

University of Massachusetts Boston

## ScholarWorks at UMass Boston

---

Graduate Doctoral Dissertations

Doctoral Dissertations and Masters Theses

---

5-31-2016

# External Validation and Cost Effectiveness Analysis of the Non-LB Framingham Cardiovascular Disease Risk Assessment Algorithm in the Atherosclerosis Risk in Communities Dataset

Jacob K. Kariuki

*University of Massachusetts Boston*

Follow this and additional works at: [https://scholarworks.umb.edu/doctoral\\_dissertations](https://scholarworks.umb.edu/doctoral_dissertations)



Part of the [Epidemiology Commons](#), and the [Health Services Administration Commons](#)

---

### Recommended Citation

Kariuki, Jacob K., "External Validation and Cost Effectiveness Analysis of the Non-LB Framingham Cardiovascular Disease Risk Assessment Algorithm in the Atherosclerosis Risk in Communities Dataset" (2016). *Graduate Doctoral Dissertations*. 280.

[https://scholarworks.umb.edu/doctoral\\_dissertations/280](https://scholarworks.umb.edu/doctoral_dissertations/280)

This Open Access Dissertation is brought to you for free and open access by the Doctoral Dissertations and Masters Theses at ScholarWorks at UMass Boston. It has been accepted for inclusion in Graduate Doctoral Dissertations by an authorized administrator of ScholarWorks at UMass Boston. For more information, please contact [scholarworks@umb.edu](mailto:scholarworks@umb.edu).

EXTERNAL VALIDATION AND COST EFFECTIVENESS ANALYSIS OF THE  
NON-LB FRAMINGHAM CARDIOVASCULAR DISEASE RISK ASSESSMENT  
ALGORITHM IN THE ATHEROSCLEROSIS RISK IN COMMUNITIES DATASET

A Dissertation Presented

by

JACOB K. KARIUKI

Submitted to the Office of Graduate Studies,

University of Massachusetts Boston,

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2016

Nursing program

© 2016 by Jacob K. Kariuki

All rights reserved

EXTERNAL VALIDATION AND COST EFFECTIVENESS ANALYSIS OF THE  
NON-LB FRAMINGHAM CARDIOVASCULAR DISEASE RISK ASSESSMENT  
ALGORITHM IN THE ATHEROSCLEROSIS RISK IN COMMUNITIES  
DATASET

A Dissertation Presented

by

JACOB K. KARIUKI

Approved as to style and content by:

---

Eileen Stuart-Shor, Associate Professor  
Chairperson of Committee

---

Suzanne Leveille, Professor  
Member

---

Philimon Gona, Associate Professor  
Member

---

Jerry Cromwell, Lecturer  
Member

---

Haeok Lee, Program Director  
Nursing Program

---

Rosanna DeMarco, Chairperson  
Nursing Department

## ABSTRACT

# EXTERNAL VALIDATION AND COST EFFECTIVENESS ANALYSIS OF THE NON-LB FRAMINGHAM CARDIOVASCULAR DISEASE RISK ASSESSMENT ALGORITHM IN THE ATHEROSCLEROSIS RISK IN COMMUNITIES DATASET

May 2016

Jacob K. Kariuki, BSN., University of Eastern Africa Baraton  
MS., University of Massachusetts Boston  
PhD., University of Massachusetts Boston

Directed by Professor Eileen Stuart-Shor

**Background:** In recent years, non-Laboratory based (non-LB) risk assessment algorithms have been developed to facilitate absolute cardiovascular disease (CVD) risk assessment in resource constrained primary care settings. The non-LB Framingham algorithm, which substitutes body mass index (BMI) for lipids, has the best discrimination and calibration among the published algorithms, but its external validity and cost-effectiveness have not been determined.

**Purpose:** External validation and comparative effectiveness analysis of the non-LB versus laboratory based (LB) Framingham algorithm in a racially diverse population, and simulated cost-effectiveness analysis focusing on a black sample.

Methods: Secondary data analysis was performed using the Atherosclerosis Risk in Communities (ARIC) dataset. Cox regression models including the non-LB and LB Framingham covariates were developed. Model discrimination was assessed using the C statistic, calibration using the goodness-of-fit test, and equivalence of regression coefficients using the z-test. Algorithms based on the models were developed and their performance assessed using the area under receiver operating characteristic curve (AUROC), and agreement using kappa statistics. Analyses using simulated incremental cost-effectiveness ratios (ICER) were focused on the black sample. IRB approval was obtained. Data were analyzed using Stata<sup>®</sup> software version 14.

Results: Among 11,601 individuals (mean age  $53.9 \pm 5.7$  years, 55% female, 24% black), the non-LB versus LB models performed as follows: C statistic (0.75 vs 0.76 for women, & 0.67 vs 0.68 for men); goodness-of-fit (14.2 vs 10.5 for women, & 25.8 vs 21.8 for men) respectively. In the black sample, regression coefficients of all covariates were similar to those generated in Framingham ( $z = \pm 1.96$ ). The two algorithms based on the models had a kappa statistic of 0.76. When used to stratify risk in the entire ARIC sample, the non-LB and LB Framingham algorithms had AUROC of 0.706 vs 0.710 respectively. Prevention program guided by the non-LB Framingham dominated those guided by individual risk factors and LB Framingham algorithm.

Conclusions: These results demonstrate the validity and cost-effectiveness of the non-LB Framingham algorithm. This approach could provide a valuable and efficient alternative to the traditional LB approaches in the ongoing efforts to address the high burden of CVD in underserved communities especially the US black population.

## ACKNOWLEDGEMENTS

This work was supported by an award from the American Heart Association. The dissertation was prepared using ARIC Research Materials obtained from the NHLBI Biologic Specimen and Data Repository Information Coordinating Center and does not necessarily reflect the opinions or views of the ARIC or the NHLBI.

## DEDICATION

I dedicate this work to the memory of my late grandmother, Wanjiru Kigo,  
who taught me the essence of life essentials.

To the love of my life, Kageha, for all the sacrifices you made.

To my parents and many siblings for the incredible support I can't even begin to quantify.

To my church community and Almighty God, for the inspiration to a life with a purpose.

I wish to thank my committee members Drs. Eileen Stuart-Shor, Suzanne Leveille, Phil Gona and Jerry Cromwell for their extraordinary support and commitment to bring the best out of me. A special thanks to Dr. Eileen Stuart-Shor, my committee chairperson for her mentorship and encouragement throughout my graduate training. Thank you to my professors, classmates and friends who inspired me to keep questioning even when the answers are not forthcoming.

## TABLE OF CONTENTS

ACKNOWLEDGMENTS/ DEDICATION.....	vi
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xiv
LIST OF ABBREVIATIONS.....	xviii
CHAPTER	Page
1. INTRODUCTION .....	1
Goals of the Study.....	4
Significance and Innovation .....	5
Defining Key terms.....	5
Conceptual Framework .....	6
The Conceptual Theoretical and Empirical (CTE) structure.....	10
2. LITERATURE REVIEW .....	12
Background: absolute CVD risk assessment .....	12
The impact of CVD in resource constrained settings .....	17
Preventing CVD in resource constrained settings .....	20
Summary .....	21



3. RESEARCH DESIGN AND METHODS .....	22
Overview .....	22
Baseline assessment .....	25
Follow-up assessment .....	26
Statistical analyses .....	28
Analysis plan by study aims .....	29
Cost-effectiveness analysis .....	31
Case example: individual vs absolute CVD risk approach ....	37
Absolute versus individual risk decision model .....	38
Cost-effectiveness framework .....	41
Interpreting the ICER.....	42
4. RESULTS .....	44
a) Sample description.....	45
Incident CVD during 12 years of follow-up .....	47
b) Mathematical models' performance .....	49
The unadjusted non-LB Framingham models.....	50
The unadjusted LB Framingham models .....	54
The adjusted non-LB models .....	58
The adjusted non-LB models .....	60

c) Applicability of the Framingham algorithms in the ARIC sample .....	62
CVD risk stratification using the frozen Framingham algorithms .....	62
CVD risk stratification using the recalibrated algorithms .....	67
CVD risk stratification using the recalibrated algorithms .....	69
d) Sensitivity/specificity analysis .....	71
Sensitivity/specificity analysis of the frozen Framingham algorithms .....	71
The frozen non-LB Framingham algorithm.....	72
The frozen LB Framingham algorithm .....	75
Sensitivity/specificity analysis of the recalibrated algorithms .....	78
The recalibrated non-LB algorithm .....	78
The recalibrated LB algorithm.....	79
Sensitivity/specificity analysis of the adjusted algorithms .....	81
The adjusted non-LB algorithm .....	81
The adjusted LB algorithm .....	82
Sensitivity/specificity analysis of all algorithms .....	84

e) Simulated cost-effectiveness analysis of the non-LB	
Framingham algorithm .....	86
Introduction .....	86
Level I: Expenses in screening for CVD .....	86
Level II: Expenses of initial and follow-up visits for positive cases .....	89
Costs associated with initial office visits .....	93
Costs associated with follow-up office visits .....	94
Costs associated preventive interventions .....	98
The Individual CVD risk factors strategy treatment expenses .....	98
The non-LB absolute CVD risk strategy treatment costs: .....	101
The LB absolute CVD risk strategy treatment costs: .....	106
Summary of level I and II expenses .....	110
Level III: Expenses in treating false negatives cases .....	110
Summary of level III expenses .....	113
Outcomes associated with the three prevention programs .....	113
Average and incremental cost effectiveness .....	116
The average cost-effectiveness ratio .....	116
The incremental cost-effectiveness ratio .....	117

5. DISCUSSION .....	119
Social determinants of health and CVD risk assessment .....	119
Incidence of CVD by sex and race .....	120
Performance of the non-LB Framingham CVD risk assessment algorithm .....	121
Comparative performance in risk stratification .....	126
Cost-effectiveness of non-LB Framingham Algorithm .....	129
Possible relevance to sub-Saharan Africa (SSA) .....	132
Conclusion .....	134
REFERENCES .....	135

## LIST OF TABLES

Table	Page
1. Sample characteristics and end points used in Framingham versus ARIC cohorts .....	16
2. WHO choice & intensity of CVD prevention strategy guided by absolute CVD risk.....	36
3.1 Baseline characteristics stratified by eligibility criteria.....	46
3.2 Baseline characteristics of the eligible sample stratified by sex .....	46
4. Regression beta coefficients of the unadjusted non-LB Framingham in ARIC and z test score of equality.....	52
5. Regression beta coefficients of the unadjusted LB Framingham in ARIC and z test score of equality.....	55
6. Regression beta coefficients of the adjusted non-LB model stratified by sex and race.....	59
7. Regression beta coefficients of the adjusted LB model stratified by sex and race.....	61

8. Kappa test frozen non-LB versus LB Framingham algorithm's risk categories .....	65
9. Sensitivity/specificity analysis of the frozen non-LB Framingham algorithm .....	72
10. Sensitivity/specificity analysis of the frozen LB Framingham algorithm .....	75
11. Sensitivity/specificity analysis of recalibrated non-LB algorithm .....	78
12. Sensitivity/specificity analysis of recalibrated LB algorithm .....	80
13. Sensitivity/specificity analysis of the adjusted non-LB algorithm .....	81
14. Sensitivity/specificity analysis of the adjusted LB algorithm .....	83
15. Marginal screening costs: individual risk factors vs absolute non-LB vs LB absolute risk approach .....	89
16. Sensitivity/specificity analysis of the individual risk factors strategy .....	90
17. Sensitivity/specificity analysis of the non-LB guided absolute CVD risk strategy .....	91

18. Sensitivity/specificity analysis of the LB guided absolute CVD risk strategy .....	92
19. Discounted costs of follow-up office visits: individual risk factors vs absolute non-LB vs LB absolute risk approach.....	97
20. Discounted costs of preventive interventions: individual risk factors vs absolute non-LB vs LB absolute risk approach.....	109
21. Initial & follow-up costs for treating false negative cases by prevention strategy .....	112
22. True positive cases stratified by CVD prevention program.....	115

## LIST OF FIGURES

Figure	Page
1. The Social Ecological & Chronic Care Model adapted to focus on CVD .....	11
2. Flowchart of inclusion and exclusion criteria and monitored outcomes .....	24
3. CVD prevention strategy based on treating diabetes/hypertension to target goals .....	34
4.1 Events trajectories associated with three CVD prevention strategies .....	40
4.2 Cost-effectiveness plane depicting acceptability of ICER ratios .....	43
5. General & cause specific CVD incidence rate in ARIC stratified by sex & race .....	48
6. General & case specific CVD incidence proportion in ARIC stratified by sex & race .....	49
7. Calibration- unadjusted non-LB based Framingham women specific model .....	53
8. Calibration- unadjusted non-LB based Framingham men specific model .....	54



9. Calibration- unadjusted LB Framingham women specific model .....	57
10. Calibration- unadjusted LB Framingham men specific model .....	57
11. Risk stratification in ARIC cohort as per the Frozen non-LB Framingham .....	64
12. Risk stratification in ARIC cohort as per the Frozen LB Framingham .....	64
13. Frozen non-LB Framingham algorithm risk categories stratified by sex & race .....	66
14. Frozen LB Framingham algorithm risk categories stratified by sex & race .....	67
15. Recalibrated non-LB algorithm risk categories stratified by sex & race .....	68
16. Recalibrated LB algorithm risk categories stratified by sex & race .....	69
17. Adjusted non-LB algorithm risk categories stratified by sex & race .....	70
18. Adjusted LB algorithm risk categories	

stratified by sex & race .....	71
19. Frozen non-LB Framingham AUROC for entire ARIC sample.....	73
20. Frozen non-LB Framingham AUROC white women .....	74
21. Frozen non-LB Framingham AUROC black women .....	74
22. Frozen non-LB Framingham AUROC white men .....	74
23. Frozen non-LB Framingham AUROC black men .....	74
24. Frozen LB Framingham AUROC for entire ARIC sample .....	76
25. Comparing frozen Framingham algorithms' AUROCs in entire ARIC sample.....	76
26. Frozen LB Framingham AUROC white women .....	77
27. Frozen LB Framingham AUROC black women .....	77
28. Frozen LB Framingham AUROC white men .....	77
29. Frozen LB Framingham AUROC black men .....	77
30. Recalibrated non-LB Framingham AUROC for entire ARIC sample.....	79

31. Recalibrated LB Framingham AUROC for entire ARIC sample.....	80
32. Adjusted non-LB algorithm AUROC for entire ARIC sample.....	82
33. Adjusted LB algorithm AUROC for entire ARIC sample.....	83
34. Comparing all non-LB algorithms .....	85
35. Comparing all LB algorithms .....	85
36. Comparing frozen LB with all non-LB based algorithms .....	85
37. Comparing frozen non-LB with all LB based algorithms .....	85

## LIST OF ABBREVIATIONS

CVD: Cardiovascular disease

LB: Laboratory

ICER: Incremental cost-effectiveness ratio

SSA: sub-Saharan Africa

## CHAPTER 1

### INTRODUCTION

Cardiovascular disease (CVD), including coronary heart disease and stroke, is now the leading cause of death globally due in part to the ongoing epidemiological transition from infectious to non-communicable diseases in developing countries (WHO, 2015). Currently over 75% of all CVD deaths occur in developing countries where CVD is taking toll on populations in resource constrained settings who rely on under-developed health care systems that are traditionally invested in treating infectious diseases (Mensah, 2008; WHO, 2015). Consequently, most of the CVD mortality and morbidity occur at younger ages in these countries (S. Mendis et al., 2007; WHO, 2015).

In developed countries such as the United States (US), underserved racial and ethnic minorities bear the highest burden of CVD (Mozaffarian et al., 2015). The US black population continues to be disproportionately affected by CVD related morbidity and mortality. For instance, the age-adjusted mortality attributable to CVD is approximately 34% higher in the black population compared to the overall US population (Mozaffarian et al., 2015). Although factors leading to these disparities are complex, barriers related to access of preventive and curative treatments are known to play a prominent role (Institute of Medicine, 2001).

The public health and socioeconomic ramifications of CVD, especially on the poor and underserved populations in both developed and developing countries require pragmatic and robust preventive initiatives. Feasible strategies that improve access and quality of CVD preventive treatments are necessary to address the burden of CVD in resource constrained settings. If well implemented, such strategies may promote cardiovascular health and economic progress of minority groups in developed countries, as well as the economically deprived populations in developing countries.

Contemporary CVD management guidelines recommend absolute risk assessment as a clinically sound guide to CVD prevention and risk surveillance (Cooney, Dudina, & Graham, 2009; World Health Organization, 2007). Absolute CVD risk, also known as total or global risk, denotes the probability that an individual will develop CVD within a given time frame, depending on the combination and severity of the risk factors present (Jilcott et al., 2007). To facilitate absolute CVD risk assessment, dozens of algorithms have been developed to predict the likelihood that a particular constellation of risk factors will contribute to occurrence of CVD related morbidity or mortality over a specific period of time (Hayman, Helden, Chyun, & Braun, 2011; D. M. Lloyd-Jones, 2010a).

For many years the available absolute CVD risk assessment algorithms were based on laboratory measures which are not readily available in resource constrained settings or for individuals with limited access to care (Beswick, Brindle, Fahey, & Ebrahim, 2008; Gaziano, Young, Fitzmaurice, Atwood, & Gaziano, 2008). However, in recent years progress has been made in developing non-LB algorithms, a move that may be helpful for management of CVD in resource constrained settings.

A recent systematic review of literature reported that of the five published non-LB risk assessment algorithms for primary prevention of CVD, the non-LB Framingham algorithm had the best sensitivity and specificity ratios (Kariuki, Stuart-Shor, Leveille, & Hayman, 2013). In the high risk category (ten-year risk threshold of 20%) the algorithm had sensitivity/specificity ratios of 0.48/0.85 and 0.58/0.83 for men and women respectfully. These sensitivity/specificity ratios were comparable to the established LB Framingham algorithm (0.49/0.85 and 0.60/0.84) for men and women respectfully (D'Agostino RB et al., 2008).

Despite the solid performance of the non-LB Framingham algorithm in its predominantly white (99.7%) derivation dataset, its performance and applicability in multiracial and black populations has not been tested. External validation is considered to be an essential process of testing the applicability of an algorithm to diverse populations with baseline characteristics which differ from those in the algorithm's derivation dataset (Cooney et al., 2009). Without external validation, the suitability of the algorithm beyond the Framingham population remains uncertain.

This study performed external validation of the non-LB Framingham algorithm in the multiracial Atherosclerotic Risk in Communities (ARIC) dataset (23% black). The external validation focused on evaluation of the algorithm's ability to optimally stratify CVD risk and predict cardiovascular events in the multiracial population that forms the ARIC dataset. A sub-analysis focusing on the black participants enlisted in the ARIC study assessed the performance of the algorithm in a group that bears the highest burden of CVD, and more likely to reside in resource constrained or underserved settings in the

US. A subsequent cost effectiveness analysis evaluated the costs and benefits associated with using the algorithm in guiding prevention of CVD in the black participants enlisted in the ARIC study. Race was self-reported and individuals who report black race or white race will hereafter be referred to as “blacks” and “whites” respectively.

### Goals of the Study

The main purpose of this study was to externally validate and cost the non-LB Framingham algorithm by testing four hypotheses to achieve three aims: -

Specific aim 1: Assess the accuracy of the non-LB Framingham algorithm in stratifying risk and predicting CVD events a racially diverse population.

Hypothesis 1: Non-LB Framingham algorithm will have adequate discrimination (Harrell’s C statistic greater than 0.75) and calibration (Hosmer-Lemeshow goodness of fit ( $\chi^2$ ) below 20 ( $p > 0.05$ ) in the multiracial ARIC dataset.

Specific aim 2: Compare the performance of the non-LB Framingham algorithm in black versus white participants of the ARIC study.

Hypothesis 2: There will be no significant difference in discrimination and calibration of the non-LB Framingham algorithm between the black and white participants of the ARIC study.

Specific aim 3: Establish the cost feasibility of using the non-LB Framingham algorithm in guiding prevention of CVD among the black participants enlisted the ARIC study.

Hypothesis 3: A CVD prevention strategy guided by the non-LB Framingham algorithm will be more cost-effective compared with treating each elevated CVD risk



factor (diabetes and/or hypertension) independently in the black subset of the ARIC cohort.

Hypothesis 4: A CVD prevention strategy guided by the non-LB Framingham algorithm will be more cost-effective compared to a strategy guided by the LB Framingham algorithm in the black subset of the ARIC cohort.

#### Significance and Innovation

Validating and costing the non-LB Framingham algorithm was an important step in availing a risk assessment tool that could guide CVD prevention in resource constrained settings. The high representation of blacks in the ARIC cohort (23%) enabled adequate evaluation of the algorithm's performance in this population that has the highest rates of CVD in the US.

#### Defining Key terms

In this study, the performance of the non-LB Framingham algorithm is assessed through external validation. External validation is the assessment of the performance of an algorithm in an external dataset. The external validation process is considered an essential step in assessing transportability of an algorithm to different populations because baseline survival and risk factors used in the test are not a perfect match for those in the algorithm's derivation dataset (Cooney et al., 2009). The main approaches for measuring the performance include discrimination, and calibration.

Discrimination is the ability of an algorithm to assign a higher risk to those who will develop the end point and a lower score to those who will not, and it is frequently measured using area under Receiver Operating Characteristic curve (AUROC) or

Harrell's C statistic. AUROC or C statistic of 1 denotes perfect discrimination whereas 0.5 equates to chance discrimination. Although the C statistic of CV risk assessment algorithms rarely exceeds 0.8, a valid algorithm should have a C statistic of 0.75 or higher (Cooney et al., 2009; May, Lawlor, Brindle, Patel, & Ebrahim, 2006). In addition, threshold discrimination operationalized by sensitivity and specificity is used to define low/high risk populations and treatment decisions are made in reference to this threshold (Cooney et al., 2009).

Calibration is a measure of the agreement between the predicted outcomes and observed outcomes. It is frequently assessed using Hosmer-Lemeshow goodness of fit testing ( $\chi^2$ ). Goodness of fit ( $\chi^2$ ) values below 20 (lack of fit,  $p > 0.05$ ) are considered good fit (Cooney et al., 2009).

### Conceptual Framework

The proposed study was guided by the Social Ecological and Chronic Care Models which are combined and adapted to provide an organizing structure for testing the validity and cost-effectiveness of the non-LB Framingham algorithm (see Figure 1).

Various socioecological models were developed after the First World War to expand understanding of the dynamic relationship between various personal and environmental factors. In 1991 Dahlgren and Whitehead published the Social Ecological Model to enhance understanding of policies and strategies to promote social equity in health. They contended that policies and strategies focusing on health equity should be based on a clear understanding of factors that threaten, promote or protect health (Dahlgren & Whitehead, 1991).

In the Social Ecological framework, the major factors that influence health are organized in hierarchical layers which include: macro-socioeconomic environment, living and work conditions, social and community networks, lifestyle choices, and genetical/constitutional factors (Dahlgren & Whitehead, 1991). The postulated influence of these interactive layers resonates with current knowledge on cardiovascular health trajectory which is known to be influenced by the interrelation between personal and environmental factors over an individual's lifetime (Hayman et al., 2011; Stuart-Shor, Berra, Kamau, & Kumanyika, 2012).

Personal factors, which include genetics and lifestyle, form the core of the model are affected by, and affect the social determinants of health which are espoused in the three outer layers of the Social Ecological model (see Figure 1). The social determinants of health include life improving resources such as food supply, education, and social relationships, and their distribution across populations is well known to impact the health trajectory (Will, Keydron, Cynthia, Luis, & Zachary, 2011).

The Chronic Care Model was developed by Dr. Edward Wagner and colleagues as part of the “Improving Chronic Illness Care initiative” supported by the Robert Wood Johnson Foundation (Bodenheimer, Wagner, & Grumbach, 2002). The initiative sought to develop innovations in primary care that would help close the quality gaps described in the 2001 Institute of Medicine report titled: *Crossing the Quality Chasm: A New Health System for the 21st Century*. In the report, the Institute of Medicine detailed many quality problems that caused a huge gap between current practices and attainable optimal chronic illness care in the US (Institute of Medicine, 2001).

To address these problems, the Chronic Care Model identifies the entire community, health care systems, and provider organizations as the three galaxies where chronic illness care occur (Bodenheimer et al., 2002). The three galaxies overlap and encompass six essential elements or pillars which may undermine or promote chronic illness care. The six pillars include: community resources and policies, self-management support, health care organization, delivery system design, decision support, and clinical information systems (Barr et al., 2003; Bodenheimer et al., 2002). Improvements in these six essential and interrelated pillars are expected to produce reformed health care systems in which informed, activated patients interact with prepared, proactive health care providers (Bodenheimer et al., 2002).

Although the six pillars of the Chronic Care Model are interrelated, only the decision support pillar is directly relevant to the objectives of this study as outlined in Figure 1. Therefore, the discussion of the other five pillars is beyond the scope of this study. The decision support pillar calls for integration of evidence based guidelines in routine protocols to help clinicians in making prudent clinical decisions in management chronic illness (Bodenheimer et al., 2002).

Decision support is not intended to substitute individualized clinical judgement, but to support it by providing real time essential data on the patient or available evidence based interventions to the clinician. Optimal decision support tools may also promote self-management by making complex concepts more concrete and comprehensible to the patient; hence increasing risk awareness and motivation to adhere to risk reduction interventions.

The value of any tool used to support clinical decisions depends on the extent to which it is valid and applicable to the relevant clinical practice. Generally clinicians are more likely to use decision support tools that are not only valid, but also quick and easy to use (Cooney et al., 2009; Gaziano et al., 2008). The need for valid but simple, user friendly decision support tools is even more acute in resource constrained settings where non-physician health workers are increasingly being entrusted with traditionally physician responsibilities such as screening for and managing CVD.

Absolute CVD risk assessment algorithms are considered as valid decision support tools appropriate for guiding CVD risk assessment and management. Consequently, the algorithms are currently used in many developed countries to support clinical decisions on CVD management by providing guidance on risk stratification and selection of treatment intensity (D. M. Lloyd-Jones, 2010b). In resource constrained settings, these algorithms are rarely used because they require laboratory measures that are usually inaccessible in these settings. Therefore, validating and costing the non-LB Framingham algorithm was an important step in availing a tool that would support implementation of evidence based guidelines in routine management of CVD.

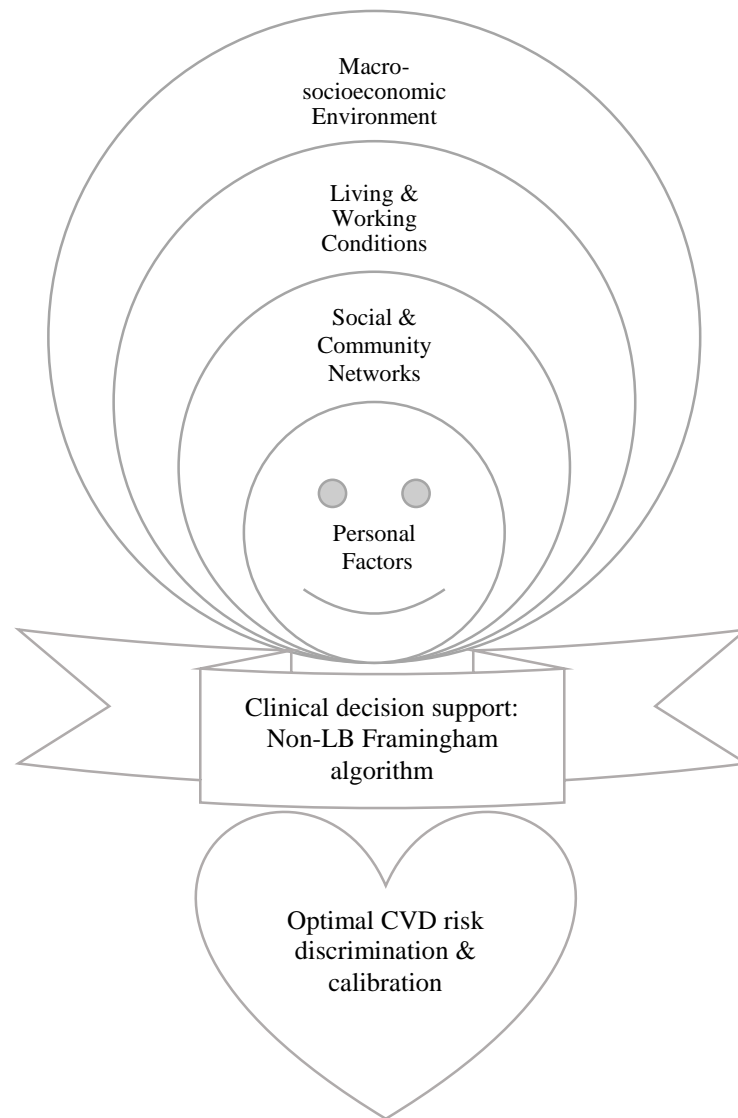
The absolute risk assessment algorithms use some covariates which are influenced by the dynamic relationship between personal and environmental factors included in the Social Ecological Model. Equipping healthcare providers with a decision support tool that enables them to have a comprehensive view of factors that threaten, promote or protect cardiovascular health is expected to make them well prepared and proactive in prevention and management of CVD. Individuals and populations served by such

proactive providers will have improved access to timely CVD risk assessment, increased risk awareness, and motivation to adherence.

