oncol. 2019 Aug 1;30(8):1404]. *Ann Oncol*. 2017;28(4):761-768. doi:10.1093/annonc/mdw695

35. Hussain N, Said ASA, Khan Z. Safety Assessment of Neoadjuvant Pertuzumab Combined with Trastuzumab in Nonmetastatic HER2-Positive Breast Cancer in Postmenopausal Elderly Women of South Asia. *Int J Breast Cancer*. 2018;2018:6106041. Published 2018 Apr 19. doi:10.1155/2018/6106041
36. Tanaka S, Iwamoto M, Kimura K, et al. Phase II Study of Neoadjuvant Anthracycline-Based Regimens Combined With Nanoparticle Albumin-Bound Paclitaxel and Trastuzumab for Human Epidermal Growth Factor Receptor 2-Positive Operable Breast Cancer. *Clin Breast Cancer*. 2015;15(3):191-196. doi:10.1016/j.clbc.2014.12.003
37. Marsh K, IJzerman M, Thokala P, et al. Multiple Criteria Decision Analysis for Health Care Decision Making--Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19(2):125-137. doi:10.1016/j.jval.2015.12.016
38. 1000minds: Decision-making software https://www.1000minds.com/ accessed on: 08/26/2020
39. R package MCDA https://cran.r-project.org/web/packages/MCDA/MCDA.pdf accessed on: 11/03/2020
40. Tervonen T, Naci H, van Valkenhoef G, et al. Applying Multiple Criteria Decision Analysis to Comparative Benefit-Risk Assessment: Choosing among Statins in

Primary Prevention. *Medical Decision Making*. 2015;35(7):859-871. doi:10.1177/0272989X15587005

41. Hsu JC, Tang DH, Lu CY. Risk-benefit assessment of oral phosphodiesterase type 5 inhibitors for treatment of erectile dysfunction: a multiple criteria decision analysis. *Int J Clin Pract*. 2015;69(4):436-443. doi:10.1111/ijcp.12548
42. Vermersch P, Martinelli V, Pfleger C, et al. Benefit-risk Assessment of Cladribine Using Multi-criteria Decision Analysis (MCDA) for Patients With Relapsing-remitting Multiple Sclerosis. *Clin Ther*. 2019;41(2):249-260.e18. doi:10.1016/j.clinthera.2018.12.015
43. Angelis A, Kanavos P. Value-Based Assessment of New Medical Technologies: Towards a Robust Methodological Framework for the Application of Multiple Criteria Decision Analysis in the Context of Health Technology Assessment. *Pharmacoeconomics*. 2016;34(5):435-446. doi:10.1007/s40273-015-0370-z
44. Zawodnik A, Niewada M. Multiple Criteria Decision Analysis (MCDA) for Health Care Decision Making – overview of guidelines. *J Health Policy Outcomes Res* 2018 DOI: 10.7365/JHPOR.2018.2.4
45. National Cancer Institue: Investigating the cardiac side effects of cancer treatments https://www.cancer.gov/news-events/cancer-currents-blog/2018/cancer-treatment-heart-side-effects accessed on: 11/09/2020
46. When cancwe ia not your only health concern https://www.cancer.net/navigating-cancer-care/older-adults/when-cancer-not-your-only-health-concern accessed on: 11/09/2020