*The Conceptual Theoretical and Empirical (CTE) structure*

The Social Ecological model's proposition that a dynamic relationship between personal factors and the social determinants of health dictates the individual's level of risk and subsequent development of disease forms the central concept of the framework. The chronic care model's conceptualization of the importance of clinical decision support in shaping the dynamic relationship between personal factors and social determinants of health in favor of optimal cardiovascular health forms the middle range theory of the framework. Empirical indicators will include discrimination and calibration statistics quantifying the contribution of clinical decision support tool (non-LB Framingham CVD risk assessment algorithm) in detecting individual's level of risk. Quantification of risk is expected to foster risk reduction discussion thereby producing proactive and well prepared healthcare providers and activate patients.

Figure 1: The Social Ecological & Chronic Care Model adapted to focus on CVD



In chapter 1, the concept of absolute CVD risk assessment and the role of valid and feasible risk assessment algorithms were introduced. The external validation process as well as the organizing framework for the study were also presented in the context of CVD. Chapter 2 will focus on the science behind these concepts and existing knowledge gaps pertaining to CVD risk assessment in resource constrained settings.

## CHAPTER 2

### LITERATURE REVIEW

#### Background

Although the morbidity and mortality associated with CVD usually occur in middle and late adulthood, the main pathological pathway leading to CVD begins early in life and progresses cumulatively through adolescence and early adulthood (World Health Organization, 2007). This lifelong cumulative process is influenced by the interaction between constitutional (genetic) and lifestyle factors, with social determinants of health such as education and socioeconomic status as exemplified in the widely published social ecological model (Golden & Earp, 2012; Whitehead & Dahlgren, 1991).

The insidious progression of CVD risk necessitates timely detection and initiation of preventive treatments. The major risk factors known to independently increase the risk of CVD include: cigarette smoking, high blood pressure, dyslipidemia, diabetes mellitus, and advancing age (Grundy, Pasternak, Greenland, Smith, & Fuster, 1999). Predisposing risk factors are known to aggravate the major CVD risk factors and include; obesity, physical inactivity, family history of premature CVD, ethnic characteristics, and psychosocial factors (Grundy et al., 1999). Co-occurrence or clustering of these risk factors is known to compound the effect of individual risk factors increasing the likelihood of developing CVD (World Health Organization, 2007)



Absolute CVD risk assessment algorithms are recommended by contemporary CVD management guidelines to facilitate assessment of “total” or “global” risk. These algorithms take into consideration the clustering of risk factors in an individual to predict their likelihood of experiencing a CVD event within a given time frame, usually 10 years (Beswick et al., 2008). The foremost absolute risk assessment algorithms were derived from the Framingham Heart Study which was inaugurated in 1948 to investigate risk factors associated with development of CVD. At the commencement of the study, the town of Framingham was an industrial trading center in North Eastern United States inhabited by white middle class families (Dawber, Meadors, & Moore, 1951). As a result, the cohort was 99.7% white. In 1971, descendants of the original cohort and their spouses were recruited to form the Framingham offspring cohort with an overarching goal of mapping the familial and genetic determinants of CVD. Similar protocols have been used in the examination of the original and offspring cohorts so as to enable combined analyses (Beswick et al., 2008).

As part of the premier cardiovascular research study, the Framingham cohorts have been instrumental in identifying many CVD risk factors. The identified risk factors have been progressively included in the Framingham based algorithms, enabling significant improvements in risk discrimination and calibration (Beswick et al., 2008). High-density lipoprotein (HDL) cholesterol was identified and incorporated into Framingham risk assessment algorithms as an independent risk factor for CVD in 1968 (Beswick et al., 2008). Traditionally, the Framingham algorithms have required

laboratory measures and were tailored to estimate the 10-year risk of developing coronary heart disease (Cooney et al., 2009).

In an effort to simplify absolute CVD risk assessment, D'Agostino RB et al. (2008) developed a risk prediction model that demonstrated that Body Mass Index (BMI) could effectively substitute total and HDL cholesterol without compromising the robustness of the Framingham model. In the same study, the focus of risk assessment was broadened from a narrow focus on hard coronary events to a broader focus that entailed the full spectrum of CVD.

The 2008 update of the Framingham model includes the simplified non-LB algorithm and the LB algorithm. Both the non-LB and LB algorithms were also modelled to predict general CVD events (coronary heart disease, cerebrovascular events, peripheral artery disease and heart failure). As a result of this broad focus, the absolute CVD risk estimated using the updated 2008 algorithms is significantly higher than when using earlier versions of Framingham model (D'Agostino RB et al., 2008).

The predicted CVD risk ( $\tilde{p}$ ) in both the non-LB and LB algorithms is calculated using the general formulae:  $\tilde{p} = 1 - S_0(t)^{\exp(\sum_{i=1}^p \beta_i X_i - \sum_{i=1}^p \beta_i \bar{X}_i)}$  where  $S_0(t)$  is baseline survival at follow-up time  $t$  (here  $t=12$  years),  $\beta_i$  is the estimated regression coefficient,  $X_i$  is the log-transformed value of the  $i^{\text{th}}$  risk factor, (if continuous),  $\bar{X}_i$  is the corresponding mean, and  $p$  denotes the number of risk factors (D'Agostino RB et al., 2008).

The 2008 non-LB and LB Framingham algorithms use similar covariates (risk factors) except for substitution of BMI for cholesterol in the non-LB model as outlined in Table 1 (D'Agostino RB et al., 2008). This substitution had a benign effect on the model because both the non-LB and LB algorithms had comparable discrimination ( $C=0.749$  vs.  $0.763$  men &  $0.785$  vs.  $0.793$  women) and equally good calibration ( $\chi^2=13.61$  vs.  $13.48$  men &  $10.24$  vs.  $7.79$  women) in their derivation dataset respectively (D'Agostino RB et al., 2008). Although the Framingham cohort is more than 99% white, previous studies have suggested that Framingham risk prediction functions generally perform well in predicting coronary heart death and myocardial infarction among the US black population (D'Agostino RB, Grundy, Sullivan, Wilson, & CHD Risk Prediction Group, 2001b).

In 2011, the 2008 LB Framingham model was tested in the multiracial third National Health and Nutrition Examination Survey (NHANES III) population, where it demonstrated optimal discrimination ( $C=0.776$  men &  $0.834$  women); but calibration was not assessed due to lack of data on clinical end points (Pandya, Weinstein, & Gaziano, 2011). The non-LB Framingham algorithm has been used to assess the effectiveness and impact of simulated national wide CVD screening strategies in Malaysia, but no published external validation studies have been found so far (Kariuki et al., 2013; Selvarajah et al., 2013). Therefore, validating and costing the non-LB algorithm in a multiracial population is an important step in availing a tool that could be instrumental in guiding prevention of CVD in resource constrained settings.

Table 1: Sample characteristics and end points used in Framingham vs ARIC cohorts

Algorithms/Dataset	Sample characteristics	Sex	Age	Smoking	BP	HTN Treatment	Cholesterol or BMI	Diabetes	Endpoints
LB-Framingham (D'Agostino RB et al., 2008).	<u>Design:</u> Prospective study of the Framingham Heart Study and Framingham Offspring cohorts.	M or F	30-74	<ul style="list-style-type: none"> <li>• Yes, self-reported current smoker (<i>must be smoking at the time of assessment</i>)</li> <li>• No, none of the above criteria</li> </ul>	Systolic <i>average of two physician obtained measures</i>	<ul style="list-style-type: none"> <li>• Yes, self-reported current treatment supplemented by physician meds review</li> <li>• No, none of the above criteria</li> </ul>	<ul style="list-style-type: none"> <li>• Total and HDL cholesterol <i>measured by standardized enzymatic methods</i></li> </ul>	<ul style="list-style-type: none"> <li>• Yes, fasting glucose <math>\geq 126</math> mg/dL (offspring cohort) or 140 mg/dL (original cohort) or use of insulin/oral hypoglycemic medications.</li> <li>• No, none of the above criteria</li> </ul>	General CVD (coronary artery, cerebrovascular, and peripheral arterial disease and heart failure).
Non-LB Framingham (D'Agostino RB et al., 2008).	Same as above ("")	"	"	"	"	"	BMI (kg/m <sup>2</sup> ) <i>measured by anthropometry</i>	"	"
ARIC dataset (ARIC Investigators, 1989)	<u>Design:</u> Prospective study of the Atherosclerosis Risk in Communities cohort.	M or F	45-64	<ul style="list-style-type: none"> <li>• Yes, self-reported current smoker (<i>must be smoking at the time of assessment</i>)</li> <li>• No, none of the above criteria</li> </ul>	Systolic <i>average of 2<sup>nd</sup> &amp; 3<sup>rd</sup> measures obtained at baseline visit</i>	<ul style="list-style-type: none"> <li>• Yes, self-reported current treatment supplemented by thorough med review</li> <li>• No, none of the above criteria</li> </ul>	<ul style="list-style-type: none"> <li>• BMI (kg/m<sup>2</sup>) <i>measured by anthropometry</i></li> <li>• Total and HDL cholesterol, and other biomarkers <i>measured by standardized enzymatic methods</i></li> </ul>	<ul style="list-style-type: none"> <li>• Yes, fasting glucose <math>\geq 126</math> mg/dL (variable <i>D/ABTS03</i>) or 140 mg/dL (variable <i>D/ABTS02</i>) or use of insulin/oral hypoglycemic medications or history of DM.</li> <li>• No, none of the above criteria</li> </ul>	General CVD (coronary artery and cerebrovascular disease, and heart failure).
ARIC dataset (ARIC Investigators, 1989)	<u>Sample:</u> 15,792 (27% black, 55% women) aged 45-64 yrs., free of CVD at baseline exam (1987-1989) <u>follow-up:</u> Annual								

### The impact of CVD in resource constrained settings

A striking similarity in the epidemiology of CVD in both developed and developing countries pertains to its impact on underserved populations. Although developed countries have strong health care systems, advanced medical technologies and abundance of resources, they are still dominated by high CVD mortality and morbidity (D. Lloyd-Jones et al., 2010). In 2011 the US spent over three hundred billion dollars to manage CVD. However, despite the staggering healthcare expenditure, CVD continues to take a lopsided toll on underserved minority groups especially the black population (Mozaffarian et al., 2015).

In developing countries which are plagued by infectious diseases and underdeveloped fragile health care systems, CVD has significantly contributed to a protracted double burden of disease (S. Mendis et al., 2011). Despite lack of adequate resources to manage a full blown CVD epidemic, many developing countries are yet to implement feasible CVD prevention and surveillance initiatives. The inaction continues to expose masses of vulnerable populations to the dangers of cardiovascular events which are labor and resource intensive to manage.

#### The burden and impact of CVD in the black US population

According to the 2015 American Heart Association estimates, about half of all black adults in the US (48% women, 46% men) are affected by some form of CVD (Mozaffarian et al., 2015). In addition, US blacks bear a disproportionately high burden of CVD risk factors including obesity, diabetes, and hypertension. It is estimated that 46% of women and 45% of men in the adult black population have high blood pressure compared to 33% of the general US population (Mozaffarian et al., 2015). Despite this

high burden of CVD, blacks experience delays in CVD diagnosis and usually receive low quality of care leading to worse health outcomes (Bonow, Grant, & Jacobs, 2005).

Although health disparities are complex and multi-factorial, the high burden of CVD in US blacks has been related to suboptimal access to the healthcare system, primary care providers, and preventive health services (Institute of Medicine, 2003). Lack of insurance coverage and geographic location has been identified as major access barriers to quality CVD preventive and curative treatments (Escarce, 2007). Overall, blacks have the second lowest health insurance coverage in every state of the union, coming only second to the Hispanics (Wilson, 2013). With the escalating cost of health care, lack of insurance is a major deterrent to optimal health care access.

Geographic location also plays an important role in limiting access to CVD preventive and curative treatments. Due to low education and high poverty levels, many US blacks reside in rural areas or inner cities. As a result of high crime rates and/or other environmental factors, health workers avoid working in inner cities leaving these populations without adequate health care access. The same trend is observed in most rural areas which are characterized by resource constrained health care systems and physical barriers such as distance and unavailability of transportation (Kamble & Boyd, 2008). These geographic limitations arguably make US blacks more likely to experience difficulties accessing health care, leading to disparate cardiovascular health outcomes.

The US southern state of Mississippi has been cited as an example of how a geographical location can be a barrier to health care access. The state has highest proportion of rural-dwelling black women, and the highest heart disease death rate in the US (Kamble & Boyd, 2008). Notwithstanding these statistics, over 80% of the counties in

Mississippi have no physicians who specialize in heart disease (Kamble & Boyd, 2008). These geographic barriers and other challenges unique to resource constrained settings necessitates innovativeness to maximize each clinical encounter.

Current policy initiatives aimed at reducing health care access barriers include subsidies to improve insurance coverage, and incentives to encourage health workers to provide services in marginalized rural or inner city communities (Brennan, Baker, & Metzler, 2008; National Rural Health Association, 2013). To reduce the burden of CVD in the US black population, these policy initiatives need to be supplemented by pragmatic strategies that would help reduce health care costs without compromising quality.

Validating the non-LB Framingham algorithm in the black sample of ARIC cohort was an important step in availing a high quality CVD prevention tool that can be readily used in settings where laboratory measures are inaccessible due to location constraints or lack of insurance coverage. If validated, the algorithm will allow improved prediction of CVD events, enabling providers working in marginalized environments to better identify high risk individuals who require intensive preventive treatments. If used in combination with counseling, the validated algorithm may help to demonstrate a patient risk profile and indication for any proposed intervention (Shillinglaw, Viera, Edwards, Simpson, & Sheridan, 2012). Although absolute CVD risk profile is in the context of the “average person” calculating the absolute CVD risk score provides a relatively concrete basis for engaging the patient on the abstract concept of risk. The ensuing patient-provider risk reduction discussion is likely to foster improved health literacy and may increase adherence to the prescribed interventions.

### Preventing CVD in resource constrained settings

Integrating absolute CVD risk assessment into routine clinical assessment and population based surveys may foster a standardized opportunistic and proactive CVD risk surveillance and prevention in underserved populations. Availability of validated non-LB risk assessment algorithms will enable health care providers in resource constrained settings to initiate risk reduction discussion and interventions within one clinical visit. It has been estimated that an individual's absolute risk score can be calculated within five to ten minutes using these algorithms because the only data required to estimate absolute risk include: age, BMI, systolic blood pressure, antihypertensive medication use, current smoking, and diabetes status (D'Agostino RB et al., 2008; Gaziano et al., 2008). This point of care utility may add great value in underserved populations which are difficult to follow.

The proposed use of the non-LB Framingham algorithm, which was derived in a population that was 99.7% white, raises issues of applicability in black populations. Whereas poor performance of Framingham based algorithms has been reported in certain ethnicities (e.g. Hispanics), the models have performed reasonably well in predicting CVD in the US black population (Beswick et al., 2008).

The adoption of algorithms developed in significantly different settings and populations is traditionally done under the assumption that the major risk factors for CVD are fairly similar around the world (Yusuf et al., 2004). The INTERHEART investigators delineated 9 major risk factors (smoking, lipids, hypertension, diabetes, obesity, unhealthy diet, physical inactivity, harmful alcohol consumption, and



psychosocial factors) which account for over 90% of the population attributable risk of acute myocardial infarction worldwide (Yusuf et al., 2004).

### Summary

Validating and costing the non-LB Framingham algorithm could make an important contribution to the ongoing efforts to address the high burden of CVD in underserved communities especially the US black population. The challenges discussed in this review including the problem of limited access to healthcare due to location and lack of health insurance can be mitigated by availability of a valid and cost-effective risk assessment algorithm.

Such an algorithm can be deployed at the point of service in real time, without need for follow-up visits to draw laboratory specimens or to review results. The time utility would make every visit in the resource constrained settings an opportunity to initiate risk reduction discussion, and to motivate adherence and self-management since the data required is readily collected during the office visit. To improve the effectiveness of CVD prevention while using the non-LB Framingham algorithm, individuals with borderline or indeterminate risk may then be further screened using the more resource intensive laboratory measures whenever feasible.

Chapter 3 focuses on research design and methodology used in the external validation and cost-effectiveness analysis.

## CHAPTER 3

### RESEARCH DESIGN AND METHODS

#### Overview

To address the specific aims of the project, a secondary data analysis was conducted using the ARIC dataset. The dataset is organized in four cohorts, three of which are predominantly or completely white (Forsyth County, NC; Suburbs of Minneapolis, MN; and Washington County, MD) while one cohort (Jackson, MS) is composed of black participants (ARIC Investigators, 1989). This diversity facilitated an adequately powered sub-analysis of the performance of the tool in the black sample.

#### Data Source and Design

The ARIC study is a prospective epidemiologic study with an overarching goal to investigate the etiology and natural history of atherosclerosis and its clinical sequelae, and examine the distribution of cardiovascular risk factors, medical care, and disease by race, sex, living location, and time. The study includes cohort and community surveillance components conducted in four ARIC field centers that include Forsyth County, NC; Jackson, MI; Minneapolis, MN and Washington County, MD. The Cohort Component of the study commenced in 1987, with each ARIC field center using driver license lists to randomly select about 4,000 individuals aged between 45-64 years from a defined population in their community (ARIC Investigators, 1989).

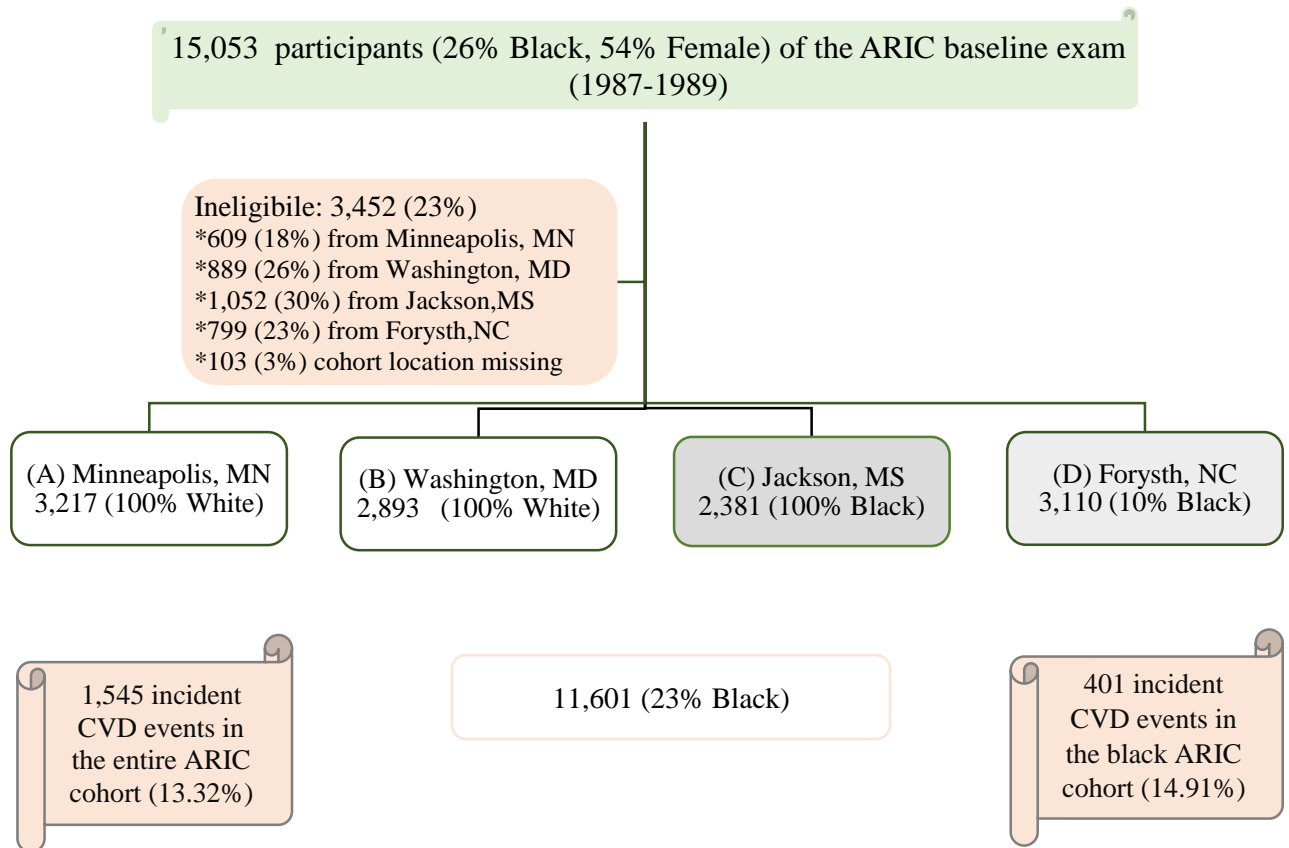
A total of 15,792 individuals were recruited, but the ARIC dataset provided by the NHLBI included 15,053 adults who had no missing variables on the identification variable. Before the eligibility criteria was employed, the dataset included a total of 8,163 women (54%) and 3,898 blacks (26%) aged 45-64 years. The sample was organized in four cohorts based on the ARIC field centers described earlier and all participants were examined at baseline between 1987 and 1989, followed by three more site-based examinations which ended in 1998. Yearly telephone follow-up interviews continue as a way to maintain contact with participants and to assess the health status of the cohort. Details of the examination procedures and criteria for the endpoints assessed have been reported elsewhere (ARIC Investigators, 1989). Table 1 summarizes sample characteristics, how essential risk factors relevant to this study were assessed, and CVD related endpoints monitored.

This secondary data analysis focused on the first 12 years of follow-up after baseline examination in ARIC. Therefore, this study's sample consists of study participants who attended baseline examination (1987-1989) and who at baseline were free of CVD, aged 45 to 64 years, and with no missing data on the variables of interest either at baseline or follow-up assessments in the next 12 years. The sample meeting these eligibility criteria includes 11,601 participants as described in Figure 2.

The 12 years follow-up employed in this study matches the follow-up time used by (D'Agostino RB et al., 2008) when generating the non-LB and LB Framingham algorithms. This congruence of follow-up time will increase comparability of the performance of the algorithms in ARIC and Framingham datasets. The covariates required to validate the non-LB Framingham algorithm include age, sex, and diabetes

status, smoking status, blood pressure, hypertension treatment and body mass index. The end points that are necessary to evaluate the calibration of the tool include the confirmed diagnosis of coronary artery disease, cerebrovascular disease, peripheral arterial disease and heart failure. Table 1 compares key aspects of the non-LB and LB Framingham algorithms and their derivation dataset, with the ARIC dataset in regard to design, sampling, assessments, data collected at baseline and the endpoints monitored. The cost-effectiveness analysis focused on the black subset of the ARIC dataset who met the eligibility criteria described above.

Figure 2: Flowchart of inclusion and exclusion criteria and monitored outcomes



### Baseline assessment

In the ARIC cohort, the baseline examination commenced with participants giving informed consent. The baseline examination assessed CVD conditions and measured key atherogenic risk factors. Key elements of the interview included the assessment of angina (Rose Questionnaire), history of diabetes, transient ischemic attack, and peripheral arterial disease, smoking status and medications use. Positive history was verified by laboratory test results and/or in the medical records abstracted by nurse researchers (ARIC Investigators, 1989).

Blood pressure was measured with the participant seated, with feet on the floor and arm at heart level, with three readings 5 minutes apart using random zero sphygmomanometer. The average of the second and third systolic blood pressure is entered into the model as a continuous variable. Anthropometric measurements were made with the participants wearing light-weight; non-constricting underwear, after emptying the bladder. Height and weight measurements were taken with the participant in light clothing and not wearing shoes. BMI was calculated as a function of height in meters and weight in kilograms, and was entered into the model as a continuous variable (ARIC Investigators, 1989).

Diabetes was operationalized by two variables in the ARIC study. The *DIABTS03* variable defined diabetes as fasting glucose greater or equal to 126 mg/dL, and use of insulin or oral hypoglycemic medications, while the *DIABTS02* variable applied a similar definition but used a fasting glucose level greater or equal to 140mg/dL. Fasting glucose in both variables was measured during the scheduled baseline assessment, after at least 8 hours of fasting. Sex and smoking status were ascertained based on self-report and

entered as dichotomous variables, while antihypertensive medication use was determined through self-report and medication review as described in Table 1 (ARIC Investigators, 1989).

Similar protocols were used to measure covariates in the Framingham cohort, with the exception of diabetes whereby fasting glucose greater or equal to 126mg/dL was used as the threshold for diabetes in the Framingham offspring cohort, and greater or equal to 140mg/dL as the threshold for the original cohort (D'Agostino RB et al., 2008). In this analysis, we used *DIABTS03* as the dichotomous variable indicating presence or absence of diabetes since it includes both thresholds used in the original and offspring cohorts of the Framingham study. Table 1 compares how selected sample characteristics, covariates and end points were assessed in the Framingham versus ARIC datasets.

#### Follow-up Assessments

All study participants were under continuous sentinel surveillance for the development of CVD events and death. The average follow-up response rate at year 12 was 95.68%, with all cohorts having response rates greater than 93% (University of North Carolina at Chapel Hill, 2013). After the baseline examination, a telephone questionnaire was administered annually, including the Rose angina questionnaire (screens for angina) and items on general health and hospitalization. Events of interest during follow-up included hospitalized and non-hospitalized myocardial infarction, coronary heart disease death, angina pectoris, stroke, and intermittent claudication (ARIC Investigators, 1989). However, the events included in the secondary dataset provided by National Heart Lung and Blood Institute (NHLBI) included; coronary artery and cerebrovascular disease, atrial fibrillation or flutter and heart failure.

In this analysis general CVD events include coronary heart and cerebrovascular disease, and heart failure. Atrial fibrillation and flutter were not included since they were not included in the Framingham cohort outcomes, while peripheral vascular disease outcomes were not available in the ARIC dataset provided by the NHLBI (see Table 1).

The cardiovascular events were ascertained through annual follow-up questionnaires, physical examinations at the study sites, communication with personal physicians and surveillance of medical records. Suspected new events were independently confirmed through review of medical charts by three experienced investigators, and neurological events were confirmed by a neurologist. Hospital records were abstracted twice independently by nurse abstractors to monitor CVD events; all substantive discrepancies were reconciled. Cardiac enzyme levels were recorded three times one day after the event, and two times for each of the next three days. The reviewed cardiac enzymes included: lactate dehydrogenase, lactate dehydrogenase subfractions, and creatinephosphokinase. Three serial electrocardiograms were reviewed, coded and interpreted at the University of Minnesota ECG Center using the full Minnesota code (ARIC Investigators, 1989).

Underlying and contributory causes for all deaths of cohort members were also investigated to determine whether the cause was CVD. Where death occurred in a hospital, the hospital record was used, but if the decedent had died outside the hospital, family interviews, physician questionnaires, and coroner records were used (ARIC Investigators, 1989).

The endpoints monitored in ARIC study follow-up closely correlate with the endpoints used in the internal validation of the non-LB and LB Framingham algorithm which

included coronary events (coronary death, myocardial infarction, coronary insufficiency, and angina), cerebrovascular events (ischemic stroke, hemorrhagic stroke, and transient ischemic attack), peripheral artery disease (intermittent claudication), and heart failure.(D'Agostino RB et al., 2008) This congruence of protocols for assessing risk factors and endpoints (with exception of failure to include peripheral vascular disease in ARIC) enables rigorous external validation of the algorithm (see Table 1). Cost-effectiveness measures are discussed separately in the analysis strategy under hypothesis 3 and 4.

### Statistical Analyses

The sex-specific non-LB Framingham algorithm was developed using sex-specific Cox proportional-hazards regression models (Cox regression). The covariates included in the model, which were also measured in the ARIC dataset, entail; age, systolic blood pressure, antihypertensive medication use, current smoking, BMI, and diagnosis of diabetes mellitus (D'Agostino RB et al., 2008). To ensure coherence and rigor of the external validity analysis, the same regression model and covariates were used in the secondary analysis of the ARIC dataset since the necessary inputs were available in the dataset.

Continuous variables (covariates) were transformed into natural logarithms to improve discrimination and calibration of the model and to minimize influence of extreme observations. Specific data on the exact days to incident CVD events since baseline examination, which are essential when using Cox regression, were included in the ARIC dataset. The incident CVD dates were ascertained through the rigorous follow-up methods detailed above, including systematic tracking of medical records and personal



communication with physicians since the participants had consented to these disclosures (ARIC Investigators, 1989).

Cox regression allowed evaluation of the effect of various independent variables on the time at which a specified event occurs without making assumptions on the baseline hazard. Cox regression was used in the survival analysis after all the covariates included in the models met the proportionality of hazards assumption (Cleves, 2008). The assumption required; a) all continuous predictors such as systolic blood pressure to have a constant effect on survival across all analysis groups during the entire period of follow-up; b) categorical predictors (e.g. smoking status) to have the same shape of hazard function within each analysis group during the entire period of follow-up.

#### Analysis plan by study aims

Specific aim 1: Assess the accuracy of the non-LB Framingham algorithm in stratifying risk and predicting CVD events in a racially diverse population.

Hypothesis 1: Non-LB Framingham algorithm will have adequate discrimination (Harrell's C statistic greater than 0.75) and calibration (Hosmer-Lemeshow goodness of fit ( $\chi^2$ ) below 20) in the multiracial ARIC dataset.

This hypothesis was tested by evaluating the frozen sex-specific non-LB Framingham algorithm's ability to: a) accurately stratify risk (discrimination) for persons who experienced a CVD event and those who did not, and b) predict CVD events (calibration), in the ARIC dataset. Discrimination (the ability of a risk prediction tool to assign a higher risk to those who will develop the end points of interest compared to those who will not) was measured using Harrell's C statistic and AUROC curve. Adequate discrimination was demonstrated by an overall C statistic of 0.75 or higher as

recommended in the literature (Cooney et al., 2009; May et al., 2006). In addition, sensitivity (proportion of individuals with CVD events who were predicted as high risk) and specificity (proportion of individuals without CVD events who are not predicted as high risk) of the tool was calculated using the *roctab* command in Stata®.

Calibration of the mathematical models was assessed by measuring the concurrence between the predicted outcomes and observed outcomes using Hosmer-Lemeshow goodness of fit statistic ( $\chi^2$ ). Kaplan-Meier survival analysis was used to obtain the observed incidence of CVD events, which was then be compared with the CVD events predicted by the non-LB Framingham algorithm. Kaplan-Meier method was preferred in this analysis because it allowed estimation of survival over time, even when some participants were censored or had varying lengths of follow-up. A Hosmer-Lemeshow goodness of fit ( $\chi^2$ ) statistic below 20 was considered a good fit as recommended in literature (Cooney et al., 2009).

The frozen Framingham model and algorithm was compared by one generated using ARIC data. In addition, re-calibration of the Framingham model using ARIC's baseline survival and risk factor means was done and new recalibrated model and algorithm developed for comparability in discrimination and calibration (D'Agostino RB, Grundy, Sullivan, Wilson, & CHD Risk Prediction Group, 2001a). In this study, the frozen Framingham algorithm refers to the unaltered Framingham algorithms with baseline survival, regression coefficients and mean of risk factors as published by D'Agostino and colleagues (D'Agostino RB et al., 2008).

Specific aim 2: Compare the performance of the non-LB Framingham algorithm in the black versus white participants of the ARIC study.

Hypothesis 2: There will be no significant difference in discrimination and calibration of the non-LB Framingham algorithm between the black and white participants of the ARIC study.

After validating the sex-specific non-LB Framingham algorithm in the entire ARIC cohort, a sub-analysis was conducted to assess the performance of the tool in the black and white cohorts. This analysis examined the ability of the tool to accurately predict CV risk in each racial subgroup stratified by sex. The Statistical analyses approach described under hypothesis 1 were used to conduct the sub-analysis by race. Re-calibration was done by substituting Framingham baseline survival and risk factor means with race specific baseline survival and risk factor means in the ARIC dataset. All the non-LB models were compared with their LB counterparts using AUROC analysis and agreement using kappa statistic. The kappa-statistic is measure of inter-rater agreement, which is 0 when agreement is random and 1 when agreement is perfect.

#### *Cost-effectiveness analysis*

Specific aim 3: Establish the feasibility of using the non-LB Framingham algorithm in guiding prevention of CVD among the black participants enlisted the ARIC study.

Hypothesis 3: A CVD prevention strategy guided by the non-LB Framingham algorithm will be more cost-effective compared with treating each elevated CVD risk factor (diabetes and/or hypertension) independently in the black subset of the ARIC cohort.

Hypothesis 4: A CVD prevention strategy guided by the non-LB Framingham algorithm will be more cost-effective compared with a strategy guided by the LB Framingham algorithm in the black subset of the ARIC cohort.

This analysis compared the cost-effectiveness of a CVD prevention strategy guided by absolute CVD risk estimates, calculated using the non-LB and LB Framingham algorithms, visa-vis an approach based on treating each elevated CVD risk factor (diabetes and/or hypertension) independently in the black subset of the ARIC cohort.

Although there are seven major risk factors for CVD that are modifiable, only two were included in costing the approach based on treating individual CVD risk factors. Hypertension and diabetes were selected due to their strong association with CVD and their significance in the global public health agenda. Hypertension is the leading cause of CVD worldwide, and diabetes is known to double the risk of CVD events (S. Mendis et al., 2011).

The thresholds for initiating therapy and treatment modalities in the individual risk factors approach e based on recommendations from the American Society of Hypertension and International Society of Hypertension (ASH-ISH) guidelines for management of hypertension, and the International Diabetes Federation (IDF) guidelines for management of diabetes (IDF Clinical Guidelines Task Force, 2006; Weber et al., 2014). The guidelines were selected for costing due to their primary focus on either diabetes or hypertension, and their international applicability. The essential components of these guidelines relevant to this analysis are discussed below and summarized in Figure 3.

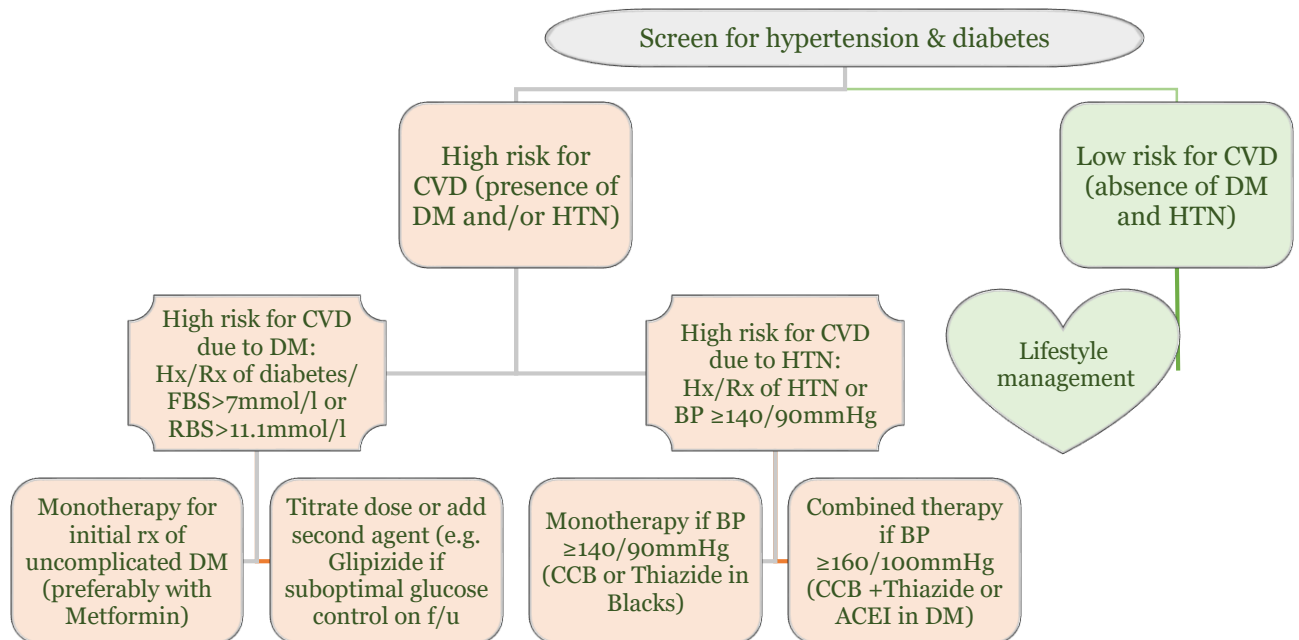
According to the ASH-ISH guidelines, hypertension is defined as systolic blood pressure greater or equal to 140 mmHg or diastolic blood pressure greater or equal to 90 mmHg or both. The guidelines recommend timing and tailoring the intensity of treatment based on the stage of hypertension. Stage 1 hypertension is defined as systolic blood pressure below 160 or diastolic blood pressure below 100, while stage 2 hypertension denotes blood pressures above these thresholds. Monotherapy with a calcium channel blocker (e.g. Amlodipine) or a thiazide diuretic (e.g. Hydrochlorothiazide) is recommended for blacks with stage 1 hypertension irrespective of their diabetes status, while addition of a second agent (angiotensin-converting enzyme (ACE) inhibitor in diabetes) is recommended in stage 2 hypertension as outlined in Figure 3.

In resource constrained settings, the guidelines recommend use of lifestyle modification for up to one year before starting drug therapy in stage 1 hypertension when no other CVD risk factors are present. All treatment modalities are focused on attaining targets below the diagnostic threshold but no explicit follow-up regimen is provided by the ASH-ISH guidelines (Weber et al., 2014).

The IDF guidelines defines diabetes as fasting blood sugar greater than 7 mmol/l (126mg/dL) or random blood sugar above 11.1mmol/l (200mg/dL). Treatment options are graded from first-line to fourth-line therapy depending on attainment of glucose control targets. First-line therapy includes monotherapy with a biguanide (e.g. Metformin) or an equivalent agent, while second-line therapy adds a sulfonylurea (e.g. Glipizide) or an equivalent agent as summarized in Figure 3. Addition of a third agent, such as a-glucosidase inhibitor, or starting insulin treatment constitutes third-line therapy. If glucose control targets are not achieved with third-line therapy, insulin treatment is

initiated alongside oral hypoglycemic agents as part of fourth-line therapy. The guidelines also recommend use of statins (based on lipid levels) to reduce risk of CVD events and periodic monitoring of glycated hemoglobin (HbA1C) as part of a comprehensive management program (IDF Clinical Guidelines Task Force, 2006). A follow-up regimen is not explicitly stipulated. Figure 3 outlines risk stratification and the basic preventive interventions prescribed by the approach guided by treating individual CVD risk factors.

Figure 3: CVD prevention strategy based on treating diabetes/hypertension to target goals



CVD prevention strategy based on treating diabetes and hypertension to target goals adapted from ASH-ISH and IDF guidelines (IDF Clinical Guidelines Task Force, 2006; Weber et al., 2014)

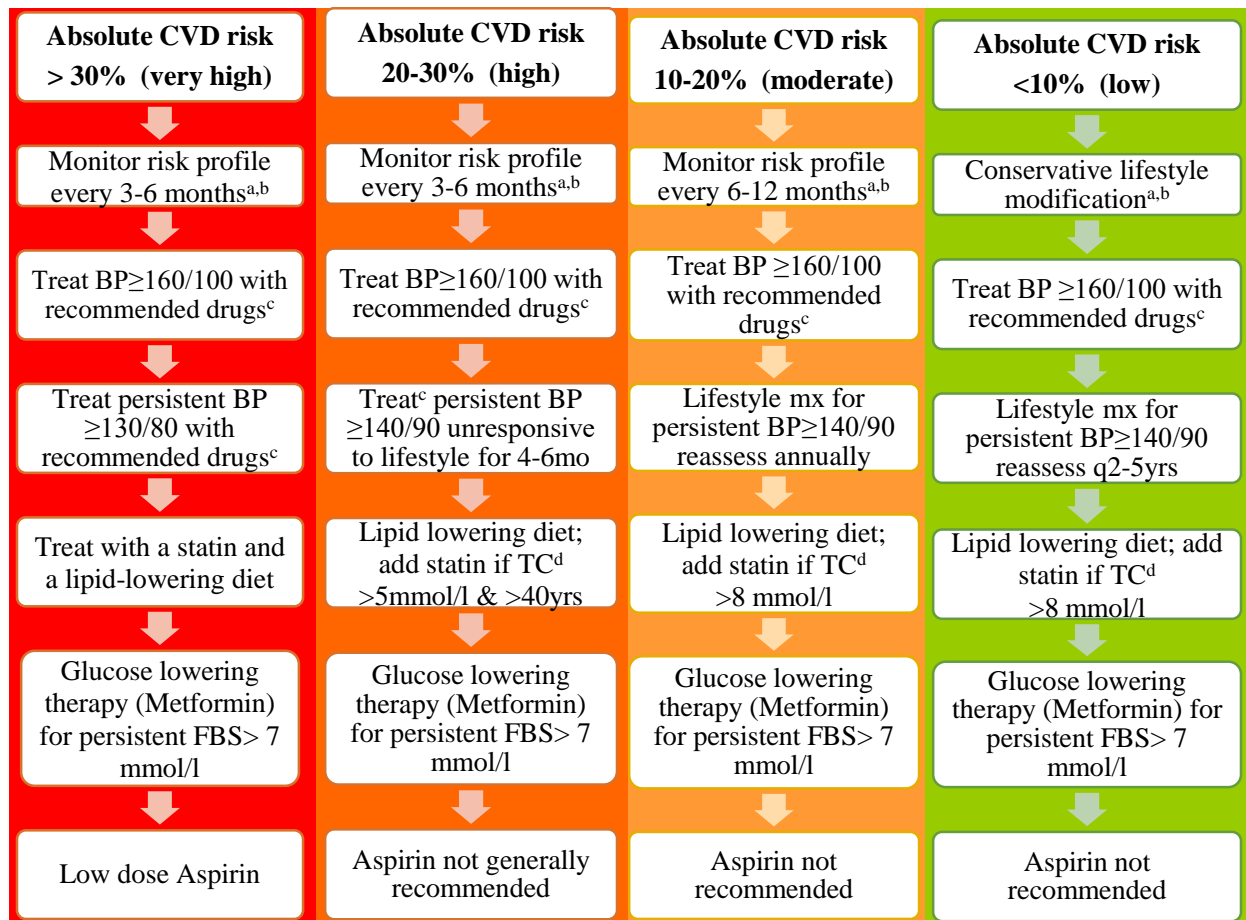
The absolute risk approach to CVD prevention is recommended by major CVD management guidelines based on the premise that whereas individual risk factors independently increase the likelihood of CVD events, clustering of multiple risk factors is known to compound the CVD risk (Beswick et al., 2008). Guidelines adopting the absolute risk approach to CVD prevention tailor the choice and intensity of recommended

treatments based on absolute CVD risk scores calculated using CVD risk assessment algorithms, such as the non-LB Framingham algorithm externally validated in this study.

The 2007 CVD prevention guidelines by the WHO were selected as the basis for costing interventions associated with the absolute risk approach due to their congruence with Framingham algorithms and relevance to primary prevention of CVD. Interventions recommended by the WHO guidelines (see Table 2) are based on absolute risk scores for general CVD events including coronary heart disease, peripheral vascular disease and cerebral vascular disease (World Health Organization, 2007). Both the non-LB and LB Framingham algorithms validated in this study were developed to predict these general CVD events, these are broader outcomes than the hard coronary events predicted by earlier versions of Framingham (D'Agostino RB et al., 2008). Currently there are no feasible alternatives to the WHO guidelines since the American Heart Association CVD prevention guidelines are dated (published in 2002) and limited by their narrow focus on hard coronary events (Pearson et al., 2002).

The WHO CVD prevention guidelines organize their recommended preventive interventions in four categories based on absolute CVD risk scores as follows; low risk (>10%), moderate risk (10% to 20%), high risk (20% to 30%) and very high risk >30%. Table 2 outlines the four risk categories and the treatment options recommended for each (World Health Organization, 2007). According to these guidelines, an individual with an absolute risk score >30% is: a) scheduled for follow-up visits at least every 6 months, b) started on antihypertensive therapy (if blood pressure is greater or equal to 130/80mmHg), c) started on a statin, d) put on glucose lowering therapy (if fasting blood sugar is equal or greater than 7mmol/l), and e) started on low dose aspirin therapy.

Table 2: WHO choice & intensity of CVD prevention strategy guided by absolute CVD risk



CVD prevention strategy based absolute risk score, adapted from the WHO CVD prevention guidelines (World Health Organization, 2007).

Key: TC (total cholesterol); FBS (fasting blood sugars)

<sup>a</sup> Smoking cessation and lifestyle management recommended across risk profiles

<sup>b</sup> Abstaining or reducing alcohol intake to <3units per day recommended across risk profiles

<sup>c</sup> First line antihypertensive therapy includes: thiazide-like diuretic, ACE inhibitor, calcium channel blocker

<sup>d</sup> Measures of TC may not be accessible in resource constrained regions.



### Case example: individual vs absolute CVD risk approach

The differences between the individual risk factors approach and the absolute risk approach to CVD prevention are evident in the treatment modalities recommended in Figure 3 and Table 2 for each strategy respectively. A case example is Mr. Q, a 57 years old male, who is a smoker with no history of diabetes or hypertension. He presents with a blood pressure of 138/88 mmHg, fasting blood sugar of 6 mmol/l (108mg/dL), HDL of 1.16 mmol/l (45mg/dL), total cholesterol of 5.84 mmol/l (226 mg/dL) and a BMI of 26.

If the individual risk factors approach outlined in Figure 3 was used to manage Mr. Q, only lifestyle modification, with emphasis on smoking cessation, would be recommended since he does not meet the hypertension threshold specified by the ASH-ISH guidelines or the threshold for diabetes recommended by the IDF guidelines.

If the absolute risk approach was used to manage Mr. Q, he would have an absolute CVD risk score of 30.4% and 30.1% according to the online interactive non-LB and LB Framingham absolute CVD risk calculators respectively (D'Agostino & Pencina, 2016). These absolute CVD risk calculators use the general formulae for predicting CVD events outlined in the literature review section. The general formulae combine sundry CVD risk factors to calculate the probability of a CVD event occurring within a maximum time frame of 12 years.

Mr. Q absolute CVD risk score is very high despite apparently normal or near normal individual risk factors because the score appreciates the additive nature of subtle elevations in CVD risk factors (e.g. blood pressure, BMI and total cholesterol). Consequently, in addition to lifestyle modification, the absolute risk approach based on either the non-LB or LB Framingham algorithm absolute risk score will prescribe Mr. Q

the relevant interventions under the “very high risk category” (>30%) in Table 2. The relevant interventions for Mr. Q would include: follow-up visits at least every 6 months, antihypertensive therapy (since BP>130/80mmHg), statin therapy, and low dose aspirin.

#### Absolute versus individual risk decision model

In order to fully appreciate the differences between the individual risk factors and absolute risk approaches to CVD prevention, the black cohort in ARIC was stratified by the type of screening algorithm used in the preventive approach. The different categories for each CVD prevention approach are summarized in Figure 4.1 and detailed below.

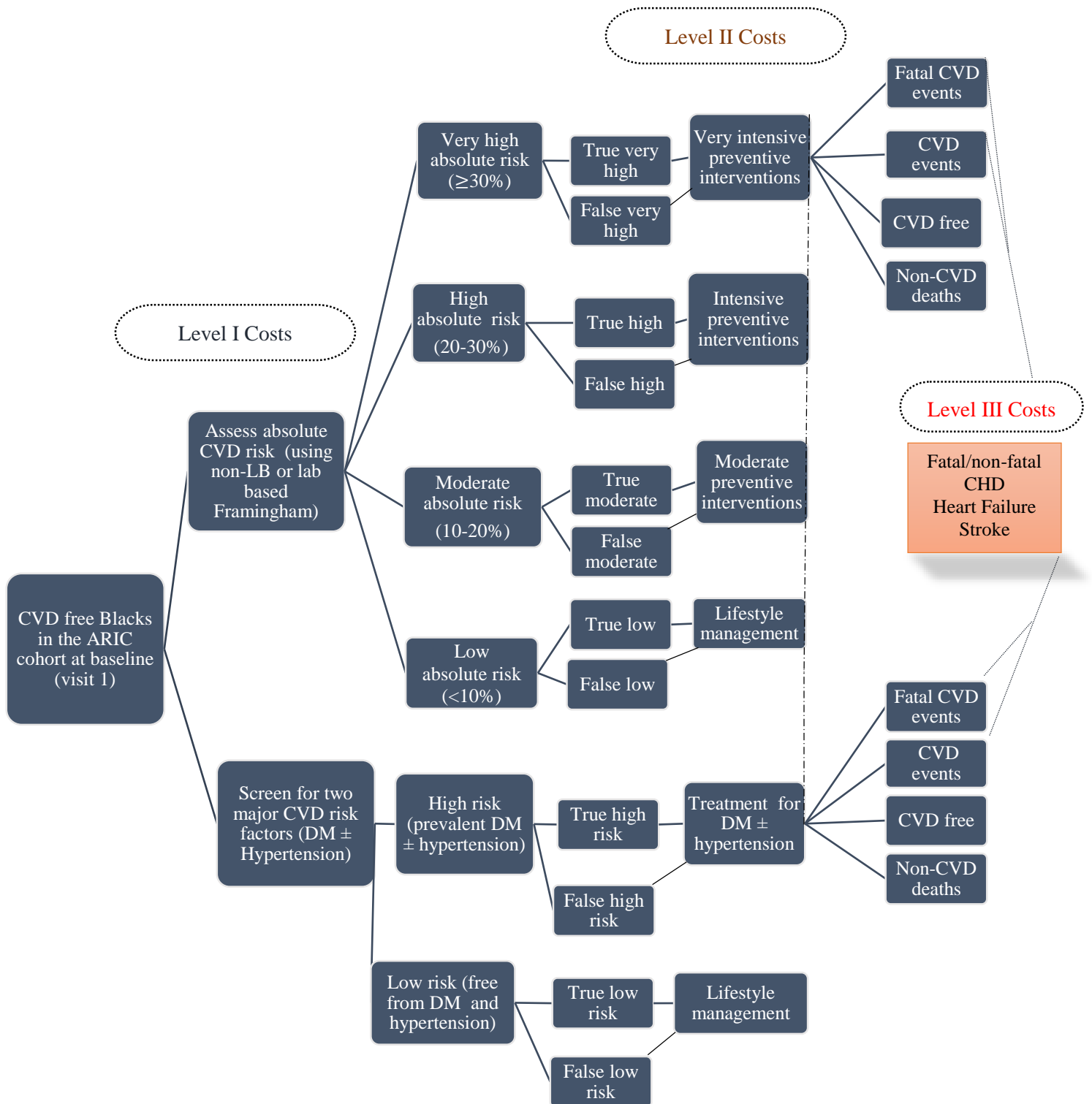
When the individual CVD risk factor approach was employed (lower arm in Figure 4.1), the black cohort free of CVD at baseline was stratified into high and low CVD risk categories depending on presence or absence of diabetes and/or hypertension. The true and false high risk categories were both prescribed the intensive preventive interventions outlined in Figure 3, while the true and false negatives were prescribed lifestyle management. The false positives ended up receiving unnecessary intensive treatment, while the false negatives missed essential treatment culminating in CVD events. In the simulated analysis, a high number of false positives were expected to increase level II expenses, while a high number of false negatives would increase level III expenses due to treatment and rehabilitative costs associated with CVD events that occur as a result of missing preventive interventions.

When the absolute CVD risk approach was employed guided by either the non-LB or LB Framingham algorithm, the black cohort free of CVD at baseline was stratified into four CVD risk categories based on their absolute risk score. When a specific risk category was selected as the threshold for initiating treatment based on

sensitivity/specificity analysis, individuals below the threshold were assumed to be low risk, while those above the threshold were considered high risk and put under the treatments prescribed for their respective category.

For instance, if the moderate risk category (10-20%) in Figure 4.1 was set as the treatment threshold, the sample with absolute risk score below 10% would be exempted from treatment, while individuals in other risk categories would receive the appropriate therapy based on their risk score as outlined in Table 2. As a result, the false positives receive unnecessary treatments and increase level II expenses, while the false negatives miss essential preventive treatments culminating in CVD events which incur level III treatments outlined in Figure 4.1.

Figure 4.1: Event trajectories associated with three CVD prevention strategies



### *Cost-effectiveness framework*

The incremental cost effectiveness tested under hypothesis 3 and 4 was done under the framework of the cost-effectiveness model below adapted from the methods described by Drummond for evaluating incremental costs and effects of a program (Drummond & Drummond, 2005). To adapt the equation for hypothesis 4, the individual risk factors (r) approach was substituted with the LB Framingham (l) approach).

The incremental cost effectiveness ratio (ICER) model:

$$ICER = \frac{\Delta TC_{a-r}}{\Delta E_{a-r}} = \frac{\sum_{t=1}^{12} \sum_{j=1}^2 [CRx_{ajt} + CUSE_{ajt} - CRx_{rjt} - CUSE_{rjt}] / (1+d)^t}{\sum_{t=1}^{12} \sum_{j=1}^2 [CVD_{ajt} - CVD_{rjt}] / (1+d)^t}$$

*ICER model Key:*

- ICER = incremental cost-effectiveness ratio
- $\Delta TC_{a-r}$  = the discounted difference between the total costs incurred in 12 years to manage CVD in the absolute (a) versus the individual risk factors (r) approach.
- $\Delta E_{a-r}$  = the discounted difference between true positives predicted in the absolute (a) vs. the individual risk factors (r) approach.
- $CRx_{ajt}$ ,  $CRx_{rjt}$  = the discounted cost of preventive interventions (see Table 2) prescribed to the j-th risk group predicted by absolute (a) versus by individual risk factors (r) approach in year t (same for  $CRx_{rjt}$ , use Figure 3).
- $CUSE_{ajt}$ ,  $CUSE_{rjt}$  = the discounted cost of treating 3 major CVD events occurring (false negatives in Figure 4.1) in the j-th risk group associated with absolute (a) vs individual risk factors (r) approach in year t (same for  $CUSE_{rjt}$ )
- $CVD_{ajt}$ - $CVD_{rjt}$  = the discounted difference between true positives predicted in the j-th absolute or individual risk factors group in year t.
- $d = 3\%$  discount rate as recommended by the US Panel on Cost-Effectiveness in Health and Medicine (Weinstein, Siegel, Gold, Kamlet, & Russell, 1996).

Each preventive approach was costed on three levels as outlined in Figure 4.1. Level I expenses includes screening costs, level II expenses included the cost of preventive interventions prescribed in Table 2 for the absolute risk approach, and Figure 3 for the individual risk factors approach, while level III expenses included the cost of treating CVD events that occurred in the false negative group. Outcomes in this analysis included three CVD events (fatal and non-fatal CHD, heart failure, and stroke) expected to be prevented over 12 years after implementation of each preventive strategy. The sum of

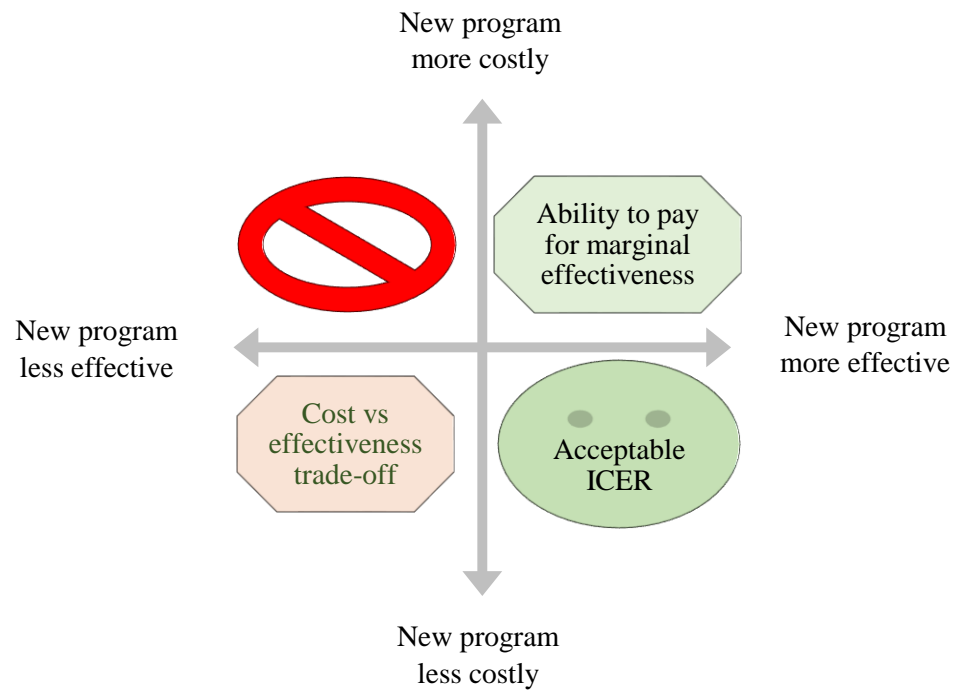
level I and II expenses constituted CRx, while level III expenses were represented by CUSE in the ICER model. All costs and outcomes were discounted at the rate of 3% as recommended by the US Panel on Cost-effectiveness in health and medicine (Weinstein et al., 1996).

Discounting refers to the practice of weighting future gains and losses less heavily than those that occur in the present. The concept is based on the premise that a dollar or a life is worth more today than it would be in the future (Smith & Gravelle, 2001). In this analysis, discounting is used to estimate the present value of future costs and CVD events associated with each of the three CVD prevention programs.

#### *Interpreting the ICER*

The incremental cost-effectiveness ratio was interpreted using a cost-effectiveness plane adapted from Drummond (Drummond & Drummond, 2005). In this analysis, the existing programs were considered to be the individual risk factors and LB absolute risk approaches to CVD prevention, while the new program was considered to be the non-LB absolute risk approach. In the plane outlined in Figure 4.2, an ideal ICER ratio would be a more effective and less costly new program as depicted in the lower right quadrant. The ratio could also be acceptable if the new CVD prevention program is slightly less effective but way less costly (see left lower quadrant), or more expensive but highly effective compared to the existing programs (see right upper quadrant). An expensive but less effective program is unacceptable under all circumstances. At origin, the new program is similar in costs and effects with the existing program.

Figure 4.2: Cost-effectiveness plane depicting acceptability of ICER ratios



## CHAPTER 4

### RESULTS

The external validation described in this chapter entails evaluation of mathematical performance and applicability of the Framingham non-LB algorithms in the ARIC dataset. Mathematical performance is evaluated by reproducing the underlying Framingham mathematical models in ARIC and comparing their regression coefficients, discrimination and calibration with those derived in the Framingham dataset. Applicability of the Framingham algorithms in ARIC is evaluated by comparing the AUROC's and sensitivity/specificity ratios of CVD risk stratification based on the published Framingham algorithms versus other models generated in the ARIC dataset.