47. Maher M. Current and Emerging Treatment Regimens for HER2-Positive Breast Cancer. *P T*. 2014;39(3):206-212.
48. DeKoven M, Bonthapally V, Jiao X, Ganguli A, Pathak P, Lee WC, Ray S. Treatment pattern by hormone receptors and HER2 status in patients with metastatic breast cancer in the UK, Germany, France, Spain and Italy (EU-5): results from a physician survey. J Comp Eff Res. 2012 Sep;1(5):453-63. doi: 10.2217/cer.12.43. PMID: 24236422
49. Alatawi YM, Hansen RA. Empirical estimation of under-reporting in the U.S. Food and Drug Administration Adverse Event Reporting System (FAERS). Expert Opin Drug Saf. 2017 Jul;16(7):761-767. doi: 10.1080/14740338.2017.1323867. Epub 2017 May 9. PMID: 28447485.
50. Hsu JC, Lin JY, Lin PC, Lee YC. Comprehensive value assessment of drugs using a multi-criteria decision analysis: An example of targeted therapies for metastatic colorectal cancer treatment. PLoS One. 2019 Dec 12;14(12):e0225938. doi: 10.1371/journal.pone.0225938. PMID: 31830075; PMCID: PMC6907782.
51. Priester R. A values framework for health system reform. Health Aff (Millwood). 1992 Spring;11(1):84-107. doi: 10.1377/hlthaff.11.1.84. PMID: 1577390.

CHAPTER 4: CONCLUSION

SUMMARY OF FINDINGS:

Our project characterized the population with cancer and cardiovascular diseases (CVD) both as compared to those only with cancer. We found that mainly patients who were younger and employed were less likely to be diagnosed with CVD. Other results might have been attributable to a lower sample size and thus have lower power. We then evaluated the use of machine learning algorithms in the prediction of cardiovascular risk among cancer patients as compared to conventional regression techniques. The machine learning models used in our study had higher predictive power as compared to the standard regression technique. Out of all the machine learning algorithms that we compared; random forest models for acute and chronic CVD were associated with the highest c-statistic indicating that they were better in predicting CVD risk in cancer patients. We validated these models using internal and external validation techniques. The results of these suspected some overfitting involved in the random forest models. We used these models to then build the web-based applications predicting probability of chronic and acute CVD risk in cancer patients. These applications used the most important predictors that were identified by the random forest algorithm. However, given the overfitting issue, the models might need to be further revised by calibration and acquiring more data to then create applications that can be used in a real-world setting. Future studies in this case to evaluate model calibration would help in further increasing the clinical utility of the model. Incorporating long term outcomes and additional data from longitudinal files might be helpful in increasing the generalizability and thus validity of the models.

We then focused on breast cancer patients and described certain cancer therapy and adverse event related factors with cardiotoxic adverse events. Targeted therapy was a primary suspect in most of the cardiotoxic events with an oral route of administration. Adjusted analyses suggested that patients receiving targeted therapy were more likely to witness a cardiovascular adverse event. However, the clinical significance of thisassociation might be low since the effect size of the odds ratio was small. We also evaluated application of multiple criteria decision analysis (MCDA) model to conduct benefit-risk assessment of HER 2 positive breast cancer therapy regimens. We conducted benefit-risk assessment using MCDA model to score and weight breast cancer therapy regimens to rank the most preferred treatment regimen. Given the benefits and risks (other adverse events along with cardiovascular implications) associated with the therapies in our MCDA model, the therapy regimen with trastuzumab, cyclophosphamide/carboplatin and a taxane was the most preferred regimen. On conducting sensitivity analyses, we found that our model was robust to most of the changes.

We thus implemented newer analytical techniques to evaluate certain cardiovascular outcomes among cancer patients. Based on our study results, the newer techniques showed potential to be used in real-world practice. Due to data limitations, our models might not be ready to be used directly in the clinical practice and in the real world, although they do suggest that with further revisions, these might help in changing the current landscape. With more data getting available with each day, techniques like machine learning can handle the 'Big Data' more efficiently than the standard regression approaches. Research in oncology treatment planning has also been growing and a

decision analytic tool like MCDA model can help in making decision-making easier as more options get available. Future research can be encouraged in these fields of machine learning and decision analytics to make healthcare more efficient.

Our study aim 1 has multiple implications. Firstly, characterizing the population that has cancer and CVD both would help in understanding the underlying factors responsible for putting cancer patients at a higher risk for developing CVD and managing the condition more efficiently. Secondly, the use of machine learning algorithms in predicting CVD risk would help in identifying future risk of CVD early in cancer patients. Machine learning algorithms as suggested above have a good predictive power as compared to the regression approaches. Based on the current predictors, the trained model might be able to predict a 5 or 10 year CVD risk which can then be utilized by physicians to plan the treatment better. Based on the probability of the risk the physicians might consider involving a multidisciplinary team with a cardiologist to monitor the cardiac health of cancer patients better. It would help in deciding the cancer treatment regimen to reduce the risk of acute and chronic CVD events. It would also help in encouraging prophylactic CVD care in cancer patients. Currently, the physicians are starting to get more worried about the cardiovascular health of cancer patients while planning treatment since the mortality rates due to cardiac conditions is increasing in cancer patients.[1 - 4] Given the concerns, a predictive model like ours with revisions made to it can help in understanding the risk earlier in time and planning the treatment accordingly.

Our study aim 2 also has multiple implications for breast cancer patients. Identifying cancer therapy and population characteristics associated with cardiotoxic adverse events in breast cancer patients would help in preventing these adverse events in the future. The

predictors we identified to be significant such as targeted therapy and route of administration can be paid more attention to in mitigating the cardiovascular implications on cancer patients. Our MCDA model would help in choosing a specific breast cancer therapy regimen given the benefits and the risks associated. MCDA modeling being a transparent technique (stating criteria and the weights upfront) of consolidating outcomes, it would help in reducing the information asymmetry and guide the decision-making in a formal and timely manner. Cancer treatment planning requires a multidisciplinary approach. A MCDA model would give a guided approach to all the providers on the multidisciplinary team to ease communication and come up with a treatment plan more efficiently. Currently there is no guided decision-making approach to cancer treatment planning, more studies like ours can help in creating one. Our model includes the criteria that we had evidence on, although in the future more criteria can be included in this model that might be necessary in decision making to increase the utility of the model that can then be used in actual practice.

FUTURE RESEARCH:

In our study, we developed machine learning models to predict the risk of CVD in cancer patients. Our model was validated using internal and external validation techniques. Our external validation results suggested that there could have been some overfitting involved with the random forest model. Future studies can be conducted to assess model calibration to further increase the reliability, validity and clinical utility of the model.[5 - 7] Calibration refers to the agreement of predicted probabilities of a model and observed outcomes. It thus evaluates if patients who were predicted to have an event did actually have an event down the line. Calibration techniques can help in increasing the

validity of the model and the confidence in the predictive power of the model. There are multiple calibration techniques such as calibration curves, cost and goodness-of-fit functions. Calibration curves are plots of the observed frequency versus the predicted frequency, how well these curves overlap is a measure of well calibrated model.[8] Cost functions on the other hand are the 'distance' between the observed and the predicted values. Higher the cost function higher is the inaccuracy.[9] The goodness-of-fit criteria comprises of measures such as sum of squared errors [SSE] and Pearson chi-square which quantify the difference between model and observed outcomes.[10] Model calibration can thus be conducted using any of these techniques. This can be achieved by conducting a longitudinal long-term study using healthcare data to evaluate observed outcomes in the same patients that were predicted using the model. In order to increase the validity of the model, further hypertuning of parameters can be carried out by changing the learning rate, number of bootstraps, etc.[11] In addition to these algorithms, literature has also suggested good performance of some other algorithms such as support vector machines (SVM).[12] SVM is a linear model for classification and regression problems. The algorithm creates a line or a hyperplane which separates the data into classes. These can also be built and compared against the algorithms already tested in our study. We can also build weighted predictive models for predicting CVD risk by using recursive partitioning for modeling survey (RPMS) data techniques that can then account for complex sampling survey weights.[13] These RPMS models allow to input survey weights independent of the predictors. With further research these can be used as predictive models too.