This chapter is organized in five parts to address; a) description of the ARIC sample baseline characteristics and incident CVD events, b) mathematical performance of the Framingham models in ARIC dataset, c) performance of the Framingham algorithms in the ARIC sample, d) sensitivity/specificity analysis of the algorithms applied to predict risk in the ARIC sample, and e) simulated cost-effectiveness analysis of the non-LB Framingham algorithm in the black cohort of the ARIC sample.



a) Sample description

This section describes the baseline characteristics of the eligible and ineligible sample based on the criteria discussed in the methods section, and the observed incidence of CVD events within the 12 years of follow stratified by sex and race.

Sample characteristics

The ARIC dataset provided by the NHLBI consisted 15,053 participants who completed the ARIC baseline exam between 1987 and 1989. Twenty-six percent of these participants were black and 55% were female. When the eligibility criteria described in the methods section were applied, 23% (3,452) of the original ARIC cohort was excluded. Therefore, the sample in this study includes 11,601 participants (23% black, 55% female).

The baseline characteristics of the eligible and ineligible sample are described in Table 3.1. Overall, 82% of those excluded had CVD at baseline examination. When the eligible sample was stratified by sex, men had a higher profile of CVD risk factors (diabetes, hypertension, smoking, HDL and total cholesterol) compared to women except for BMI and diabetes. Table 3.2 outlines the variables of interest stratified by gender.

Table 3.1: Baseline characteristics stratified by eligibility criteria

	Eligible	Ineligible	P (95% CI)
N	11601	3452	
Male (n, %)	5246 (45.22)	1541 (46.01)	0.417
white (%)	8911 (76.81)	2141 (63.93)	0.000
Age (mean, sd)	53.91 (5.74)	55.43 (5.69)	0.000
Diabetes (n, %)	1152 (9.93)	658 (20.48)	0.000
Current smoker (n, %)	2970 (25.60)	990 (29.69)	0.000
BMI (mean, sd)	27.39 (5.14)	28.84 (5.96)	0.000
SBP (mean, sd)	120.27 (17.99)	125.35 (21.90)	0.000
Hypertension (n, %)	3,866 (33.32)	2116 (64.28)	0.000
BP treatment (n, %)	2775 (23.92)	1858 (55.61)	0.000
HDL (mean, sd)	52.15 (17.08)	48.92 (16.93)	0.000
Total chol (mean, sd)	214.21 (41.05)	218.19 (45.36)	0.000
Baseline CVD (n, %)	0 (0)	2801 (81.52)	0.000

Baseline characteristics recorded at the baseline exam for entire ARIC cohort between 1987 and 1989

Data presented as mean/SD for continuous variables and as frequencies (%) for categorical variables

Chi<sup>2</sup> and t test were used as the tests of homogeneity for categorical and continuous variables respectively

Table 3.2: Baseline characteristics of the eligible sample stratified by sex and race

	White female	Black female	p (95% CI)	White male	Black male	p (95% CI)
N	4758	1597		4153	1093	
Diabetes ( %)	6.85	18.79	0.000	8.72	15	0.000
Current smoker (%)	24.88	23.67	0.329	24.01	37.60	0.000
BMI (mean, sd)	26.3 (5.2)	30.6 (6.3)	0.000	27.3 (3.9)	27.5 (4.9)	0.213
SBP (mean, sd)	116.4 (17.4)	126.5 (19.5)	0.000	120 (15.8)	129 (20.1)	0.000
Hypertension (%)	27.87	53.73	0.000	26.99	51.33	0.000
BP treatment (%)	21.21	42.08	0.000	18.28	30.65	0.000
HDL (mean, sd)	58 (17)	58.3 (17.3)	0.525	43.2 (12.4)	51.3 (17.2)	0.000
Total chol (mean, sd)	217.4 (41.4)	216.5 (44.4)	0.464	210.3 (38)	211.6 (43.1)	0.319
Age (mean, sd)	53.8 (5.7)	53.1 (5.8)	0.000	54.4 (5.7)	53.5 (6)	0.000
Intermediate educ. (%)	51.16	29.59	0.000	39.43	26.72	0.000
Advanced educ. (%)	33.85	32.79	0.000	44.61	31.22	0.000

Baseline characteristics recorded at the baseline exam for entire ARIC cohort between 1987 and 1989

Data presented as mean/SD for continuous variables and as frequencies (%) for categorical variables

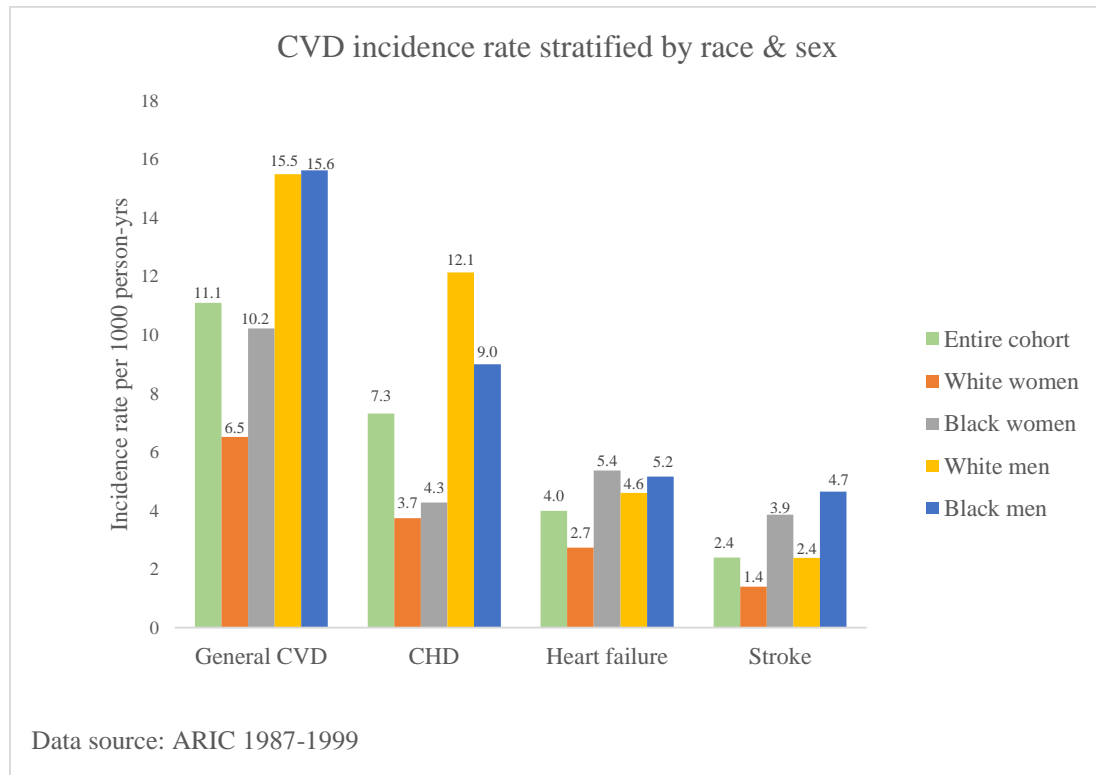
Chi<sup>2</sup> and t test were used as the tests of homogeneity for categorical and continuous variables respectively

### *Incident CVD during 12 years of follow-up*

During the 12 years of follow-up included in this analysis 1,545 new cases of CVD occurred in the eligible ARIC cohort comprised of 11,601 individuals. This translated into 11.1 incident CVD cases per 1000 persons-years. Stratified by sex and race, the incidence rate among white versus black women was 6.5 versus 10.2 cases per 1000 person-years respectively, and 15.5 versus 15.6 cases per 1000 person-years among white versus black men respectively.

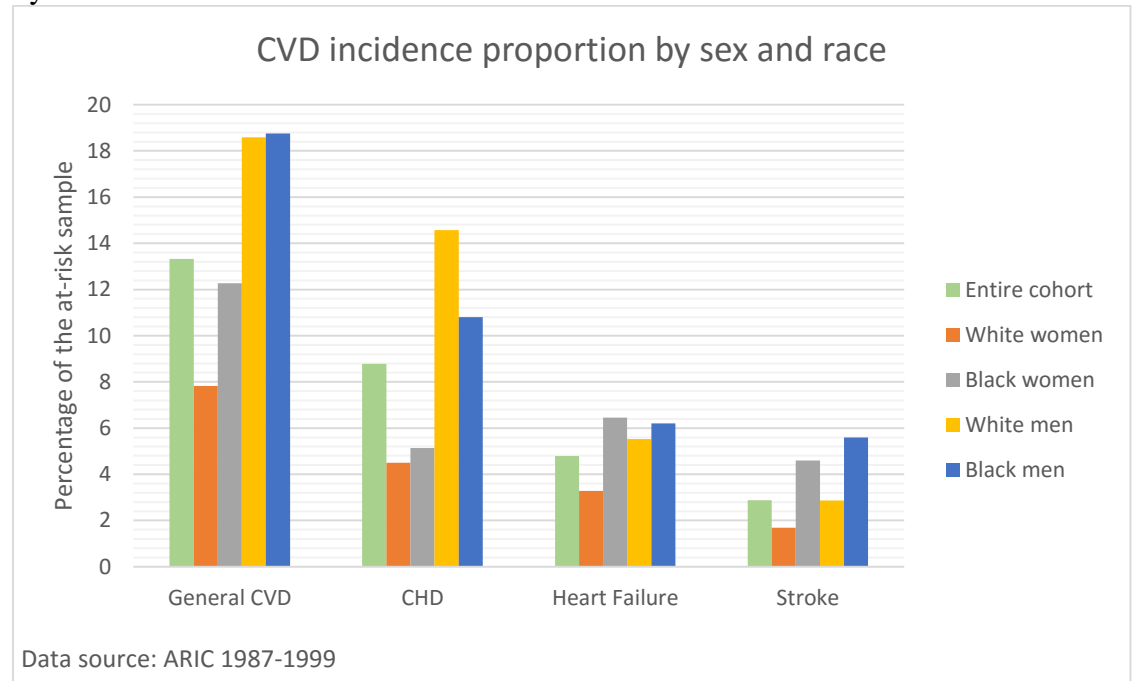
white and black men had the highest incidence rate of CHD (12.1 and 9.0 cases per 1000 person-years respectively), while black and white women had the lowest incidence (4.3 and 3.7 per 1000 person-years respectively). The incidence of heart failure was highest among black women and men (5.4 and 5.2 cases per 1000 person-years respectively), while white men and women had the lowest incidence (4.6 and 2.7 cases per 1000 person-years respectively). The incidence of stroke was highest among black men and women (4.7 and 3.9 case per 1000 person-years respectively). White men and women incidence of Stroke was low at 2.4 and 1.4 cases per 1000 person-years respectively. Figure 5 shows the incidence rate of general and cause specific CVD events stratified by sex and race.

Figure 5: General & cause specific CVD incidence rate in ARIC stratified by sex & race



The incidence proportion of CVD during the 12 years of follow-up was 13.3% in the entire ARIC cohort. Stratified by sex and race, the incidence proportion was 7.8% versus 12.3% among white versus black women respectively ( $p=0.000$ ), and 18.6% versus 18.8% among white and black men respectively ( $p=0.900$ ). Figures 6 depict the incidence proportion of general and cause specific CVD during the 12 years of follow-up stratified by sex and race.

Figure 6: General & case specific CVD incidence proportion in ARIC stratified by sex & race



#### b) Mathematical models' performance

In this section, two mathematical models, namely unadjusted and adjusted Framingham models, were developed and their mathematical performance compared to the published (frozen) Framingham model. The unadjusted model was generated by running a Cox regression model that included only the covariates used to derive the published Framingham algorithms, while the adjusted model included additional covariates postulated to improve the discrimination and calibration of the published Framingham models. The mathematical model performance was evaluated using the metrics of discrimination and calibration as discussed in the methods section. In addition,

the regression coefficients generated in the unadjusted model were compared to the published (frozen) Framingham coefficients to compare the effect size of each covariate in the Framingham versus ARIC cohorts.

Both the unadjusted and adjusted models met the proportional hazards assumption (global test  $>0.05$ ) and did not indicate presence of multicollinearity among the covariates. To compare the effect of each covariate in the Framingham versus ARIC cohorts, a statistical test of the difference in their respective regression coefficients ( $\beta$ ) was done using the z-score formula:  $z = (b_1 - b_2) / \sqrt{se_{b_1}^2 + se_{b_2}^2}$ , where  $b_1$  and  $b_2$  are the unstandardized regression coefficients ( $\beta$ ), while  $se_{b_1}$  and  $se_{b_2}$  are the standard errors of the regression coefficients (Paternoster, Brame, Mazerolle, & Piquero, 1998). Regression coefficients with similar effects across cohorts are expected to have a z score in the interval -1.96 and +1.96. A positive z means that the regression coefficient in ARIC was greater than Framingham and vice versa.

#### *The unadjusted non-LB Framingham models*

The Cox regression analysis including the non-LB covariates (sex, age, smoking status, diabetes status, antihypertensive medication use, systolic blood pressure and BMI) satisfied the proportional hazards assumption with a global test of  $p=0.2384$  among women and  $p=0.2873$  among men. All the covariates included in the unadjusted non-LB Framingham model were statistically significant ( $p<0.05$ ) in both sexes and in the white cohort, but BMI was not significant in the black cohort ( $p=0.071$  for women and  $0.128$  for men). In addition, untreated systolic blood pressure was not statistically significant among black men ( $p=0.065$ ). Table 4 describes the regression coefficients for the unadjusted non-LB Framingham model, stratified by sex and race.

Stratified by sex only, the regression coefficients generated in the unadjusted non-LB Framingham model were similar to the frozen non-LB Framingham regression coefficients, except for the smoking covariate among men and women, and the diabetes covariate among men. The effect of smoking was higher among women (z score of 3.28) but lower among men (z score of -2.842) in ARIC compared to the Framingham cohort. The effect of diabetes was higher among men (z score of 2.188) but similar among women in ARIC compared to the Framingham cohort.

Stratified by sex and race, the regression coefficients for all strata were similar to the frozen non-LB Framingham regression coefficients, except for the smoking and diabetes coefficients which were different in the white cohort. Smoking had a higher effect among white women (z score of 3.327) and a lower effect among white men (z score of -2.637) in ARIC compared to the Framingham cohort. Diabetes had a higher effect among white men (z score of 2.174) but similar effect among white women in ARIC compared to the Framingham cohort. Table 4 shows the regression coefficients for the unadjusted non-LB Framingham model and the associated z scores, stratified by sex and race.

Table 4: Regression beta coefficients of the unadjusted non-LB Framingham in ARIC and z test score of equality

Non-LB based covariates	Female coefficients stratified by race			Male coefficients stratified by race		
	Women	white women	black women	Men	white men	black men
N	6355	4758	1597	5246	4153	1093
Current smoker	1.053*** [0.0885] z = 3.28†	1.106*** [0.108] z = 3.327†	0.914*** [0.157] z = 1.592	0.412*** [0.0707] z = -2.842†	0.411*** [0.0826] z = -2.637†	0.501*** [0.145] z = -1.269
Diabetes	1.073*** [0.101] z = 1.65	1.073*** [0.136] z = 1.467	1.044*** [0.154] z = 1.25	0.831*** [0.0814] z = 2.188†	0.848*** [0.0948] z = 2.174†	0.831*** [0.161] z = 1.537
Log of Age	3.129*** [0.428] z = 0.788	3.333*** [0.545] z = 0.991	2.841*** [0.723] z = 0.154	3.050*** [0.321] z = -0.164	2.881*** [0.372] z = -0.540	3.330*** [0.654] z = 0.316
Log of BMI	0.801*** [0.219] z = 0.828	0.797** [0.284] z = 0.726	0.668 <sup>γ</sup> [0.370] z = 0.341	0.793*** [0.222] z = 0.000	0.864** [0.265] z = 0.180	0.631 [0.415] z = -0.320
Log of non-treated SBP	1.958*** [0.353] z = -1.701	1.808*** [0.435] z = -1.786	2.158*** [0.633] z = -0.902	1.499*** [0.270] z = -0.897	1.790*** [0.323] z = -0.151	1.018 <sup>γ</sup> [0.551] z = -1.341
Log of treated SBP	3.106*** [0.389] z = 0.425	3.054*** [0.600] z = 0.247	3.033*** [0.531] z = 0.235	1.834*** [0.399] z = -0.188	1.719*** [0.497] z = -0.361	2.246** [0.693] z = 0.425

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001 <sup>γ</sup> p<0.1

Standard errors in brackets

z = score comparing published non-LB Framingham coefficients with those generated in ARIC using the formulae  $f(z) = (b_1 - b_2) / \sqrt{(se_{b1}^2 + se_{b2}^2)}$ ; † denotes z score >1.96 or < -1.96. Positive z means the regression coefficient in ARIC was greater than Framingham & vice versa

In regards to risk stratification and congruence between predicted and observed CVD events, the unadjusted non-LB Framingham model had a higher discrimination statistic and better calibration among women compared to men. However, no significant differences in discrimination were observed between blacks and whites within their respective sexes. Women had a C statistic of 0.75 (95% confidence interval [CI], 0.73-0.77) compared to men's 0.67 (95% CI, 0.65-0.68). When stratified by sex and race, white women had a C statistic of 0.746 (95% CI, 0.72-0.77) compared to black women's 0.745 (95% CI, 0.71-0.78), while white men had a C statistic of 0.664 (95% CI, 0.65-



0.68) compared to black men's 0.683 (95% CI, 0.65-0.72). The published (frozen) non LB Framingham model had a C statistic of 0.785 (95% CI, 0.764-0.806) for women and 0.749 (95% CI, 0.731-0.767) for men.

Calibration was good ( $p > 0.05$ ) among women but poor among men with Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 14.2$  ( $p = 0.1154$ ) versus 25.8 ( $p = 0.0022$ ) respectively as depicted in Figure 7 and 8 below. When stratified by sex and race, calibration was good among all ARIC cohorts except white men with Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 19.6$  ( $p = 0.0208$ ). The published Framingham model had a Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 10.24$  for women ( $p = 0.33$ ) and 13.61 ( $p = 0.14$ ) for men.

Figure 7: Calibration- unadjusted non-LB based Framingham women specific model

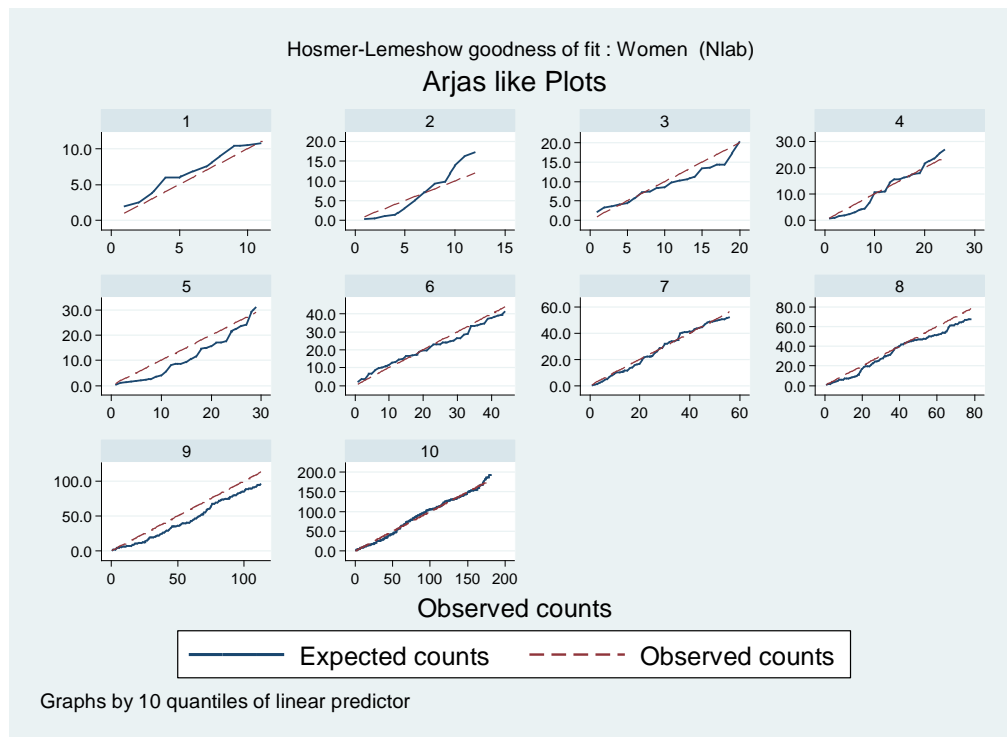
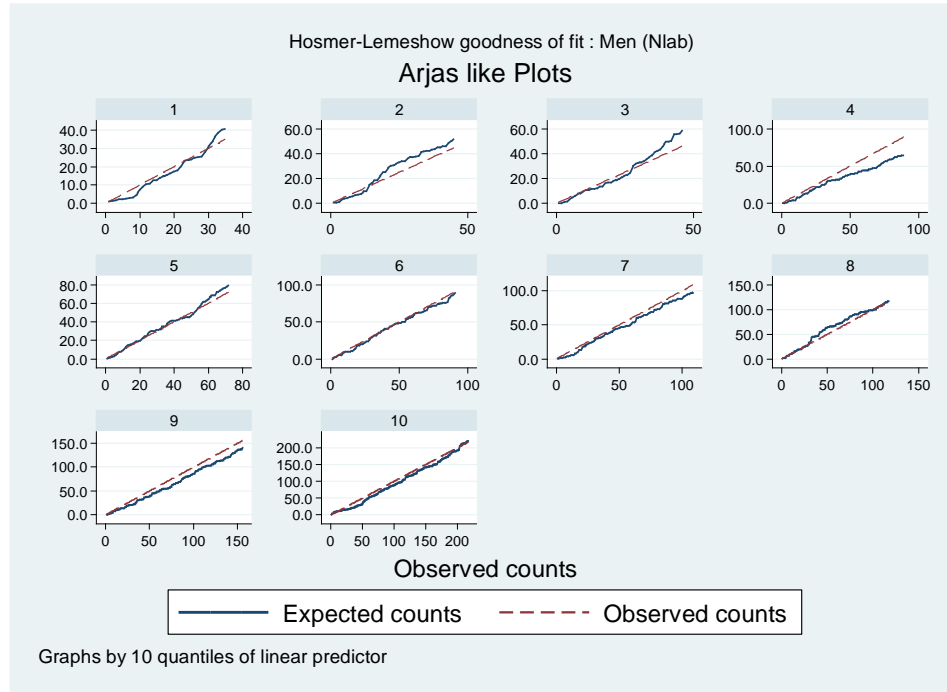


Figure 8: Calibration- unadjusted non-LB based Framingham men specific model



### *The unadjusted LB Framingham models*

The Cox regression analysis including the LB covariates (sex, age, smoking status, diabetes status, antihypertensive medication use, systolic blood pressure total cholesterol and HDL) also satisfied the proportional hazards assumption over the duration of follow-up with a global test of  $p=0.2999$  among women and  $p=0.4111$  among men. All the covariates included in the unadjusted LB Framingham model were statistically significant ( $p<0.05$ ) in both sexes and in the white cohort, but total cholesterol was not statistically significant in the black cohort ( $p=0.356$  for women and  $0.867$  for men). Table 5 presents the regression coefficients for the unadjusted LB Framingham model, stratified by sex and race.

When stratified by sex, all the regression coefficients generated in the unadjusted LB Framingham model were similar to the frozen LB Framingham regression

coefficients, except for the smoking status among men and women. Smoking had a higher effect among women (z score of 2.946) but a lower effect among men (z score of -2.720) in ARIC compared to the Framingham cohort. Table 5 below describes the regression coefficients for the unadjusted LB Framingham model and the associated z scores stratified by sex.

Table 5: Regression beta coefficients of the unadjusted LB Framingham in ARIC and z test score of equality

LB covariates	Female coefficients stratified by race			Male coefficients stratified by race		
	Women	white women	black women	Men	white men	black men
N	6355	4758	1597	5246	4153	1093
Current smoker	0.918*** [0.0879] z= 2.946†	0.966*** [0.108]	0.780*** [0.155]	0.370*** [0.0700] z= -2.720†	0.326*** [0.0824]	0.471*** [0.143]
Diabetes	1.004*** [0.0999] z= 1.745	0.931*** [0.136]	0.983*** [0.155]	0.793*** [0.0808] z= 1.609	0.803*** [0.0944]	0.803*** [0.159]
Log of Age	2.795*** [0.431] z= 0.884	3.121*** [0.551]	2.579*** [0.723]	2.966*** [0.320] z= -0.24	2.851*** [0.371]	3.337*** [0.655]
Log of T. cholesterol	0.552* [0.217] z= -1.923	0.736** [0.278]	0.320 [0.346]	0.957*** [0.173] z= -0.618	1.254*** [0.202]	0.0581 [0.346]
Log of HDL	-0.987*** [0.146] z= -1.232	-1.087*** [0.178]	-0.900*** [0.267]	-0.896*** [0.112] z= 0.203	-0.947*** [0.132]	-0.736** [0.236]
Log of non-treated SBP	2.156*** [0.349] z= -1.213	1.909*** [0.432]	2.273*** [0.633]	1.793*** [0.273] z= -0.350	1.902*** [0.322]	1.270* [0.555]
Log of treated SBP	3.209*** [0.386] z= 0.741	2.911*** [0.593]	3.065*** [0.525]	2.019*** [0.400] z= 0.041	1.799*** [0.495]	2.421*** [0.701]

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

Standard errors in brackets

z = score comparing published non-LB Framingham coefficients with those generated in ARIC using the formulae  $f(z) = (b_1 - b_2) / \sqrt{(se_{b1}^2 + se_{b2}^2)}$ ; † denotes z score >1.96 or < -1.96. Positive z means the regression coefficient in ARIC was greater than Framingham & vice versa

The unadjusted LB Framingham model also had a higher discrimination statistic and better calibration among women compared to men, with no significant differences in discrimination between blacks and whites within their respective sexes. Women had a C statistic of 0.76 (95% CI, 0.74-0.78) compared to men's 0.68 (95% CI, 0.67-0.70). When stratified by sex and race, white women had a C statistic of 0.754 (95% CI, 0.73-0.78) compared to black women's 0.750 (95% CI, 0.72-0.78), while white men had a C statistic of 0.685 (95% CI, 0.67-0.70) compared to black men's 0.69 (95% CI, 0.65-0.73). The published LB Framingham model had a C statistic of 0.793 (95% CI, 0.772-0.814) in women and 0.763 (95% CI, 0.746-0.780) for men.

Calibration was good ( $p > 0.05$ ) among women but poor among men with Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 10.5$  ( $p = 0.3084$ ) versus 21.8 ( $p = 0.0095$ ) respectively as depicted in Figure 9 and 10. When stratified by sex and race, calibration was good among all ARIC cohorts except white men with Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 25.5$  ( $p = 0.0024$ ). The published LB Framingham model had a Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 7.79$  for women ( $p = 0.56$ ) and 13.48 ( $p = 0.14$ ) for men.

Figure 9: Calibration- unadjusted LB Framingham women specific model

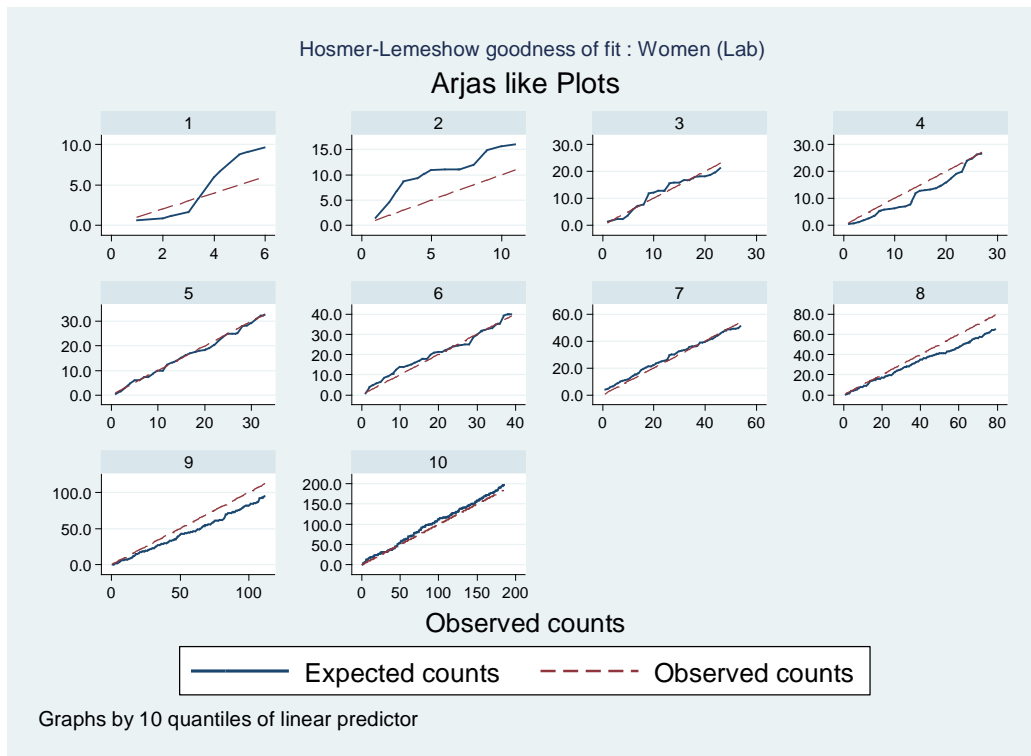
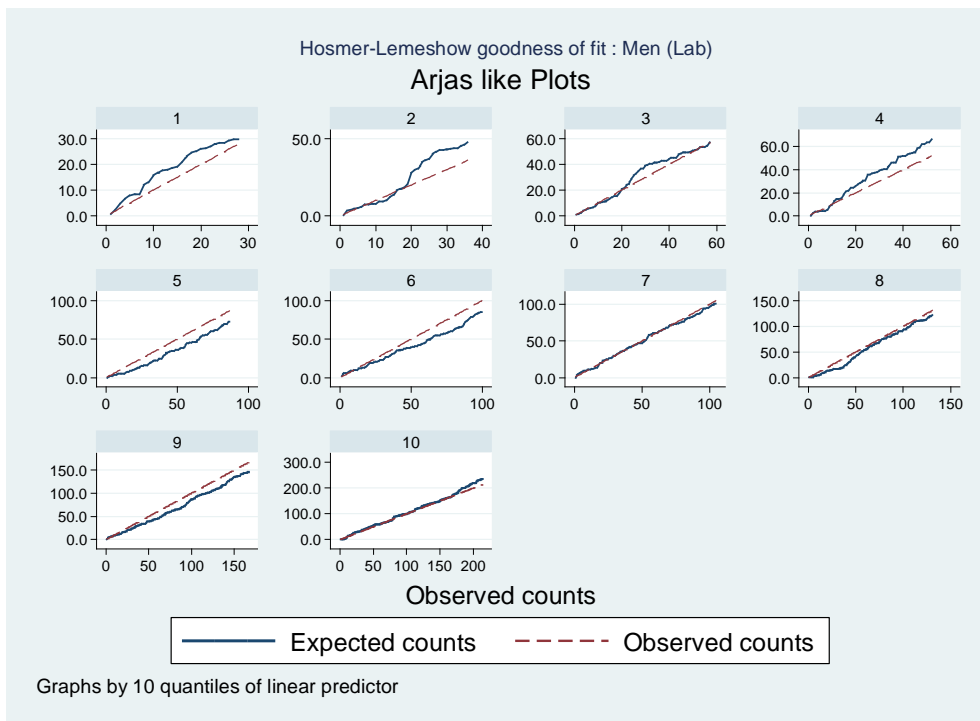


Figure 10: Calibration- unadjusted LB Framingham men specific model



### *The adjusted non-LB model*

In a bid to improve discrimination and calibration, additional variables known to be CVD risk factors were sequentially added to the covariates included in the published non-LB Framingham model, and their statistical significance evaluated. The additional variables included waist hip ratio and family history of premature CHD. A new model (adjusted non-LB model) including covariates that were statistically significant in multivariate regression was generated and its performance evaluated through the metrics of discrimination and calibration.

Family history of premature CHD was statistically significant among men, but not among women ( $p=0.068$ ). Waist hip ratio was marginally significant among women ( $p=0.052$ ), but the addition of waist hip ratio rendered BMI no longer significant ( $0.946$ ) among men. Table 6 describes each covariate's regression coefficient and the corresponding standard error and p value.

When stratified by sex and race, BMI was significant among white women but statistically not significant among white men ( $p=0.34$ ), black women ( $p=0.445$ ) and black men ( $p=0.535$ ). On the other hand, waist hip ratio was significant among blacks and white men, but statistically not significant among white women ( $p=0.338$ ). Family history of premature CHD was statistically not significant among white women ( $p=0.127$ ), black women ( $p=0.195$ ) and black men ( $p=0.971$ ), but significant among white men. In addition, untreated systolic blood pressure was not statistically significant among black men ( $p=0.101$ ). Table 6 describes the regression coefficients for the adjusted non-LB model, stratified by sex and race.

Table 6: Regression coefficients of the adjusted non-LB model stratified by sex and race

Adjusted non-LB model covariates	Female coefficients stratified by race				Male coefficients stratified by race		
	women	white women	black women		men	white men	black men
N	6353	4758	1597		4406	3612	1093
Current smoker	1.023*** [0.0895]	1.106*** [0.108]	0.830*** [0.154]		0.373*** [0.0771]	0.365*** [0.0882]	0.475*** [0.142]
Diabetes	1.028*** [0.103]	1.073*** [0.136]	0.982*** [0.156]		0.844*** [0.0897]	0.881*** [0.101]	0.773*** [0.161]
Log of Age	2.947*** [0.441]	3.333*** [0.545]	2.433*** [0.723]		2.772*** [0.354]	2.706*** [0.398]	2.997*** [0.658]
Log of BMI	0.597* [0.242]	0.797** [0.284]	N/A		N/A	N/A	N/A
Log of waist hip ratio	1.141 [0.587]	N/A	2.462** [0.906]		3.631*** [0.660]	3.799*** [0.757]	3.848** [1.229]
Family hx of premature CHD	N/A	N/A	N/A		0.334** [0.109]	0.359** [0.114]	N/A
Log of non-treated SBP	1.906*** [0.355]	1.808*** [0.435]	2.071** [0.635]		1.500*** [0.297]	1.641*** [0.350]	0.909 [0.554]
Log of treated SBP	3.097*** [0.390]	3.054*** [0.600]	3.092*** [0.531]		1.800*** [0.434]	1.803*** [0.519]	2.196** [0.698]

Standard errors in brackets, N/A= not applicable

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001, γ p≤0.1

There was no difference in discrimination between the adjusted and unadjusted non-LB models. Women maintained a C statistic of 0.75 (95% CI, 0.73-0.77) compared to men's 0.67 (95% CI, 0.66-0.69). When stratified by sex and race, there was no significant difference in discrimination. white women had a C statistic of 0.746 (95% CI, 0.72-0.77) compared to black women's 0.748 (95% CI, 0.71-0.78), while white men had a C statistic of 0.675 (95% CI, 0.65-0.69) compared to black men's 0.689 (95% CI, 0.65-0.72).

There was minimal improvement in calibration with women still maintaining good calibration compared to men with Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2$ = 12.4 (p=0.189) versus 21.2 (p=0.0115) respectively. When stratified by sex and race, calibration was good in all groups stratified by sex and race.

### *The adjusted LB model*

To improve the LB Framingham model, variables that included BMI, waist hip ratio, family history of premature CHD, apolipoprotein A and apolipoprotein B were added to the model. A new model (adjusted LB model) including covariates that were statistically significant in multivariate regression was generated and its performance evaluated through the metrics of discrimination and calibration.

Total cholesterol, waist hip ratio and family history of premature CHD were statistically significant in predicting CVD events among men, but not among women ( $p=0.765$  and  $0.327$  and  $0.106$  respectively). On the other hand, apolipoprotein B was significant among women but not among men ( $p=0.969$ ) as shown in Table 7. BMI and apolipoprotein A were not significant among women ( $p=0.592$  and  $0.585$  respectively) or men ( $p=0.288$  and  $0.938$  respectively). All other variables were significant in both sexes as described in Table 7.

When stratified by sex and race, apolipoprotein B was significant among white women but not significant among black women ( $p=0.238$ ), white men ( $p=0.953$ ) and black men ( $p=0.853$ ), waist hip ratio was significant among blacks and white men, but statistically not significant among white women ( $p=0.897$ ). Family history of premature CHD was statistically not significant among white women ( $p=0.148$ ), black women ( $p=0.224$ ) and black men ( $p=0.955$ ), but was significant among white men. In addition, untreated systolic blood pressure was marginally significant among black men ( $p=0.055$ ).



Table 7: Regression coefficients of the adjusted LB model stratified by race and sex

Adjusted LB model covariates	Female coefficients stratified by race			Male coefficients stratified by race		
	women	white women	black women	men	white men	black men
Current smoker	0.916*** [0.0880]	0.960*** [0.108]	0.771** [0.155]	0.368*** [0.0773]	0.332*** [0.0890]	0.475*** [0.142]
Diabetes	1.002*** [0.100]	0.935*** [0.136]	0.910*** [0.158]	0.787*** [0.0903]	0.806*** [0.102]	0.740*** [0.161]
Log of Age	2.809*** [0.429]	3.164*** [0.548]	2.482*** [0.722]	2.878*** [0.356]	2.814*** [0.399]	3.125*** [0.664]
Log of total cholesterol	N/A	N/A	N/A	0.812*** [0.192]	1.043*** [0.217]	N/A
Log of HDL	-0.877*** [0.151]	-0.955*** [0.185]	-0.770** [0.271]	-0.800*** [0.131]	-0.847*** [0.149]	-0.561* [0.244]
Log of ApoB	0.368** [0.141]	0.490** [0.183]	N/A	N/A	N/A	N/A
Log of waist hip ratio	N/A	N/A	1.980* [0.922]	2.360*** [0.706]	2.723*** [0.800]	2.958* [1.297]
Family hx of premature CHD	N/A	N/A	N/A	0.297** [0.109]	0.322** [0.114]	N/A
Log of non-treated SBP	2.154*** [0.349]	1.893*** [0.432]	2.119*** [0.641]	1.670*** [0.303]	1.628*** [0.353]	1.076 [0.561]
Log of treated SBP	3.181*** [0.387]	2.794*** [0.597]	3.019*** [0.525]	1.920*** [0.437]	1.838*** [0.519]	2.301** [0.705]
N	6352	4757	1597	4406	3612	1093

Standard errors in brackets, N/A= not applicable

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001, γ p<0.1

There was no difference in discrimination between the adjusted and unadjusted LB models. Women had a C statistic of 0.76 (95% CI, 0.74-0.78) compared to men's 0.69 (95% CI, 0.67-0.71). When stratified by sex and race, there was no significant difference in discrimination. white women had a C statistic of 0.755 (95% CI, 0.73-0.78) compared to black women's 0.753 (95% CI, 0.72-0.79), while white men had a C statistic of 0.693 (95% CI, 0.67-0.71) compared to black men's 0.693 (95% CI, 0.66-0.73).

Both sexes had similar calibration with Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 18.6$  among women (p=0.0289) versus 18.1 (p=0.0343) among men. When stratified by sex and race, calibration was good in all groups except among white men with Hosmer-Lemeshow goodness of fit (df 8)  $\chi^2 = 20.3$  (p=0.0164).

c) Applicability of the Framingham algorithms in the ARIC sample

In this section, the published (frozen) non-LB and LB Framingham algorithms are imputed in the ARIC dataset and applied to stratify CVD risk in the ARIC sample. In addition, two more risk prediction functions (recalibrated Framingham and adjusted algorithms) are derived using ARIC dataset, and their performance in stratifying CVD risk compared to the frozen Framingham algorithms using kappa statistic and sensitivity/specificity analysis. The three CVD risk prediction algorithms adhere to the general formulae outlined in the literature review section [  $\tilde{p} = 1 - S_0(t)^{\exp(\sum_{i=1}^p \beta_i X_i - \sum_{i=1}^p \beta_i \bar{X}_i)}$  ] which is widely used in Framingham and other studies to generate CVD risk prediction algorithms.

The frozen Framingham algorithms simply adopt all features of the published Framingham risk prediction functions, while the recalibrated algorithms alter the published Framingham functions by substituting their baseline survival and mean of risk factors with ARIC generated baseline survival and mean of risk factors. The adjusted algorithms alter the published Framingham risk function by substituting their baseline survival, covariates and regression coefficients with those generated in ARIC.

*CVD risk stratification using the frozen Framingham algorithms*

The frozen non-LB Framingham algorithm calculates the 10-year CVD risk for women as  $\tilde{p} = 1 - 0.94833^{\exp(\sum_{i=1}^p \beta_i X_i - 26.0145)}$ ; while the risk for men is calculated as

$\tilde{p} = 1 - 0.88431^{\exp(\sum_{i=1}^p \beta_i X_i - 23.9388)}$ . On the other hand, the frozen LB Framingham algorithm calculates the risk for women as  $\tilde{p} = 1 - 0.95012^{\exp(\sum_{i=1}^p \beta_i X_i - 26.1931)}$ , and the risk for men as  $\tilde{p} = 1 - 0.88936^{\exp(\sum_{i=1}^p \beta_i X_i - 23.9802)}$ . In addition to using the Framingham's baseline survival and mean of risk factors, the frozen Framingham algorithms also use the Framingham generated regression coefficients ( $\beta_i$ ) in the equations above.

When the frozen Framingham algorithms were applied to stratify CVD risk in the entire ARIC sample, the outcome was comparable in all risk categories. For instance, the non-LB Framingham algorithm classified 18% of the sample as high or very high risk (see Figure 11 below) compared to 17% by the frozen LB Framingham algorithm (Figure 12 below). When the four risk categories stratified by the two frozen Framingham algorithms were compared by kappa test, there was an overall agreement of 92.76% and a kappa statistic of 0.7624 as described in Table 8 below.

Figure 11: Risk stratification in ARIC cohort as per the Frozen non-LB Framingham

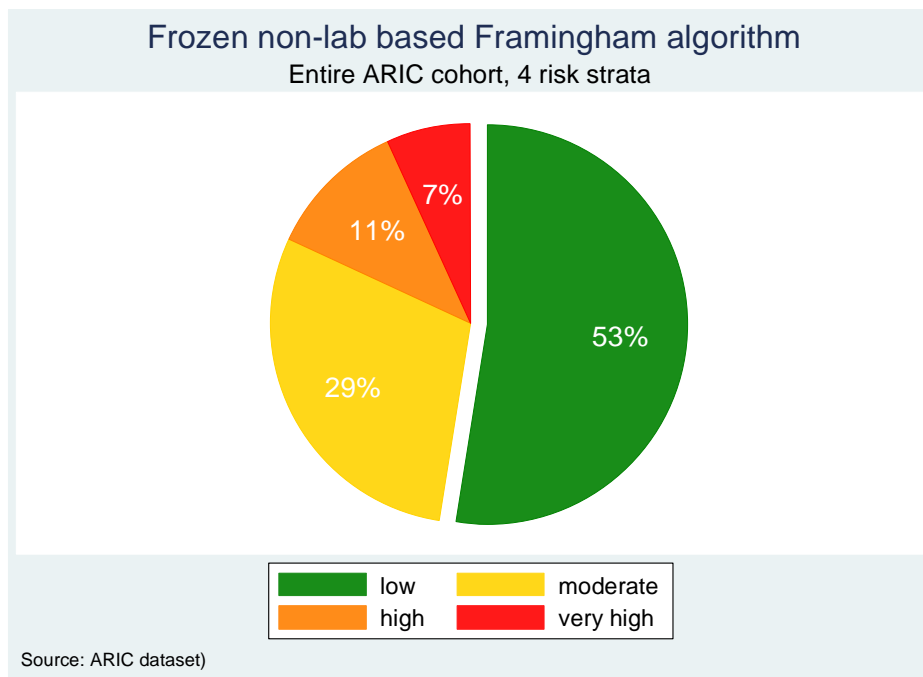


Figure 12: Risk stratification in ARIC cohort as per the Frozen LB Framingham

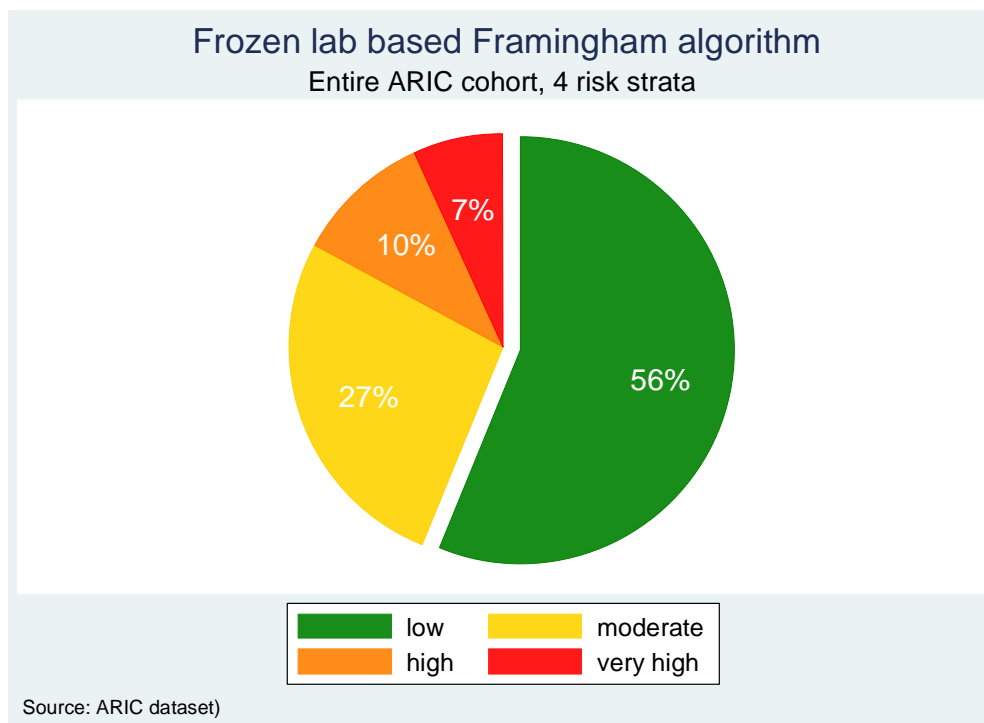


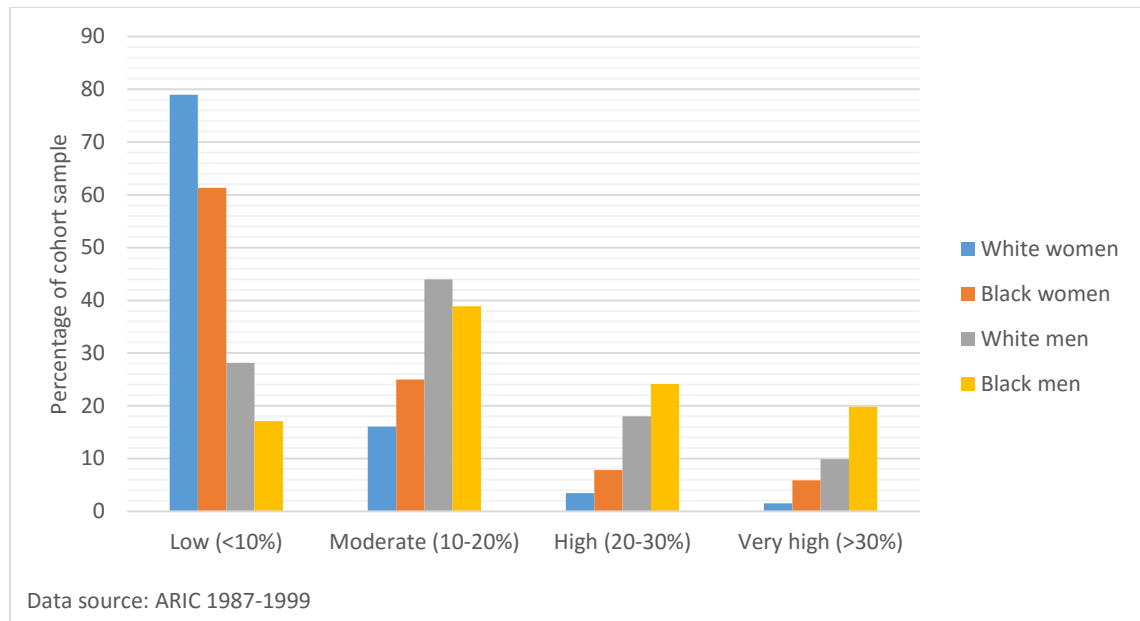
Table 8: Kappa test frozen non-LB versus LB Framingham algorithm's risk categories

Frozen non-LB based	Frozen LB Framingham algorithm				
	low	moderate	high	very	Total
low	5,696	809	9	0	6,514
moderate	392	2,241	443	31	3,107
high	3	350	642	198	1,193
very high	0	16	208	563	787
total	6,091	3,416	1,302	792	11,601
Ratings weighted by:					
1.0000	0.6667	0.3333	0.0000		
0.6667	1.0000	0.6667	0.3333		
0.3333	0.6667	1.0000	0.6667		
0.0000	0.3333	0.6667	1.0000		
Agreement	Expected Agreement	Kappa	Std. Err.	Z	Prob>Z
92.76%	69.56%	0.7624	0.0069	111.28	0.0000

When the frozen non-LB Framingham algorithm was used to classify CVD risk in the ARIC sample stratified by sex and race, a greater proportion of black women and black men were classified as high or very high risk compared to their white counterparts. The proportion of black women in the high risk category was 4.3% more than the proportion of white women (7.8% versus 3.5% respectively), while the proportion of black women in the very high risk category was 4.4% more than the proportion of white women (5.9% versus 1.5% respectively) as depicted in Figure 13 below.

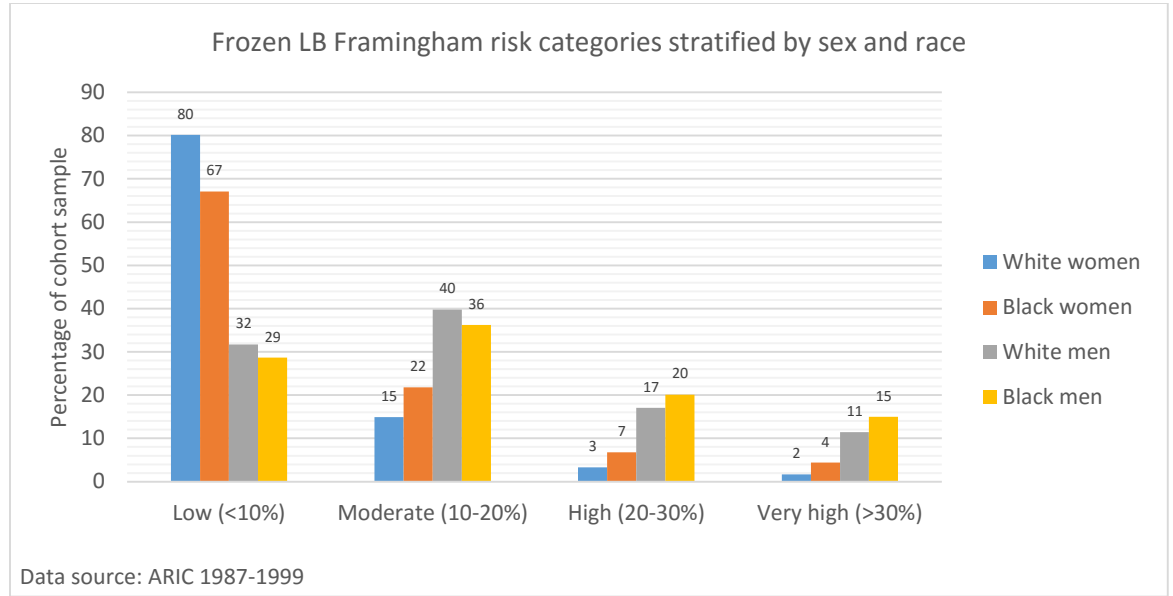
Among men, blacks dominated the high and very high risk categories. The proportion of black men in the high risk category was 6.2% higher compared to white men (24.2% versus 18.0% respectively). In the very high risk category, the proportion of black men was 10% higher than the proportion of white men (19.9% versus 9.9% respectively) as described in Figure 13.

Figure 13: Frozen non-LB Framingham algorithm risk categories stratified by sex & race



When the frozen LB Framingham algorithm was applied to classify CVD risk in the ARIC sample stratified by sex and race, higher proportions of blacks dominated the high and very high risk categories compared to whites. The proportion of black women in the high risk category was 3.5% higher than the proportion of white women (6.8% versus 3.3% respectively), while the proportion of black women in the very high risk category was 2.8% higher than the proportion of white women (4.4% versus 1.6% respectively) as depicted in Figure 14 below. Among men, the differences in the high and very high risk categories were minimal, with 20.1% of black men and 17.1% of white men being classified as high risk (3% difference), and 15% of black men versus 11.4% of white men classified in the very high risk category (3.6% difference) as depicted in Figure 14.

Figure 14: Frozen LB Framingham algorithm risk categories stratified by sex & race



### *CVD risk stratification using the recalibrated Framingham algorithms*

Recalibration of the Framingham algorithms was done by replacing their baseline survival and mean of risk factors with values generated in the ARIC sample, but the Framingham generated regression coefficients ( $\beta_i$ ) were retained. In the ARIC sample, the baseline survival adjusted for the non-LB covariates was 0.9111038 for women and 0.8335824 for men. The baseline survival adjusted for the LB covariates was 0.9033732 and 0.8461685 for women and men respectively in the ARIC sample.

Therefore, the recalibrated non-LB algorithm calculates the 10-year CVD risk for women as  $\tilde{p} = 1 - 0.9111038^{\exp(\sum_{i=1}^p \beta_i X_i - 27.127605)}$ ; while the risk for men is calculated as  $\tilde{p} = 1 - 0.8335824^{\exp(\sum_{i=1}^p \beta_i X_i - 24.78308)}$ . On the other hand, the recalibrated LB algorithm calculates the risk for women as  $\tilde{p} = 1 - 0.9033732^{\exp(A \sum_{i=1}^p \beta_i X_i - 27.18902)}$ , and the risk for men as  $\tilde{p} = 1 - 0.8461685^{\exp(\sum_{i=1}^p \beta_i X_i - 24.780015)}$ .

When the recalibrated algorithms were applied to stratify CVD risk in the entire ARIC sample, the outcome was comparable in all risk categories. For instance, the recalibrated non-LB algorithm classified 7.44% of the sample as high or very high risk compared to 8.09% by the recalibrated LB algorithm. When the four risk categories stratified by the two recalibrated algorithms were compared by kappa test, there was an overall agreement of 94.26% and a kappa statistic of 0.72.

When the recalibrated algorithms were applied to classify CVD risk in the ARIC sample stratified by sex and race, there was no difference between the non-LB and LB algorithm, or between blacks and whites. However, higher proportions of men were classified in the high and very high risk strata compared to women in both racial groups by both algorithms as described in Figures 15 and 16.

Figure 15: Recalibrated non-LB algorithm risk categories stratified by sex & race

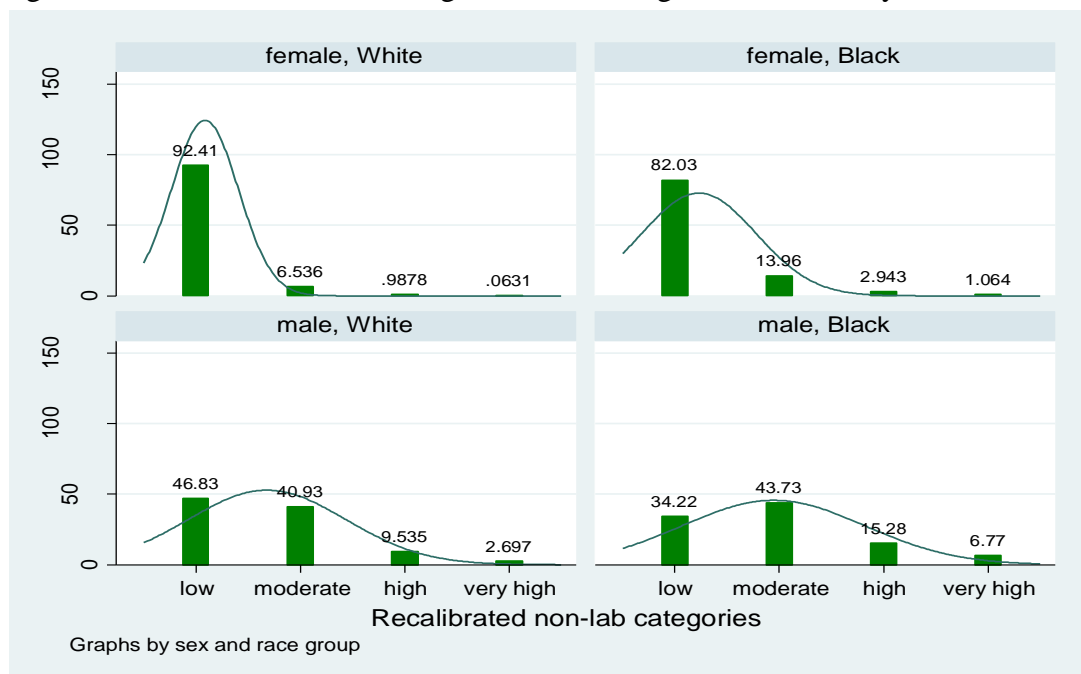
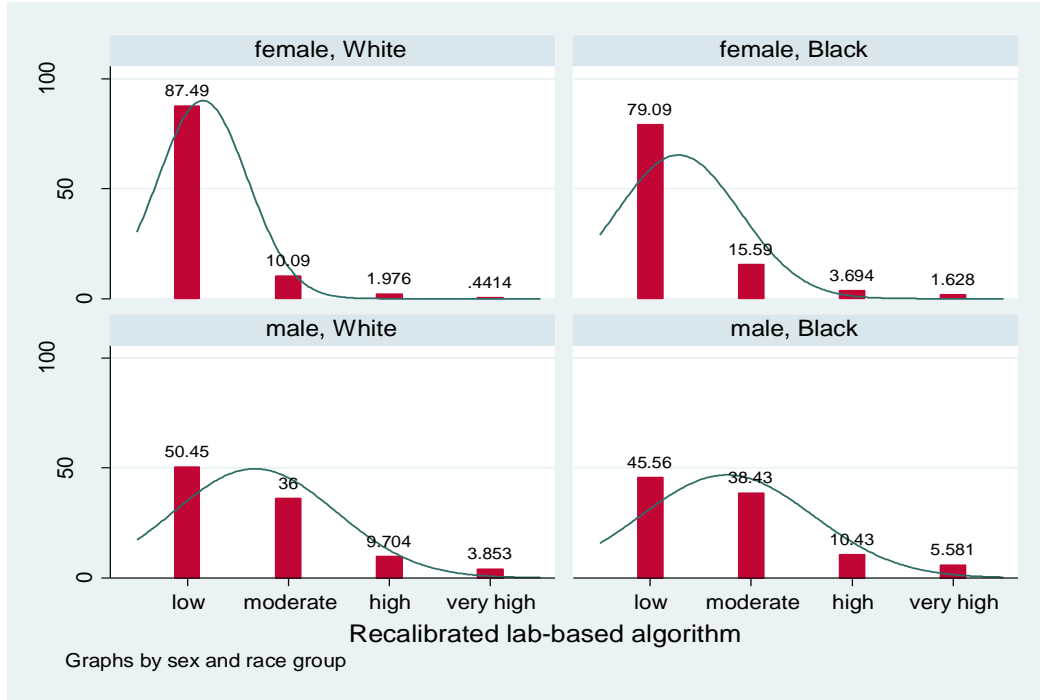




Figure 16: Recalibrated LB algorithm risk categories stratified by sex & race



### *CVD risk stratification using the adjusted algorithms*

The adjusted algorithms were generated by replacing the baseline survival, mean of risk and regression coefficients of the general CVD formulae with values generated in the ARIC sample. In the ARIC sample, the baseline survival adjusted for the additional non-LB covariates (described in the mathematical performance section) was 0.908755 for women and 0.8518082 for men. The baseline survival for the adjusted LB covariates was 0.9032351 and 0.8568301 for women and men respectively in the ARIC sample.