Our MCDA model has ranked preferences for the breast cancer therapy regimens based on their benefits and risks involved. Further studies can be conducted to evaluate

patient preferences involved in these therapy regimens and incorporate those in the MCDA model too. A sample or an expert panel can be used to rank the criteria importance and assign weights other than those include in the sensitivity analyses. The criteria can be provided as a survey to a sample of physicians to assess the value they would give to each criteria and use those as weights. Larger longitudinal databases such as claims files can also be used to evaluate the long-term effects of these therapy regimens to be incorporated into the model to make the results more generalizable and robust. We used criteria in our model that we had available evidence on from the literature. However, if we elicit more criteria to be deemed important from the physician perspective, in the future we can try to obtain information on these criteria through database analysis and revise the model further. Based on evidence, we can also try to include more treatment alternatives in the model. A choice based decompositional approach such as a discrete choice experiment could also be used in the future to identify more criteria that would be essential in decision-making and make the model more robust. In a discrete choice experiment, the study sample is provided with the therapy alternatives to choose from. Based on which alternative they chose, weights are derived for all the criteria they considered important while taking that decision.[14] This can be conducted to elicit patient and provider preferences both which can then be included in the model to conduct a benefit-risk assessment to rank the therapy regimen alternatives using evidence-based approach.

REFERENCES:

1. Liu D, Ma Z, Yang J, et al. Prevalence and prognosis significance of cardiovascular disease in cancer patients : a population-based study. 2019;11(18):7948-7960.
2. Lin PL, Hao Y, Xie J, et al. Physician experiences and preferences in the treatment of HR+/HER2- metastatic breast cancer in the United States: a physician survey. *Cancer Med*. 2016;5(2):209-220. doi:10.1002/cam4.580
3. National Cancer Institue: Investigating the cardiac side effects of cancer treatments https://www.cancer.gov/news-events/cancer-currents-blog/2018/cancer-treatment-heart-side-effects accessed on: 11/09/2020
4. When cancwe ia not your only health concern https://www.cancer.net/navigating-cancer-care/older-adults/when-cancer-not-your-only-health-concern accessed on: 11/09/2020
5. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12-22. doi:10.1016/j.jclinepi.2019.02.004
6. Gravesteijn BY, Nieboer D, Ercole A, et al. Machine learning algorithms performed no better than regression models for prognostication in traumatic brain injury. *J Clin Epidemiol*. 2020;122:95-107. doi:10.1016/j.jclinepi.2020.03.005
7. Model Calibration http://www.coastalwiki.org/wiki/Model_calibration accessed on: 11/09/2020
8. Machine Learning Mastery: How and when to use a calibrated classification model with scikit-learn https://machinelearningmastery.com/calibrated-classification-

model-in-scikit-learn/ accessed on:11/09/2020

9. Kaggle: Classification probability calibration https://www.kaggle.com/residentmario/notes-on-classification-probability-calibration accessed on:11/09/2020
10. Van der Steen A, van Rosmalen J, Kroep S, van Hees F, Steyerberg EW, de Koning HJ, van Ballegooijen M, Lansdorp-Vogelaar I. Calibrating Parameters for Microsimulation Disease Models: A Review and Comparison of Different Goodness-of-Fit Criteria. Med Decis Making. 2016 Jul;36(5):652-65. doi: 10.1177/0272989X16636851. Epub 2016 Mar 8. PMID: 26957567.
11. DataCamp: Machine Learning with caret in R
12. Gopi Battineni, Nalini Chintalapudi, Francesco Amenta, Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM), Informatics in Medicine Unlocked, Volume 16, 2019, 100200, ISSN 2352-9148, https://doi.org/10.1016/j.imu.2019.100200.
13. CRAN: Recursive Partitioning for Modeling Survey Data https://cran.r-project.org/web/packages/rpms/index.html accessed on: 11/09/2020
14. YHEC: Discrete Choice Experiment https://yhec.co.uk/glossary/discrete-choice-experiment accessed on : 11/09/2020