Therefore, the adjusted non-LB algorithm calculates the 10-year CVD risk for women as  $\tilde{p} = 1 - 0.908755 \exp(\sum_{i=1}^p \beta_i X_i - 26.73064)$ ; while the risk for men is calculated as  $\tilde{p} = 1 - 0.8518082 \exp(\sum_{i=1}^p \beta_i X_i - 21.26414)$ . On the other hand, the adjusted LB algorithm calculates the risk for women as  $\tilde{p} = 1 - 0.9032351 \exp(\sum_{i=1}^p \beta_i X_i - 27.747355)$ , and the risk for men as  $\tilde{p} = 1 - 0.8568301 \exp(\sum_{i=1}^p \beta_i X_i - 23.347965)$ .

When the adjusted algorithms were applied to stratify CVD risk in the entire ARIC sample, there was no difference between categories. For instance, the adjusted non-LB algorithm classified 6.88% of the sample as high or very high risk compared to 5.90% by the adjusted LB algorithm. When the four risk categories stratified by the two algorithms were compared by kappa test, there was an overall agreement of 96.12% and a kappa statistic of 0.7996.

When the adjusted algorithms were applied to classify CVD risk in the ARIC sample stratified by sex and race, there was no difference between the non-LB and LB algorithm, or between blacks and whites. However, higher proportions of men were classified in the high and very high risk strata compared to women in both racial groups by both algorithms as described in Figures 17 and 18.

Figure 17: Adjusted non-LB algorithm risk categories stratified by sex & race

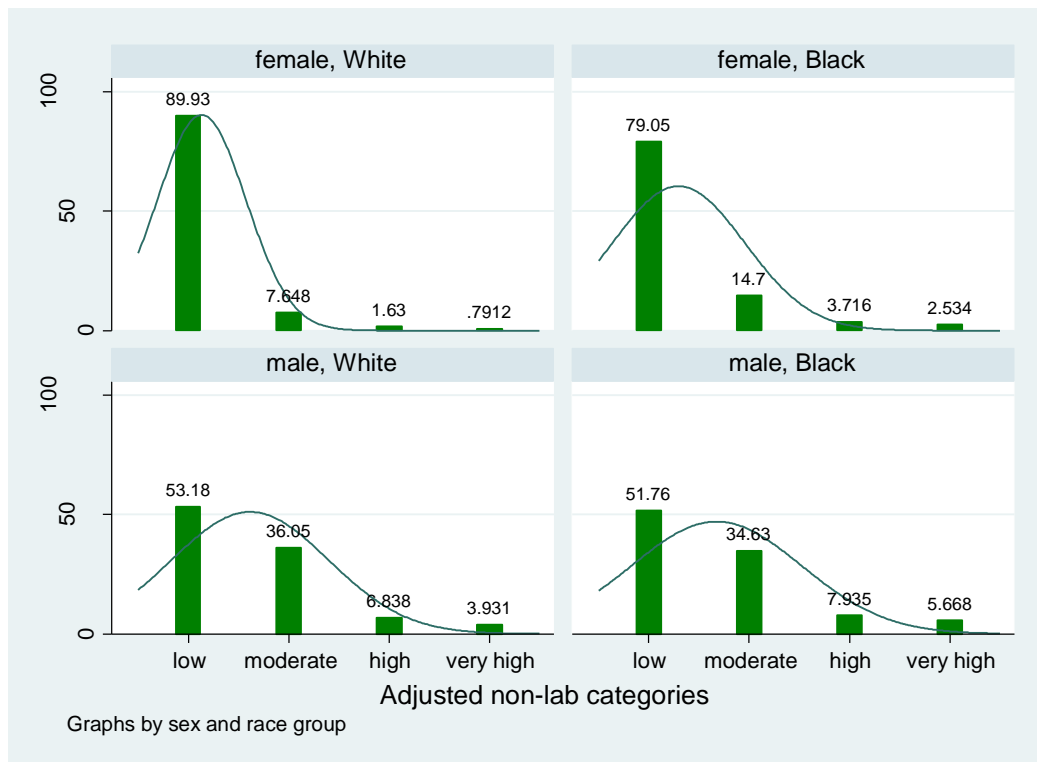
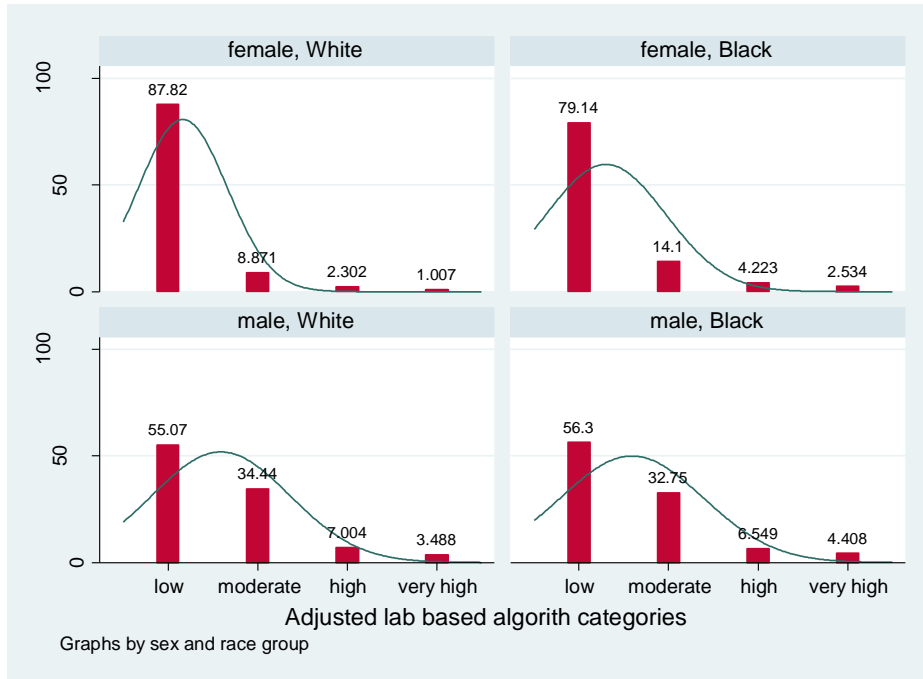


Figure 18: Adjusted LB algorithm risk categories stratified by sex & race



#### d) Sensitivity/specificity analysis

Sensitivity/specificity analysis was done to determine the clinical usefulness of the non-LB and LB versions of the three algorithms described above. Sensitivity of each algorithm was determined by the proportion of the sample with incident CVD who are correctly identified as high risk by the algorithm. Specificity was dictated by the proportion of the sample without incident CVD who were classified as low risk by the algorithm. The roctab command in Stata© was used to calculate and plot the nonparametric AUROC based on sensitivity/specificity analysis for each algorithm.

#### *Sensitivity/specificity analysis of the frozen Framingham algorithms*

Although the Framingham algorithms are ideally intended to predict CVD risk within a time frame of 10 years, they can be extended to a maximum of a 12 years as

discussed in the methods section. Therefore, the analysis described in this section focuses on their sensitivity/specificity analysis within 12 years' time frame.

*The frozen non-LB Framingham algorithm*

In the entire ARIC sample, the frozen non-LB Framingham algorithm had an overall AUROC of 0.7063. At the moderate risk category (10-20%), the algorithm had a sensitivity of 76% and specificity of 57.03%. When the high risk (20-30%) cut point was used as the threshold for predicted incident CVD, the algorithm had a sensitivity of 41.10% versus specificity of 85.49%. Table 9 below describes a detailed report of the algorithm's sensitivity and specificity at the four different risk cut points described in the methods section, while Figure 19 plots the resultant AUROC.

Table 9: Sensitivity/specificity analysis of the frozen non-LB Framingham algorithm

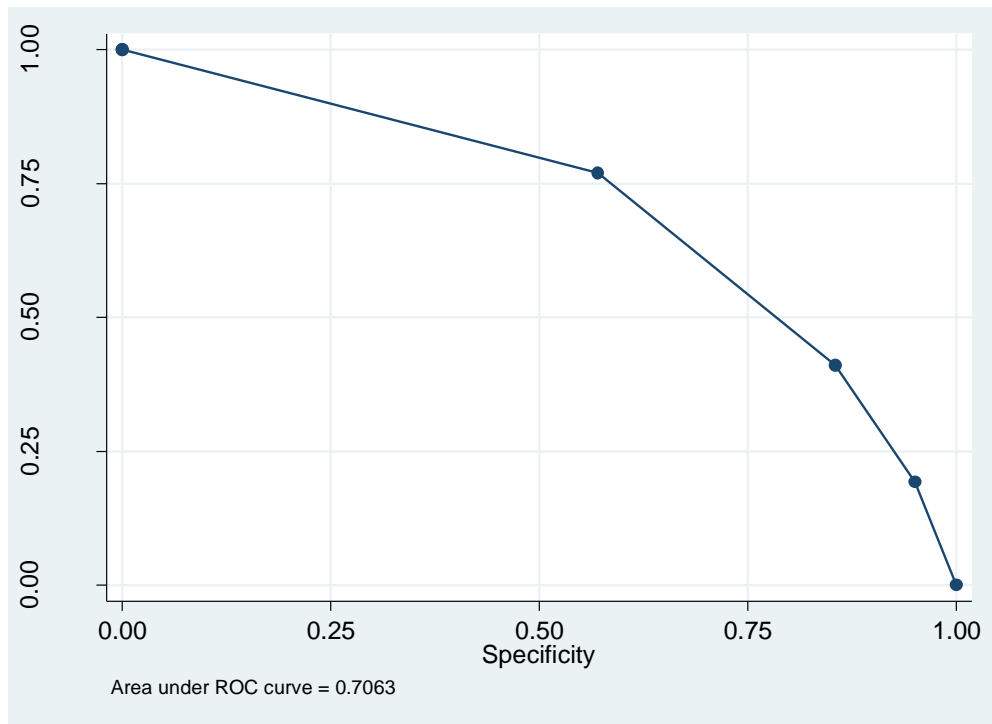
Any CVD event within first 12yrs of follow-up	Frozen non-LB based Framingham risk categories in ARIC				
	low	moderate	high	very high	Total
CVD free	5,735	2,862	965	494	10,056
Incident CVD	356	554	337	298	1,545
Total	6,091	3,416	1,302	792	11,601
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
>=low risk (<10%)	100.00%	0.00%	13.32%	1.0000	
>=moderate risk (10-20%)	76.96%	57.03%	59.68%	1.7910	0.4040
>=high risk (20-30%)	41.10%	85.49%	79.58%	2.8328	0.6890
>=very high risk (>30% )	19.29%	95.09%	84.99%	3.9263	0.8488
> very high risk	0.00%	100.00%	86.68%	1.0000	
ROC -Asymptotic Normal--					
	Obs	Area	Std. Err.	[95% Conf. Interval]	
	11,601	0.7063	0.0068	0.69294	0.71969

Sensitivity= fraction of true positive cases

Specificity= fraction of true negative cases

Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

Figure 19: Frozen non-LB Framingham AUROC for entire ARIC sample



When the ARIC sample was stratified by sex and race, the algorithm's AUROCs for women and blacks were higher compared to men and whites respectively. Figures 20-23 plots the frozen non-LB based Framingham AUROCs stratified by sex and race.

Figure 20: Frozen non-LB Framingham AUROC white women

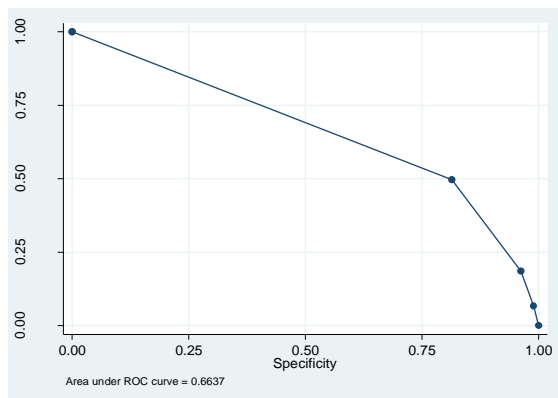


Figure 21: Frozen non-LB Framingham AUROC black women

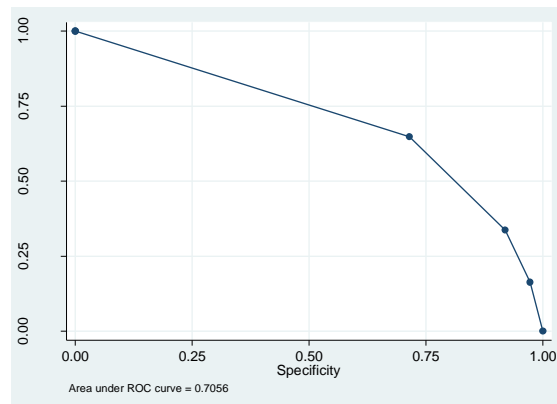


Figure 22: Frozen non-LB Framingham AUROC white men

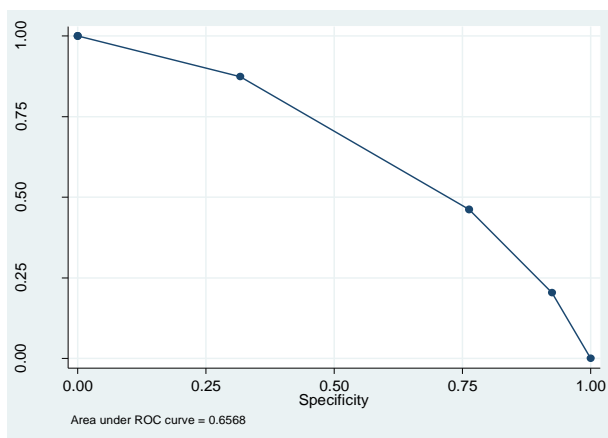
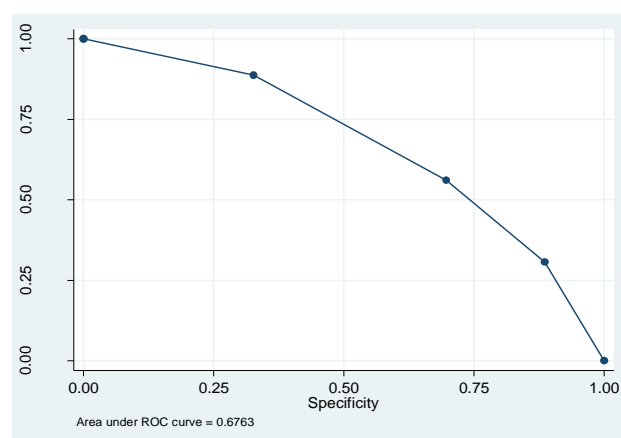


Figure 23: Frozen non-LB Framingham AUROC black men



### *The frozen LB Framingham algorithm*

In the entire ARIC sample, the frozen LB Framingham algorithm had an overall AUROC of 0.71. At the moderate risk (10-20%) threshold, the algorithm had a sensitivity of 74.63% and specificity of 60.88%. When the high risk (20-30%) cut point is used as the threshold for predicted incident CVD, the algorithm had a sensitivity of 40.32% versus specificity of 86.51%. Table 10 below describes a detailed report of the algorithm's sensitivity and specificity at the four different risk cut points described in the methods section, while Figure 24 plots the resultant AUROC. Both frozen Framingham algorithms (non-LB and LB) had similar AUROCs with no statistical difference in their sensitivity and specificity ( $p=0.3661$ ) as depicted in Figure 25.

Table 10: Sensitivity/specificity analysis of the frozen LB Framingham algorithm

Any CVD event within first 12yrs of follow-up	Frozen LB based Framingham risk categories in ARIC				
	low	moderate	high	very high	Total
CVD free	6,122	2,577	880	477	10,056
Incident CVD	392	530	313	310	1,545
Total	6,514	3,107	1,193	787	11,601
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
$\geq$ low risk (<10%)	100.00%	0.00%	13.32%	1.0000	
$\geq$ moderate risk (10-20%)	74.63%	60.88%	62.71%	1.9076	0.4168
$\geq$ high risk (20-30%)	40.32%	86.51%	80.36%	2.9882	0.6899
$\geq$ very high risk (>30% )	20.06%	95.26%	85.24%	4.2300	0.8392
> very high risk	0.00%	100.00%	86.68%		1.0000
ROC -Asymptotic Normal--					
	Obs	Area	Std. Err.	[95% Conf. Interval]	
	11,601	0.7100	0.0068	0.69661	0.72345

Sensitivity= fraction of true positive cases

Specificity= fraction of true negative cases

Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

Figure 24: Frozen LB Framingham AUROC for entire ARIC sample

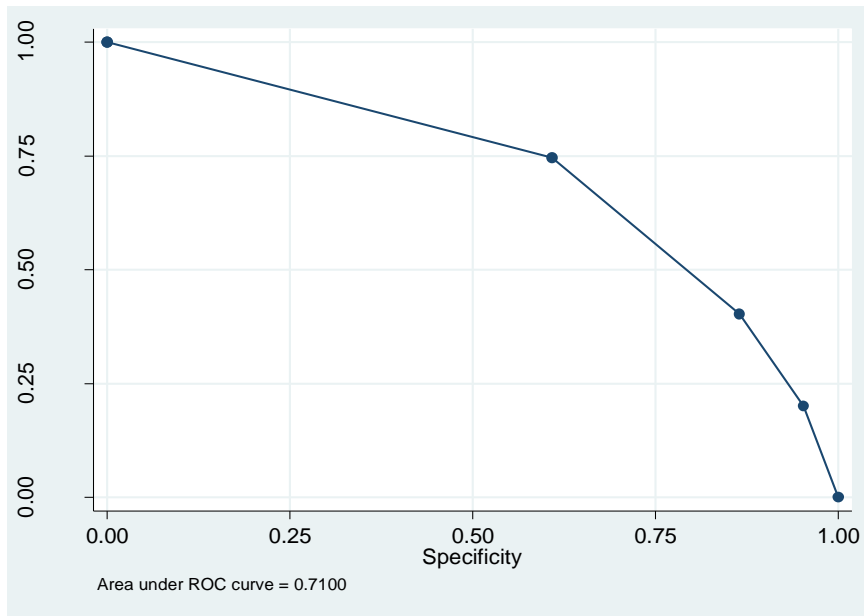
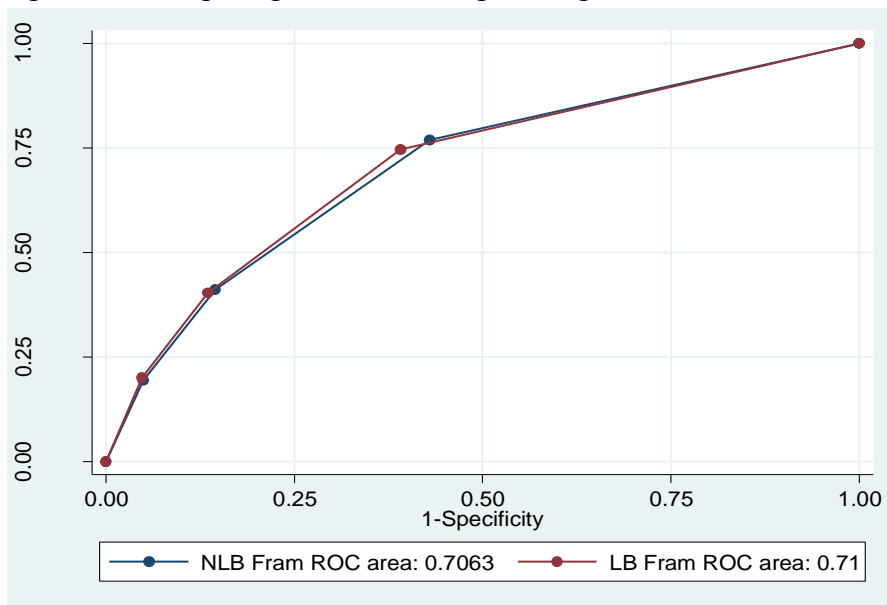


Figure 25: Comparing frozen Framingham algorithms' AUROCs in entire ARIC sample



When the ARIC sample was stratified by sex and race, the frozen LB Framingham algorithm's AUROCs for blacks were higher compared to whites. Figures 26-28 plots these AUROCs stratified by sex and race.



Figure 26: Frozen LB Framingham AUROC  
white women

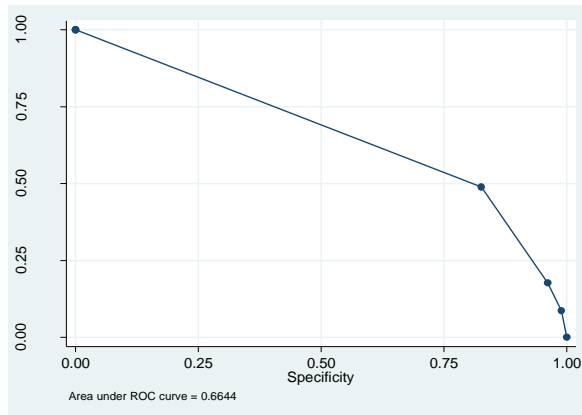


Figure 27: Frozen LB Framingham AUROC  
black women

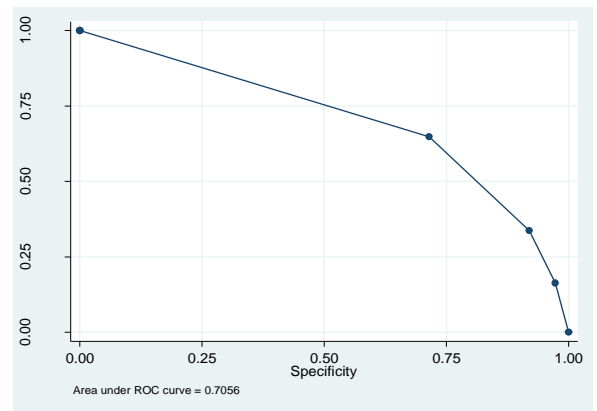


Figure 28: Frozen LB Framingham AUROC  
white men

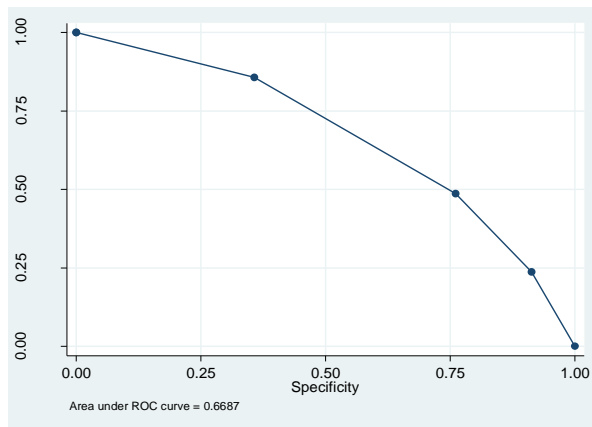
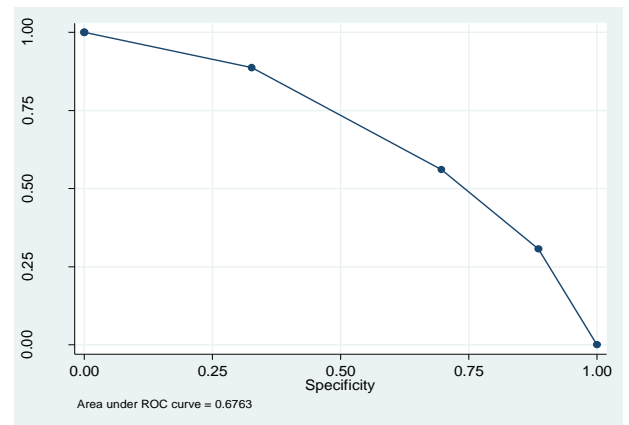


Figure 29: Frozen LB Framingham AUROC  
black men



## *Sensitivity/specificity analysis of the recalibrated algorithms*

### *The recalibrated non-LB algorithm*

In the entire ARIC sample, the recalibrated non-LB algorithm had an overall ROC of 0.6711. At the moderate risk category (10-20%), the algorithm had a sensitivity of 58.58% and specificity of 73.45%. When the high risk (20-30%) cut point was used as the threshold for predicted incident CVD, the algorithm had a sensitivity of 20.13% versus specificity of 94.51%. Table 11 below describes a detailed report of the algorithm's sensitivity and specificity at the four different risk cut points, while Figure 30 plots the resultant AUROC.

Applied to the ARIC sample stratified by sex and race, the recalibrated algorithm had the highest AUROC for black women (0.6691) followed by black men (0.6577). white men had a higher AUROC (0.6383) compared to white women (0.5890).

Table 11: Sensitivity/specificity analysis of the recalibrated non-LB algorithm

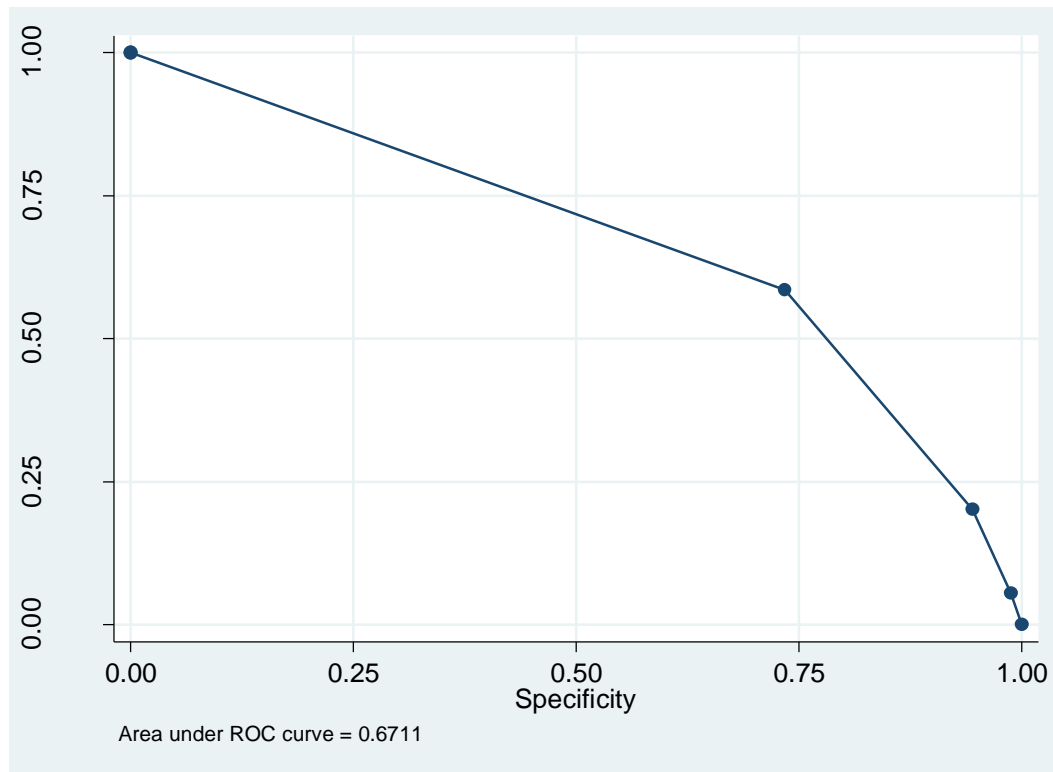
Any CVD event within first 12yrs of follow-up	Recalibrated non-LB algorithm risk categories in ARIC				
	low	moderate	high	very high	Total
CVD free	7,386	2,118	431	121	10,056
Incident CVD	640	594	226	85	1,545
Total	8,026	2,712	657	206	11,601
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
>=low risk (<10%)	100.00%	0.00%	13.32%	1.0000	
>=moderate risk (10-20%)	58.58%	73.45%	71.47%	2.2061	0.5640
>=high risk (20-30%)	20.13%	94.51%	84.60%	3.6671	0.8451
>=very high risk (>30% )	5.50%	98.80%	86.37%	4.5723	0.9565
> very high risk	0.00%	100.00%	86.68%		1.0000
ROC -Asymptotic Normal--					
	Obs	Area	Std. Err.	[95% Conf. Interval]	
	11,601	0.6711	0.0069	0.65745	0.68469

Sensitivity= fraction of true positive cases

Specificity= fraction of true negative cases

Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

Figure 30: Recalibrated non-LB Framingham AUROC for entire ARIC sample



#### *The recalibrated LB algorithm*

In the entire ARIC sample, the recalibrated LB algorithm had an overall AUROC of 0.6851. At the moderate risk category (10-20%), the algorithm had a sensitivity of 60.84% and specificity of 73.73%. When the high risk (20-30%) cut point was used as the threshold for predicted incident CVD, the algorithm had a sensitivity of 22.46% versus specificity of 94.12%. Table 12 below describes a detailed report of the algorithm's sensitivity and specificity at the four different risk cut points, while Figure 31 plots the resultant AUROC.

Applied to the ARIC sample stratified by sex and race, the recalibrated LB algorithm had similar AUROC for black men (0.6648) and white men AUROC (0.6644). black women had a higher AUROC (0.6622) compared to white women (0.6230).

Table 12: Sensitivity/specificity analysis of the recalibrated LB algorithm

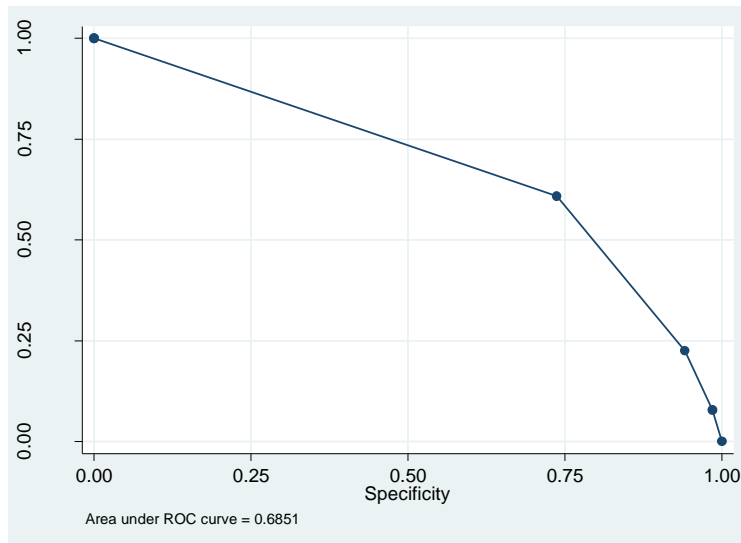
Any CVD event within first 12yrs of follow-up	Recalibrated LB risk categories in ARIC				
	low	moderate	high	very high	Total
CVD free	7,414	2,051	444	147	10,056
Incident CVD	605	593	226	121	1,545
Total	8,019	2,644	670	268	11,601
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
>=low risk (<10%)	100.00%	0.00%	13.32%	1.0000	
>=moderate risk (10-20%)	60.84%	73.73%	72.01%	2.3158	0.5311
>=high risk (20-30%)	22.46%	94.12%	84.58%	3.8215	0.8238
>=very high risk (>30% )	7.83%	98.54%	86.46%	5.3575	0.9354
> very high risk	0.00%	100.00%	86.68%		1.0000
<div>ROC -Asymptotic Normal--</div> <div>Obs Area Std. Err. [95% Conf. Interval]</div> <div>-----</div> <div>11,601 0.6851 0.0069 0.67158 0.69868</div>					

Sensitivity= fraction of true positive cases

Specificity= fraction of true negative cases

Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

Figure 31: Recalibrated LB Framingham AUROC for entire ARIC Cohort



## *Sensitivity/specificity analysis of the adjusted algorithms*

### *The adjusted non-LB algorithm*

In the entire ARIC sample, the adjusted non-LB algorithm had an overall AUROC of 0.6768. At the moderate risk category (10-20%), the algorithm had a sensitivity of 56.77% and specificity 76.20%. When the high risk (20-30%) cut point was used as the threshold for predicted incident CVD, the algorithm had a sensitivity of 21.02% versus specificity of 95.23%. Table 13 below describes a detailed report of the algorithm's sensitivity and specificity at the four different risk cut points, while Figure 32 plots the resultant AUROC.

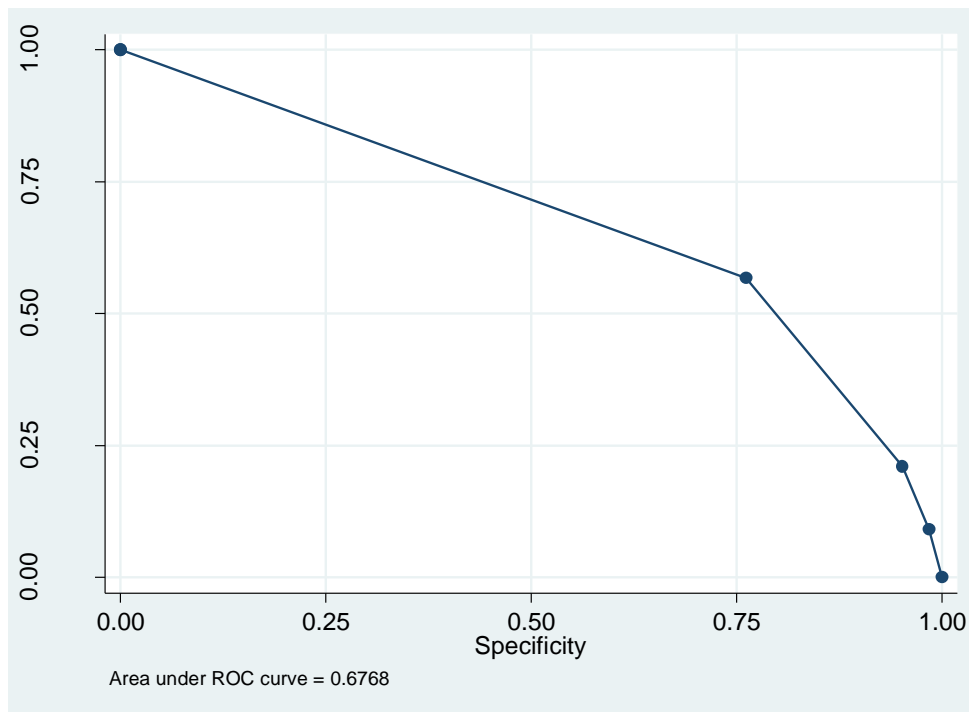
When applied to the ARIC sample stratified by sex and race, the adjusted non-LB algorithm had the highest AUROC in black women (0.6793), while white women had the lowest AUROC (0.6066). White men had similar AUROC (0.6568) with black men (0.6580).

Table 13: Sensitivity/specificity analysis of the adjusted non-LB algorithm

Any CVD event within first 12yrs of follow-up	Adjusted non-LB categories in ARIC				
	low	moderate	high	very high	Total
CVD free	6,470	1,616	271	134	8,491
Incident CVD	549	454	151	116	1,270
Total	7,019	2,070	422	250	9,761
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
>=low risk (<10%)	100.00%	0.00%	13.01%	1.0000	
>=moderate risk (10-20%)	56.77%	76.20%	73.67%	2.3852	0.5673
>=high risk (20-30%)	21.02%	95.23%	85.58%	4.4077	0.8293
>=very high risk (>30% )	9.13%	98.42%	86.80%	5.7877	0.9232
> very high risk	0.00%	100.00%	86.99%		1.0000
<div style="display: flex; justify-content: space-between;"> <div>ROC</div> <div>-Asymptotic Normal--</div> </div> <div style="display: flex; justify-content: space-between;"> <div>Obs</div> <div>Area</div> <div>Std. Err.</div> <div>[95% Conf. Interval]</div> </div> <div style="border-top: 1px dashed black; display: flex; justify-content: space-between; margin-top: 5px;"> <div>9,761</div> <div>0.6768</div> <div>0.0077</div> <div>0.66183</div> <div>0.69187</div> </div>					

Sensitivity= fraction of true positive cases; Specificity= fraction of true negative cases  
 Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

Figure 32: Adjusted non-LB algorithm AUROC for entire ARIC sample



#### *The adjusted LB algorithm*

In the entire ARIC sample, the adjusted LB algorithm had an overall AUROC of 0.6908. At the moderate risk category (10-20%), the algorithm had a sensitivity of 58.98% and specificity of 76.73%. When the high risk (20-30%) cut point was used as the threshold for predicted incident CVD, the algorithm had a sensitivity of 22.05% versus specificity of 95.24%. Table 14 below describes a detailed report of the algorithm's sensitivity and specificity at the four different risk cut points, while Figure 33 plots the resultant AUROC.

Applied to the ARIC sample stratified by sex and race, the adjusted LB algorithm had the highest AUROC for black women (0.7033) followed by white men (0.6783).

black men had a higher AUROC (0.6411) compared to white women (0.6317).

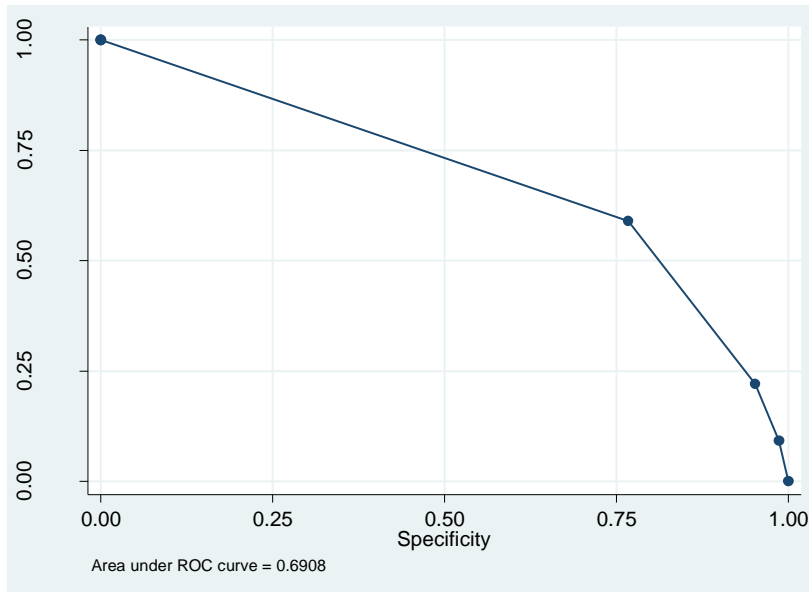
Table 14: Sensitivity/specificity analysis of the adjusted LB algorithm

Any CVD event within first 12yrs of follow-up	Adjusted LB categories in ARIC				
	low	moderate	high	very high	Total
CVD free	6,515	1,572	288	116	8,491
Incident CVD	521	469	163	117	1,270
Total	7,036	2,041	451	233	9,761
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
>=low risk (<10%)	100.00%	0.00%	13.01%	1.0000	
>=moderate risk (10-20%)	58.98%	76.73%	74.42%	2.5343	0.5347
>=high risk (20-30%)	22.05%	95.24%	85.72%	4.6337	0.8185
>=very high risk (>30% )	9.21%	98.63%	87.00%	6.7435	0.9204
> very high risk	0.00%	100.00%	86.99%		1.0000
<div>ROC</div> <div>Obs    Area    Std. Err.    -Asymptotic Normal--</div> <div>   [95% Conf. Interval]</div> <div>-----</div> <div>9,761   0.6908   0.0076   0.67592   0.70574</div>					

Sensitivity= fraction of true positive cases; Specificity= fraction of true negative cases

Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

Figure 33: Adjusted LB algorithm AUROC for entire ARIC sample



*Comparing sensitivity/specificity of all algorithms*

When all non-LB algorithms (frozen, recalibrated and adjusted) were compared in the entire ARIC sample, the frozen non-LB Framingham had the highest AUROC (0.7086), while the recalibrated model had the lowest (0.6697) as described in Figure 34 below. Among the LB algorithms, the frozen non-LB algorithm had the highest AUROC (0.7141) while the recalibrated algorithm had the lowest (0.6868) as depicted in Figure 35 below.

When the frozen LB Framingham algorithm was compared to all non-LB algorithms in the entire ARIC sample, the frozen LB Framingham had the highest AUROC (0.7141) followed by the frozen non-LB Framingham (0.7086). The recalibrated non-LB model had the lowest AUROC (0.6697) as described in Figures 36 and 37.



Figure 34: Comparing all non-LB algorithms

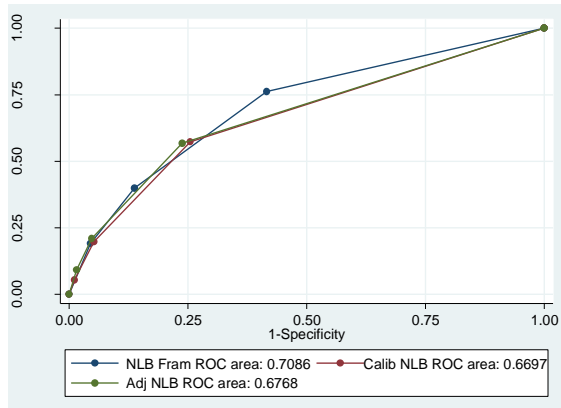


Figure 35: Comparing all LB algorithms

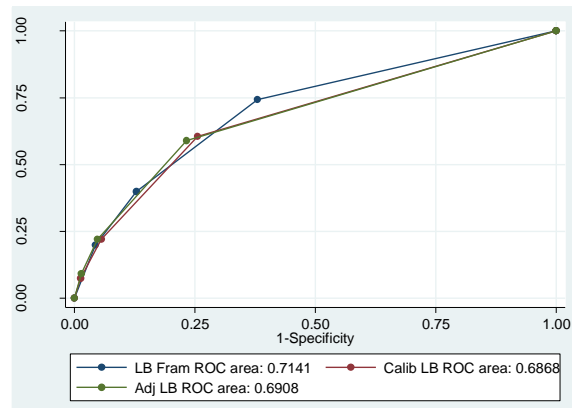


Figure 36: Comparing frozen LB with all non-LB algorithms

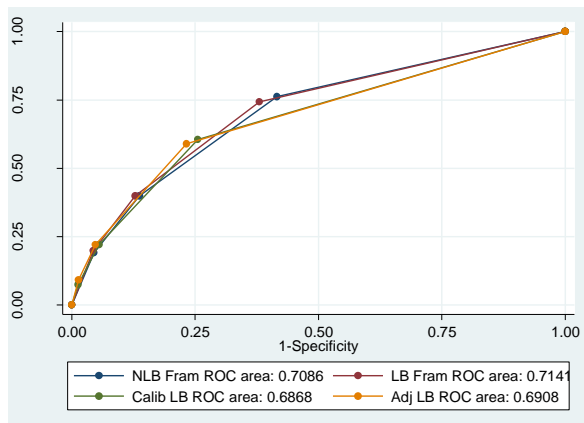
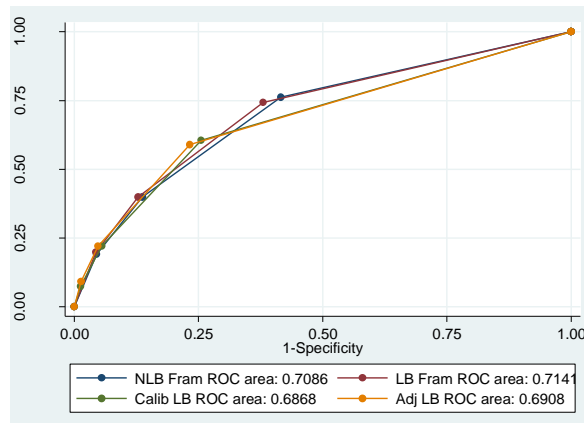


Figure 37: Comparing frozen non-LB with all LB algorithms



e) Simulated cost-effectiveness analysis of the non-LB Framingham algorithm

*Introduction*

To determine the cost-effectiveness of the non-LB Framingham algorithm, the expenses and outcomes associated with the three simulated CVD prevention programs described in the methods were calculated. The expenses were costed at three levels which are consistent with the steps inherent in CVD prevention programs. Level I expenses are the screening costs and are determined by the unit cost of screening an individual in each program. Level II expenses are the costs of preventive interventions prescribed in each program and are driven by the number of true and false positive cases associated with each program. Level III expenses includes the downstream costs of treating false negative cases associated with each preventive program. The outcomes describing the effectiveness of each program were quantified by identifying the true positive cases associated with each preventive program.

The costs and outcomes associated with each program were used to compute their respective average cost-effectiveness ratios. Finally, incremental cost-effectiveness analysis was done where existing programs (status quo) were considered to be the individual risk factors and LB absolute risk approaches to CVD prevention, while the new program was considered to be the non-LB absolute risk approach.

*Level I: Expenses in screening for CVD*

Level I expenses are the over-time, non-recurring marginal costs incurred during a patient's first routine office visit when the provider screens a patient to assess CVD risk. These expenses include the extra Registered Nurses (RN) hours spent taking a patient's medical history and obtaining physiological measures to appraise CVD risk. The office

visit itself is not costed because the cost is incurred regardless of whether CVD screening takes place or not.

It has been estimated that an individual's absolute risk score can be calculated within ten minutes using non-LB algorithms because the only data required include: sex, age, smoking status, diabetes status, antihypertensive medication use, systolic blood pressure and BMI (D'Agostino RB et al., 2008; Gaziano et al., 2008). For purposes of this study, the RN time required to appraise CVD risk based on individual risk factors (diabetes and/or hypertension) was estimated to be 5 minutes because only the demographic and relevant history/physiologic data are collected without calculating any risk scores. The RN screening time using the LB Framingham algorithm was assumed to be similar to the non-LB algorithm since the extra time for assessing lipids was costed under the laboratory expenses.

The basic equipment needed for either preventive approach (i.e. weighing scale, tape measure, glucometer and sphygmomanometer) are readily available in most primary care offices; hence no additional capital inputs are required for screening. The diagnostic cost of blood glucose testing for diabetes is required for the three CVD prevention programs. The LB absolute risk approach has additional diagnostic costs for HDL and total cholesterol tests. The absolute CVD risk score can be calculated for both approaches by an interactive online calculator or an offline calculator embedded in a downloadable excel spread sheet. Alternatively, the absolute CVD risk scores can be calculated manually by a paper based tool that aggregates points associated with each covariate included in the Framingham algorithms. Both the online and offline calculators, and the paper tool can be accessed for free at the Framingham heart study website. This analysis

assumed use of the offline calculators, because of their efficiency and applicability to settings without internet connection.

All costs were based on Mississippi payment rates because most of the black participants in ARIC were recruited from Jackson, Mississippi (ARIC Investigators, 1989). The RN hourly wage is based on Bureau of Labor Statistics median RN hourly wage for the state of Mississippi which is \$27.19/hr. (Bureau of Labor Statistics, 2014). Since the Centers for Medicare and Medicaid (CMS) have well established reimbursement rates for diagnostic tests, screening for diabetes and lipids were costed based on the 2016 clinical diagnostic laboratory fee schedule in Mississippi. Reimbursement for diabetes was calculated using Current Procedure Terminology (CPT code 82962 (\$3.19) while HDL and total cholesterol tests were costed under CPT codes 83718 (\$11.16) and 82465 (\$5.92) respectively (Centers for Medicare & Medicaid Services, 2015a). Medicare rates were assumed close to true marginal costs because many tests are conducted at these rates which are lower than with private insurance.

The screening costs are calculated for each CVD prevention strategy in equations 1a-c and summarized in Table 15. Both approaches will incur costs for screening the 2,690 eligible black sample in the ARIC dataset.

Equation 1a: The individual CVD risk factors (DM/HTN) approach:

$$\text{Screen}_r = 2690 \text{ individuals} * [(5\text{min} * \$0.45) + \$3.19] = \$14,634$$

Equation 1b: The non-LB absolute CVD risk approach:

$$\text{Screen}_a = 2690 \text{ individuals} * [(10\text{min} * \$0.45) + \$3.19] = \$20,686$$

Equation 1c: The LB absolute CVD risk approach:

$$\text{Screen}_l = 2690 \text{ individuals} * [(10\text{min} * \$0.45) + \$3.19 + \$11.16 + \$5.92] = \$66,631$$

Table 15: Marginal screening costs: individual risk factors vs absolute non-LB vs LB absolute risk approach

Costed screening items	Individual risk factors (DM/HTN)	Absolute CVD risk (non-LB Framingham )	Absolute CVD risk (LB Framingham )
RN screening min/person	5 mins	10 mins	10 mins
RN hourly wage (\$27.19/hr.)	\$0.45/min	\$0.45/min	\$0.45/min
Fasting glucose test cost	\$3.19	\$3.19	\$3.19
HDL test	\$0	\$0	\$11.16
Total cholesterol test	\$0	\$0	\$5.92
Total Screening costs/person	\$5.44	\$7.69	\$24.77
Total Screening costs	\$14,634	\$20,686.1	\$66,631.3

Screening time is estimated as suggested by Gaziano et al. (2008)

RN wages are the 2014 average RN wages for Mississippi reported by Bureau of Labor Statistics.

Fasting glucose, HDL and total cholesterol test costs are based on reimbursement for CPT codes 82962, 83718 and 82465 respectively in Mississippi

### *Level II: Expenses of initial and follow-up visits for positive cases*

If a patient is determined to be at risk for CVD, guidelines recommend a battery of follow-up steps. Initial and follow-up office visits are costed using the Medicare physician fee schedule. The drugs prescribed by each CVD prevention approach (see Figure 3 and Table 2) are identified by their National Drug Code (NDC), and costed using the National Average Drug Acquisition Cost (NADAC) compiled by Medicaid in the last week of December 2015 (Centers for Medicare & Medicaid Services, 2015c; Medicaid.gov, 2016).

Level II expenses are influenced by the number of participants classified above a certain risk threshold and the cost of interventions prescribed by each CVD preventive strategy (see Table 2 for absolute CVD risk programs and Figure 3 for individual risk factors program). A CVD preventive strategy with many false positives unnecessarily increases level II expenses with fewer prevented CVD events.

The moderate risk category (absolute CVD risk score  $\geq 10$ ) was selected as the optimal risk threshold for the absolute CVD risk programs and the high risk category (presence of diabetes and/or hypertension) for the individual CVD risk factor program. These thresholds were based on sensitivity/specificity analyses outlined in Table 16 (for the individual CVD risk factors program) and Table 17 and 18 (for the absolute CVD risk programs).

When the high risk category was applied as the treatment threshold in the approach based on treating individual CVD risk factors, there were 1045 true negatives and 88 false negatives, along with 313 true positives and 1244 false positives as detailed in Table 16. The false and true negatives (1133 cases) were not be prescribed the preventive treatments outlined in Figure 3, and did not contribute to level II expenses. On the other hand, the false and true positives (1557) received the preventive interventions and contributed to level II expenses.

Table 16: Sensitivity/specificity analysis of the individual risk factors strategy

Any CVD event within first 12yrs of follow-up			Individual risk factors (DM/HTN) risk categories in ARIC			
			low	High	Total	
CVD free			1,045	1,244	2,289	
Incident CVD			88	313	401	
Total			1,133	1,557	2,690	
Detailed report of sensitivity and specificity						
Cut point	Sensitivity	Specificity	Correctly Classified		LR+	LR-
>=low risk	100.00%	0.00%	14.91%		1.0000	
>=high risk	78.05%	45.65%	50.48%		1.4362	0.4807
> high risk	0.00%	100.00%	85.09%			1.0000
ROC -Asymptotic Normal--						
Obs		Area	Std. Err.	[95% Conf. Interval]		
2,690		0.6185	0.0116	0.59584	0.64124	

Sensitivity= fraction of true positive cases

Specificity= fraction of true negative cases

Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

Using the non-LB moderate risk category, the absolute CVD risk prevention strategy has 1094 true negatives and 72 false negatives, along with 329 true positives and 1195 false positives as detailed in Table 17. The true and false positives (1524) contribute to level II interventions because they receive the preventive interventions described in Table 2 depending on absolute CVD risk score. The false positive category drives up level II expenses while not preventing CVD events.

Table 17: Sensitivity/specificity analysis of the non-LB guided absolute CVD risk strategy

Any CVD event within first 12yrs of follow-up	Frozen non-LB based Framingham risk categories in ARIC				
	low	moderate	high	very high	Total
CVD free	1,094	704	296	195	2,289
Incident CVD	72	120	93	116	401
Total	1,166	824	389	311	2,690
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
>=low risk (<10%)	100.00%	0.00%	14.91%	1.0000	
>=moderate risk (10-20%)	82.04%	47.79%	52.90%	1.5716	0.3757
>=high risk (20-30%)	52.12%	78.55%	74.61%	2.4298	0.6096
>=very high risk (>30% )	28.93%	91.48%	82.16%	3.3957	0.7769
> very high risk	0.00%	100.00%	85.09%		1.0000
ROC -Asymptotic Normal--					
	Obs	Area	Std. Err.	[95% Conf. Interval]	
	2,690	0.7061	0.0136	0.67938	0.73277

Sensitivity= fraction of true positive cases

Specificity= fraction of true negative cases

Correctly classified= percentage correctly classified in their true disease state (CVD or no CVD)

When the moderate risk category of the LB Framingham algorithm was applied as the treatment threshold, the absolute CVD risk prevention strategy guided by the algorithm had 1,292 true negatives and 92 false negatives, along with 309 true positives and 997 false positives as detailed in Table 18. The false and true negatives (1,384 cases) did not contribute to level II expenses, since no preventive interventions were prescribed

for this group. The false and true positives (1,306 cases) received the preventive interventions described in Table 2 and hence contributed to level II expenses.

Table 18: Sensitivity/specificity analysis of the LB guided absolute CVD risk strategy

Any CVD event	Frozen lab based Framingham risk categories in ARIC				
	low	moderate	high	very high	Total
CVD free	1,292	616	242	139	2,289
Incident CVD	92	128	86	95	401
Total	1,384	744	328	234	2,690
Detailed report of sensitivity and specificity					
Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
>=low risk (<10%)	100.00%	0.00%	14.91%	1.0000	
>=moderate risk (10-20%)	77.06%	56.44%	59.52%	1.7692	0.4065
>=high risk (20-30%)	45.14%	83.36%	77.66%	2.7118	0.6582
>=very high risk (>30% )	23.69%	93.93%	83.46%	3.9013	0.8124
> very high risk	0.00%	100.00%	85.09%		1.0000
ROC -Asymptotic Normal--					
	Obs	Area	Std. Err.	[95% Conf. Interval]	
	2,690	0.7077	0.0136	0.68097	0.73441

Whereas the true and false negatives in all programs do not contribute to level II expenses because follow-up is not recommended, the false negative cases end up missing the preventive interventions required to prevent CVD events. Treatment of CVD events observed among the false negative cases contribute to level III expenses.

Level II expenses are subcategorized into costs associated with initial and follow-up office visits, and costs associated with the drugs prescribed by each preventive strategy. These costs are discussed and calculated below. Although in clinical settings the CPT coding varies depending on specific problems and complexity of the office visit, this analysis assumed uniform complexity of all visits using CPT code 99203 that requires medical decision making of moderate complexity. The visit typically lasts for 45 minutes and includes face-to-face counseling and/or coordination of care with other providers



(Centers for Medicare & Medicaid Services, 2015b). The 2015 physician office reimbursement for CPT code 99203 was \$100.28 in Mississippi (Centers for Medicare & Medicaid Services, 2015c).

*Costs associated with initial office visits*

The initial office visits in this analysis are expected to occur immediately after screening for all individuals who meet the treatment threshold. The initial office visits are costed on an annual basis for each CVD prevention strategy in equation 2a-c below.

The individual CVD risk factors approach **initial** office visit costs:

For the individual risk factors approach, the initial visits included a total of 1557 individuals (true and false positives) who met the treatment threshold for the strategy discussed above. The total costs of these initial office visits (IVrc) are estimated in equation 2a:

$$f(IVrc) = 1557 \text{ individuals} * 1 \text{ visit} * \$100.28 = \$156,135.96$$

The non-LB absolute CVD risk strategy **initial** office visit costs:

For the non-LB absolute CVD risk strategy, initial visits included a total of 1524 individuals (true and false positives) who met the treatment threshold. The total costs of these initial office visits (IVac) are estimated in equation 2b:

$$f(IVac) = 1524 \text{ individuals} * 1 \text{ visit} * \$100.28 = \$152,826.72$$

The LB absolute CVD risk strategy **initial** office visit costs:

For the LB absolute CVD risk strategy, initial visits included a total of 1,306 individuals (true and false positives) who met the treatment threshold. The total costs of these initial office visits (IVlc) are estimated in equation 2c:

$$f(IVlc) = 1306 \text{ individuals} * 1 \text{ visit/yr} * \$100.28 = \$130,965.68$$

### *Costs associated with follow-up office visits*

The follow-up office visits in this analysis were expected to occur after the initial office visit for all positive cases in each program. The follow-up office visits were scheduled at different intervals based on each individual's absolute CVD risk score (for the absolute CVD risk based programs) as summarized in Table 2, or as per the IDF and ISH-ASH guidelines (for the individual risk factors program) outlined in Figure 3. The follow-up office visits are detailed and costed on an annual basis for each CVD prevention strategy in equation 3-5.

The individual CVD risk factors approach **follow-up** costs:

For the individual CVD risk factors strategy, follow-up office visits were only relevant to individuals with diabetes and/or hypertension (high CVD risk category). Since the IDF and ISH-ASH guidelines do not explicitly recommend a specific follow-up regimen, this analysis used the follow-up schedule recommended by the American Diabetes Association (ADA). The association recommends twice a year office visits for diabetic patients with stable glycemic control (American Diabetes Association, 2014). Follow-up for hypertension is assumed to follow the diabetes schedule. As explained earlier, this analysis assumed similar complexity of all office visits. Therefore, all the follow-up office visits were costed under CPT code 99203.

The high risk category (presence of diabetes and/or hypertension) identified 1557 true and false positive cases eligible for follow-up office visits scheduled every 6 months. The annual costs for the office visits are estimated in equation 3 and shown in Table 19 as discounted costs.

$$FV_{rc} = 1557 \text{ individuals} * 2 \text{ visits/yr} * \$100.28 = \$312271.92$$

The non-LB absolute CVD risk strategy **follow-up** costs

Follow-up office visits includes 824 individuals in the moderate risk category, 389 in the high risk category and 311 in the very high risk category. The follow-up costs (FV<sub>ac</sub>) associated with each risk category are estimated in equations (4a-c) and summarized in Table 19 as discounted costs.

For the moderate risk category (10-20%), 824 follow-up office visits for treatment and CVD risk reduction discussion with a primary care provider are scheduled annually. The annual cost for these office visits is estimated in equation 4a:

$$FV_{ac-mod} = 824 \text{ individuals} * 1 \text{ visit/yr} * \$100.28 = \$82630.72$$

For the high risk category (20-30%), 389 follow-up office visits for treatment and CVD risk reduction discussion with a primary care provider are scheduled every 6 months. The annual costs for these office visits are estimated in equation 4b:

$$FV_{ac-high} = 389 \text{ individuals} * 2 \text{ visits/yr} * \$100.28 = \$78017.84$$

For the very high risk category (>30), 311 follow-up office visits for treatment and CVD risk reduction discussion with a primary care provider are scheduled every 6 months. The annual costs for these office visits are estimated in equation 4c:

$$FV_{ac-vhigh} = 311 \text{ individuals} * 2 \text{ visits/yr} * \$100.28 = \$62374.16$$

Total annual cost of the follow-up office visits in all risk categories are calculated by summation of equations 4a-c and shown in Table 19 as discounted costs.

$$\sum \text{eq } 3_{a-c} = \$82630.72 + \$78017.84 + \$62374.16 = \$223,022.72$$

The LB absolute CVD risk strategy **follow-up** costs

Follow-up office visits includes 744 individuals in the moderate risk category, 328 individuals in the high risk category and 234 individuals in the very high risk category. The follow-up costs (FV<sub>lc</sub>) associated with each risk category are estimated in equations 5a-c and summarized in Table 19 as discounted costs.

For the moderate risk category (10-20%), 744 Follow-up office visits for treatment and CVD risk reduction discussion with a primary care provider are scheduled annually. The annual costs for these office visits are estimated in equation 5a:

$$FV_{lc-mod} = 744 \text{ individuals} * 1 \text{ visit/yr} * \$100.28 = \$74608.32$$

For the high risk category (20-30%), 328 Follow-up office visits for treatment and CVD risk reduction discussion with a primary care provider are scheduled every 6 months. The annual costs for these office visits are estimated in equation 5b:

$$FV_{lc-high} = 328 \text{ individuals} * 2 \text{ visits/yr} * \$100.28 = \$65783.68$$

For the very high risk category (>30), 234 Follow-up office visits for treatment and CVD risk reduction discussion with a primary care provider are scheduled every 6 months. The annual costs for these office visits are estimated in equation 5c:

$$FV_{lc-vhigh} = 234 \text{ individuals} * 2 \text{ visits/yr} * \$100.28 = \$46931.04$$

Total annual cost of the follow-up office visits in all risk categories are calculated by summation of equations 5a-c:

$$\sum \text{eq } 15_{a-c} = \$74608.32 + \$65783.68 + \$46931.04 = \$187,323.04$$

The annual costs associated with each year's follow-up office visits are divided by the 3% discounting rate discussed under the analysis framework. Therefore, in Table 19 the discounted annual costs for year 1 visits are divided by  $1.03^1$  and the costs for year 12 by  $1.03^{12}$  to get the respective discounted total costs for follow-up office visits. All costs are based on 2015 prices.

Table 19: Discounted costs of follow-up office visits: individual risk factors vs absolute non-LB vs LB absolute risk approach

Year	Individual CVD risk factors approach	non-LB absolute CVD risk approach	LB absolute CVD risk approach
	Discounted annual costs	Discounted annual costs	Discounted annual costs
Year 1	\$303,177	\$216,527	\$181,867
Year 2	\$294,346	\$210,220	\$176,570
Year 3	\$285,773	\$204,097	\$171,427
Year 4	\$277,450	\$198,153	\$166,434
Year 5	\$269,369	\$192,381	\$161,586
Year 6	\$261,523	\$186,778	\$156,880
Year 7	\$253,906	\$181,338	\$152,311
Year 8	\$246,510	\$176,056	\$147,875
Year 9	\$239,330	\$170,928	\$143,568
Year 10	\$232,360	\$165,950	\$139,386
Year 11	\$225,592	\$161,116	\$135,326
Year 12	\$219,021	\$156,424	\$131,385
Total	\$3,108,356	\$2,219,969	\$1,864,614

Annual visits=True & false positives adjusted by recommended frequency of follow-up visits annually

Discounted costs=Annual Follow-up costs discounted by  $(1.03)^t$  where  $t$ = year 1 through 12

Source: Visits based on total annual office visits expected in the ARIC cohort based on their risk profile; cost based on Medicare physician fees schedule for CPT code 99203.

Annual visits=True & false positives adjusted by recommended frequency of follow-up visits annually

Discounted costs=Annual follow-up costs discounted by  $(1.03)^t$  where  $t$ = year 1 through 12

Source: Visits based on total annual office visits expected in the ARIC cohort based on their risk profile; cost based on Medicare physician fees schedule for CPT code 99203.

### *Costs associated with preventive interventions*

Besides the follow-up office visits, each CVD prevention strategy has its own set of preventive interventions as described in the methods section. The costs associated with these preventive interventions are described below.

#### *The Individual CVD risk factors strategy treatment expenses*

##### i. Antihypertensive therapy expenses:

The ISH-ASH guidelines recommend different treatment options depending on the stage of hypertension (Weber et al., 2014). In stage I hypertension (BP>140/90), monotherapy with either a thiazide-like diuretic or a CCB is recommended for blacks with or without comorbid diabetes. In stage II hypertension (BP>160/100), combined therapy is recommended with the second drug being an ACEI for individuals with comorbid diabetes (Weber et al., 2014).

To ensure consistency in this analysis, the costing of the monotherapy treatment outlined in Figure 3 was done using Hydrochlorothiazide 25mg/day, while the combined therapy included Hydrochlorothiazide 25mg/day with Amlodipine 5mg/day in absence of comorbid diabetes, or Hydrochlorothiazide 25mg/day with Lisinopril 10mg/day in comorbid diabetes. The NADAC of Hydrochlorothiazide ((NDC 00143125601) is \$0.01192 per 25mg tablet, Lisinopril (NDC 00143126701) \$0.02011 per 10mg tablet and Amlodipine (NDC 76282023890) \$0.01839 per 5mg tablet (Medicaid.gov, 2016).

For the individual CVD risk factors prevention program, the high risk category (treatment threshold) included 1,419 individuals with blood pressure greater or equal to 140/90 mmHg. Of these individuals, 1,230 had stage I hypertension (BP  $\geq$  140/90 < 160/100mmHg) while 189 had stage II hypertension (BP  $\geq$  160/100mmHg).

In the group with stage I hypertension, 959 individuals were diabetes free, while 271 had comorbid diabetes. The group with stage II hypertension had 134 diabetes free individuals and 55 with comorbid diabetes. The annual costs for treating hypertension in the high risk (DM/HTN) category are estimated in equations 6-7 and shown in Table 20 as discounted costs.

Stage I hypertension category was prescribed monotherapy with Hydrochlorothiazide 25mg/day even if they had comorbid diabetes as explained above. The annual costs for treating stage I hypertension are calculated in equation 6:

$$RxhtnI_{\pm dm} = 1230 \text{ individuals} * \$0.01192 * 365 \text{ days} = \$5351.484$$

Stage II hypertension category was prescribed combined therapy with addition of Amlodipine or Lisinopril depending on whether or not they have comorbid diabetes as explained above. The annual costs for treating stage II hypertension are calculated in equation 7a-b.

Equation 7a: Combined antihypertensive therapy for hypertension stage II without comorbid diabetes:

$$RxhtnII_{-dm} = 134 \text{ individuals} * \$0.01192 + \$0.01839 * 365 \text{ days} = \$1482.4621$$

Equation 7b: Combined antihypertensive therapy for hypertension stage II with comorbid diabetes:

$$RxhtnII_{+dm} = 55 \text{ individuals} * \$0.01192 + \$0.02011 * 365 \text{ days} = \$643.00225$$

The total annual costs of antihypertensive therapy in the individual CVD risk factor approach were calculated by summing up equations 6-7 and summarized in Table 20 as discounted costs.

$$\Sigma eq_{6-7} = \$5351.48 + \$1482.46 + \$643 = \$7,476.9484$$

ii. Diabetes treatment and monitoring expenses

The IDF guidelines recommend initiating pharmacotherapy for type 2 diabetes at the threshold of  $>7\text{mmol/l}$  ( $>126\text{ mg/dL}$ ) as outlined in Figure 3. The recommended first line oral hypoglycemic agent is metformin (IDF Clinical Guidelines Task Force, 2006). To ensure consistency in this analysis, costing for diabetes treatment was done using the common prescribed first line drug start dose, i.e. Metformin 850mg/day (NDC 00093104910). Metformin NADAC is \$0.03292 per 850mg tablet (Centers for Medicare & Medicaid Services, 2015a).

Although the frequency and intensity of glucose monitoring varies depending on the plan of care, the IDF guidelines recommend periodic monitoring of glycated hemoglobin (HbA1C) in all people with type 2 diabetes as part of comprehensive management program (IDF Clinical Guidelines Task Force, 2006). In absence of an explicit monitoring schedule by the IDF guidelines, this analysis assumed twice a year monitoring of HbA1C in diabetic patients with stable glycemic control as recommended by the ADA (American Diabetes Association, 2014). The costing for the HbA1C test was done using CPT code 83036 QW (\$13.22) in Mississippi as detailed under the absolute CVD risk factor approach.

For the individual CVD risk factor approach, the high risk category (treatment threshold) included 464 individuals with diabetes. The annual costs for treating and monitoring type 2 diabetes are calculated in equation 8a-b.

Equation 8a: Cost of treating diabetes with Metformin:

$$\text{Rxdma} = 464 \text{ individuals} * \$0.03292 * 365 \text{ days} = \$5575.3312$$



Equation 8b: The cost of monitoring diabetes using the HbA1C test:

$$\text{HbA1C}_{\text{testa}} = 464 \text{ individuals} * \$13.22 * 2 \text{ tests/year} = \$12268.16$$

The total annual costs of diabetes management in the individual CVD risk approach were calculated by summing up equations 8a-b and summarized in Table 20 as discounted costs.

$$\Sigma \text{eq}_{8a-b} = \$5575.3312 + \$12268.16 = \$17843.491$$

iii. Statin and Aspirin therapy:

In the individual CVD risk factor approach treatment with statin therapy was recommended for diabetics at high risk for CVD based on risk appraisal that includes measuring lipids and other metrics beyond the scope of this analysis. Anti-platelet therapy was not routinely recommended except for individuals with prior CVD events (not included in this analysis). Therefore, these two therapies were not included in the costs of the approach based on treating individual risk factors (DM/HTN).

*The non-LB absolute CVD risk strategy treatment costs:*

The preventive interventions employed in the non-LB absolute CVD risk strategy were based on the recommendations included in the WHO CVD prevention guidelines discussed in the methods section and outlined in Table 2. Four preventive interventions were costed: antihypertensive therapy, diabetes treatment and monitoring, statin therapy and antiplatelet therapy with aspirin. These expenses are detailed below and summarized as discounted costs in Table 20.

i. Antihypertensive therapy expenses:

Antihypertensive therapy is recommended at different thresholds based on systolic blood pressure and the patient's absolute CVD risk score. For blacks, initial

antihypertensive therapy should include monotherapy with a thiazide-like diuretic or a CCB, which are preferred over ACEIs except in cases where hypertension coexists with diabetes. Antihypertensive therapy for individuals with hypertension and diabetes should include an ACEI combined with a thiazide-like diuretic because blacks have poor response to ACEIs unless combined with a thiazide diuretic (World Health Organization, 2007).

To ensure consistency in this analysis, the costing of the hypertension treatments outlined in Table 2 was done using the commonly prescribed first line drugs' start doses, i.e. Hydrochlorothiazide 25mg/day for hypertensive patients without diabetes, or Lisinopril 10mg/day with Hydrochlorothiazide 25mg/day for hypertensive patients with diabetes. Hydrochlorothiazide NADAC is \$0.01192 per 25mg tablet, while Lisinopril is \$0.02011 per 10mg tablet (Medicaid.gov, 2016). The annual costs for antihypertensive therapy are detailed in equations 9-11 and summarized in Table 20.

For the moderate risk category (10-20%), antihypertensive therapy was indicated for 487 individuals with  $BP \geq 140/90$ , with 80 of them having co-existing diabetes. The annual costs for the antihypertensive therapy in the moderate risk category are estimated in equation 9a-b below:

Equation 9a: Hypertension without comorbid diabetes

$$Rxmod_{htn-dm} = 407 \text{ individuals} * \$0.01192 * 365 \text{ days} = \$1770.76$$

Equation 9b: Hypertension with comorbid diabetes

$$Rxmod_{htn+dm} = 80 \text{ individuals} * (\$0.01192 + \$0.02011) * 365 \text{ days} = \$935.28$$

For the high risk category (20-30%), antihypertensive therapy was indicated for 289 individuals with  $BP \geq 140/90$ , with 96 of them having co-existing diabetes. The

annual costs for antihypertensive therapy in the high risk category/year is estimated in equation 10a-b below:

Equation 10a: Hypertension without comorbid diabetes

$$Rxhigh_{htn-dm} = 193 \text{ individuals} * \$0.01192 * 365 \text{ days} = \$839.70$$

Equation 10b: Hypertension with comorbid diabetes

$$Rxhigh_{htn+dm} = 96 \text{ individuals} * (\$0.01192 + \$0.02011) * 365 \text{ days} = \$1122.33$$

For the very high risk category (>30%), antihypertensive therapy was indicated for 301 individuals with BP $\geq$ 130/80, with 150 of them having co-existing diabetes. The annual cost for treating BP $\geq$ 130/80 for the very high risk category/year is estimated in equation 11a-b below:

Equation 11a: Hypertension without comorbid diabetes

$$Rxvhigh_{htn-dm} = 151 \text{ individuals} * \$0.01192 * 365 \text{ days} = \$656.97$$

Equation 11b: Hypertension with comorbid diabetes

$$Rxvhigh_{htn+dm} = 150 \text{ individuals} * (\$0.01192 + \$0.02011) * 365 \text{ days} = \$1753.64$$

The total annual costs of antihypertensive therapy in the non-LB absolute risk strategy were calculated by summed up costs associated with the three risk categories and shown in Table 20 as discounted costs.

$$\sum eq_{9-11} = \$1770.76 + \$935.28 + \$839.70 + \$1122.33 + \$656.97 + \$1753.64 = \$7078.70$$

ii. Diabetes treatment and monitoring expenses:

The WHO CVD prevention guidelines recommends pharmacotherapeutics treatment for type 2 diabetes to commence at the same threshold (fasting >7mmol/l or >126 mg/dL) for all risk categories (World Health Organization, 2007). The

recommended first line oral hypoglycemic agent is Metformin 850mg/day (also used in the individual risk factors strategy).

In the non-LB absolute CVD risk strategy, 431 individuals with diabetes met the treatment threshold discussed above (absolute risk score  $\geq 10\%$ ). The annual cost for treating type 2 diabetes (RxDM) in the moderate, high and very high absolute CVD risk categories are calculated in equation 12a.

Equation 12a: Cost of treating diabetes with Metformin:

$$\text{RxDM} = 431 \text{ individuals} * \$0.03292 * 365 \text{ days} = \$5178.8098$$

Since the WHO CVD prevention guidelines do not give explicit recommendations on the frequency of blood glucose monitoring for diabetics, in this analysis, monitoring of HbA1C was assumed to occur twice a year during the recommended follow-up visits and as recommended by the ADA. Costing for HbA1C testing was done using CMS clinical diagnostic laboratory fee schedule for CPT code 83036 (also used in the individual risk factors strategy). The annual cost for glucose monitoring (HbA1Ct) in the three absolute CVD risk categories/year are calculated in equation 12b.

Equation 12b: Cost of monitoring diabetes with HbA1C test:

$$\text{Hb}_{\text{A1Ct}} = 431 \text{ individuals} * \$13.22 * 2 \text{ tests/year} = \$11395.64$$

The total cost of diabetes management in the three absolute CVD risk categories were calculated by summing up equations 12a-b and summarized in Table 20 as discounted costs.

$$\sum \text{eq12}_{a-b} = \$5178.8098 + \$11395.64 = \$16574.45$$

iii. Statin therapy expenses:

Treatment with statin therapy was only recommended for the very high risk category (>30%) when there was no mechanism to assess lipid levels (see Table 2). To ensure consistency in this analysis, costing was done using the common prescribed first line drug start dose, i.e. atorvastatin 10mg/day (NDC 00378395005). Atorvastatin NADAC is \$0.10714 per 10mg tablet (Medicaid.gov, 2016).

The non-LB absolute CVD risk strategy put 311 individuals in the very high risk category, hence qualifying them for statin therapy. The annual cost of the statin therapy (Rxstatin) in the very high risk category/year was estimated in equation 13 and summarized in Table 20 as discounted costs.

Equation 13: Cost of statin therapy:

$$\text{Rxstatin} = 311 \text{ individuals} * \$0.10714 * 365 \text{ days} = \$12162$$

iv. Aspirin therapy expenses:

Treatment with aspirin therapy was only recommended for the very high risk category (>30) as outlined in Table 2. To ensure consistency in this analysis, costing was done using the common first line drug start dose, i.e. enteric coated aspirin 81mg/day (NDC 00536100410). Aspirin NADAC is \$0.01117 per 81mg tablet (Medicaid.gov, 2016).

The non-LB absolute CVD risk strategy put 311 individuals in the very high risk category thus qualifying them for antiplatelet therapy. The annual cost of the Aspirin therapy in the very high risk category/year was estimated in equation 14 and shown in Table 20 as discounted costs.

Equation 13: Cost of Aspirin therapy:

$$R_{xasprin} = 311 \text{ individuals} * \$0.01117 * 365 \text{ days} = \$1,267.9626$$

*The non-LB absolute CVD risk strategy treatment costs:*

The preventive interventions employed in the absolute CVD risk strategy guided by the LB Framingham algorithm were based on the absolute CVD risk score and followed the same pattern described under the non-LB absolute CVD risk strategy.

i. Antihypertensive therapy expenses

Antihypertensive therapy was costed using the same recommendations and thresholds described under the preventive strategy guided by the non-LB Framingham algorithm. These expenses were calculated in equations 14-16 below and summarized in Table 20 as discounted costs.

For the moderate risk category (10-20%), antihypertensive therapy was indicated for 479 individuals with  $BP \geq 140/90$ , with 98 of them having co-existing diabetes. The annual costs for the antihypertensive therapy in the moderate risk category/year are estimated in equation 14a-b.

Equation 14a: Hypertension without comorbid diabetes

$$R_{xmodl_{htn-dm}} = 381 \text{ individuals} * \$0.01192 * 365 \text{ days} = \$1657.6548$$

Equation 14b: Hypertension with comorbid diabetes

$$R_{xmodl_{htn+dm}} = 98 \text{ individuals} * (\$0.01192 + \$0.02011) * 365 \text{ days} = \$1145.7131$$

For the high risk category (20-30%), antihypertensive therapy was indicated for 252 individuals with  $BP \geq 140/90$ , with 89 of them having co-existing diabetes. The annual costs for antihypertensive therapy in the high risk category/year are estimated in equation 15a-b.

Equation 15a: Hypertension without comorbid diabetes

$$Rxhighl_{htn-dm} = 163 \text{ individuals} * \$0.01192 * 365 \text{ days} = \$709.1804$$

Equation 15b: Hypertension with comorbid diabetes

$$Rxhighl_{htn+dm} = 89 \text{ individuals} * (\$0.01192 + \$0.02011) * 365 \text{ days} = \$1040.495$$

For the very high risk category (>30%), antihypertensive therapy was indicated for 200 individuals with BP $\geq$ 130/80, with 108 of them having co-existing diabetes. The annual cost for treating BP $\geq$ 130/80 for the very high risk category/year are estimated in equation 16a-b.

Equation 16a: Hypertension without comorbid diabetes

$$Rxvhighl_{htn-dm} = 92 \text{ individuals} * \$0.01192 * 365 \text{ days} = \$400.2736$$

Equation 16b: Hypertension with comorbid diabetes

$$Rxvhighl_{htn+dm} = 108 \text{ individuals} * (\$0.01192 + \$0.02011) * 365 \text{ days} = \$1262.622$$

The total annual cost of antihypertensive therapy in all the categories above were calculated by summing equations 14-16 and summarized in Table 20 as discounted costs.

$$\sum eq_{14-16} = \$1657.65 + \$1145.71 + \$709.18 + \$1040.5 + \$1040.5 + \$400.27 + \$1262.62 = \$7256.43$$

ii. Diabetes treatment and monitoring expenses

In the LB absolute CVD risk strategy, 399 individuals with diabetes met the treatment threshold (absolute risk score  $\geq$ 10%). The annual costs for treating and monitoring type 2 diabetes (RxDM<sub>I</sub>) in the moderate, high and very high risk categories are calculated in equations 17a-b.

Equation 17a: Cost of treating diabetes with Metformin:

$$RxDM_I = 399 \text{ individuals} * \$0.03292 * 365 \text{ days} = \$4,794.30$$

Equation 17b: Cost of monitoring diabetes with HbA1C test:

$$Hb_{A1CtI} = 399 \text{ individuals} * \$13.22 * 2 \text{ tests/yr} = \$10,549.56$$

The total cost of diabetes management in the three risk categories were calculated by summing up equations 17a-b and summarized as discounted costs in Table 20.

$$\sum \text{eq17}_{a-b} = \$4,794.30 + \$10,549.56 = \$15,343.86$$

iii. Statin therapy expenses

Treatment with statin therapy was recommended if total cholesterol was greater than 8 mmol/l (309 mg/dL) for the moderate risk category, and greater than 5mmol/l (193 mg/dL) for the high risk category. In addition, everyone in the very high risk category (>30%) was put on statin therapy irrespective of total cholesterol levels (see Table 2).

Based on these criteria, the LB absolute CVD risk strategy identified 25 individuals in the moderate risk category, 250 individuals in the high risk category and 234 individuals in the very high risk category as qualifying for statin therapy. Therefore, a total of 509 individuals were put on statin therapy. The associated annual cost for statin therapy are calculated in equation 18 and included as discounted costs in Table 20.

Equation 18: Cost of statin therapy

$$\text{Rxstatin}_l = 509 \text{ individuals} * \$0.10714 * 365 \text{ days} = \$19,905$$

iv. Aspirin therapy expenses

Treatment with aspirin therapy was only recommended for the very high risk category (see Table 2). The LB absolute CVD risk strategy put 234 individuals in the very high risk category thus qualifying them for antiplatelet therapy. The annual cost of the Aspirin therapy in the very high risk category was calculated in equation 19 and summarized as discounted costs in Table 20.

Equation 19: Cost of Aspirin therapy

$$\text{Rxasiprin} = 234 \text{ individuals} * (\$0.01117 * 365 \text{ days}) = \$954.0$$



Table 20: Discounted costs of preventive interventions: Individual risk factors vs non-LB vs LB absolute CVD risk programs

Year	Individual CVD risk factor program			Non-LB Absolute CVD risk program					LB Absolute CVD risk program				
	BP therapy	DM therapy	Discounted cost /year	BP therapy	DM therapy	Statin therapy	Aspirin therapy	Discounted cost/year	BP therapy	DM therapy	Statin therapy	Aspirin therapy	Discounted cost/year
Year 1	\$7,259	\$17,324	\$24,583	\$6,873	\$16,092	\$11,808	\$1,231	\$36,003	\$7,045	\$14,897	\$19,325	\$926	\$42,194
Year 2	\$7,048	\$16,819	\$23,867	\$6,672	\$15,623	\$11,464	\$1,195	\$34,954	\$6,840	\$14,463	\$18,762	\$899	\$40,965
Year 3	\$6,842	\$16,329	\$23,172	\$6,478	\$15,168	\$11,130	\$1,160	\$33,936	\$6,641	\$14,042	\$18,216	\$873	\$39,771
Year 4	\$6,643	\$15,854	\$22,497	\$6,289	\$14,726	\$10,806	\$1,127	\$32,948	\$6,447	\$13,633	\$17,685	\$848	\$38,613
Year 5	\$6,450	\$15,392	\$21,842	\$6,106	\$14,297	\$10,491	\$1,094	\$31,988	\$6,259	\$13,236	\$17,170	\$823	\$37,488
Year 6	\$6,262	\$14,944	\$21,205	\$5,928	\$13,881	\$10,185	\$1,062	\$31,057	\$6,077	\$12,850	\$16,670	\$799	\$36,397
Year 7	\$6,079	\$14,508	\$20,588	\$5,756	\$13,477	\$9,889	\$1,031	\$30,152	\$5,900	\$12,476	\$16,185	\$776	\$35,336
Year 8	\$5,902	\$14,086	\$19,988	\$5,588	\$13,084	\$9,601	\$1,001	\$29,274	\$5,728	\$12,113	\$15,713	\$753	\$34,307
Year 9	\$5,730	\$13,676	\$19,406	\$5,425	\$12,703	\$9,321	\$972	\$28,421	\$5,561	\$11,760	\$15,256	\$731	\$33,308
Year 10	\$5,564	\$13,277	\$18,841	\$5,267	\$12,333	\$9,050	\$943	\$27,593	\$5,399	\$11,417	\$14,811	\$710	\$32,338
Year 11	\$5,402	\$12,891	\$18,292	\$5,114	\$11,974	\$8,786	\$916	\$26,790	\$5,242	\$11,085	\$14,380	\$689	\$31,396
Year 12	\$5,244	\$12,515	\$17,759	\$4,965	\$11,625	\$8,530	\$889	\$26,009	\$5,090	\$10,762	\$13,961	\$669	\$30,481
Total	\$74,426	\$177,614	<b>\$252,040</b>	\$70,461	\$164,982	\$121,061	\$12,621	<b>\$369,125</b>	\$72,231	\$152,733	\$198,134	\$9,496	<b>\$432,594</b>

BP, DM, Statin & Aspirin therapies are the discounted annual costs of therapy for hypertension, diabetes, cholesterol and platelets respectively, calculated by multiplying the at-risk ARIC sample that was eligible for therapy with the NADAC of the drug (s) recommended by each CVD prevention program  
Discounted cost/year sums up the discounted costs of all therapies prescribed to the at-risk ARIC sample annually

### *Summary of level I and II expenses*

The total costs of screening and preventive interventions prescribed for the true and false positives in the individual CVD risk factor program ( $CRx_{rjt}$ ) were calculated by summing up the costs of screening (eq. 1a), initial (eq. 2a) and follow-up (Table 19) office visits, and cost of pharmacotherapeutics (Table 20) as detailed below:

$$CRx_{rjt} = \$14,634 + \$156,135.96 + \$3,108,356 + \$252,040 = \$3,531,165.96$$

The total costs of screening and preventive interventions prescribed for the true and false positives in the non-LB absolute CVD risk program ( $CRx_{ajt}$ ) were calculated by summing up the costs of screening (eq. 1b), initial (eq. 2b) and follow-up (Table 19) office visits, and cost of pharmacotherapeutics (Table 20) as detailed below:

$$CRx_{ajt} = \$20,686 + \$152,826.72 + \$2,219,969 + \$369,125 = \$2,762,606.72$$

The total costs of screening and preventive interventions prescribed for the true and false positives in the LB absolute CVD risk program ( $CRx_{ljt}$ ) were calculated by summing up the costs of screening (eq. 1c), initial (eq. 2c) and follow-up (Table 19) office visits, and cost of pharmacotherapeutics (Table 20) as detailed below:

$$CRx_{ljt} = \$66,631.3 + \$130,965.68 + \$1,864,614 + \$432,594 = \$2,494,804.98$$

### *Level III: Expenses in treating false negatives cases*

The cost of secondary prevention interventions associated with each CVD prevention strategy was dependent on the sensitivity of the screening algorithm used. A screening algorithm with low sensitivity led to a high number of false negatives requiring treatment and rehabilitation services for the CVD events occurring in the group. Since the initial follow-up in ARIC did not include the survival status of each incident CVD event,

costing level III expenses was done using the average costs associated with CVD events in general rather than the average cost for specific CVD events.

Although the costs associated with treating the CVD events included in this study vary greatly within the US, this analysis used the estimated direct average initial and follow-up costs for treating CVD events in the US published by Chapman and colleagues in 2011 (Chapman, Liu, Girase, & Straka, 2011). In their retrospective matched cohort analysis of commercially insured managed care population, Chapman et al. (2011) estimated that initial inpatient management of a CVD event would have an average cost of \$16,981 (SD \$20,474), while the first year follow-up costs would average \$16,582 (SD \$34,425) per case.

Due to data limitations, this analysis estimated level III expenses using these average direct medical and pharmacological expenses published by Chapman et al. (2011) despite the expected great variation in event specific costs. Although Chapman et al. (2011) demonstrated that the follow-up costs would increase in subsequent years of follow-up, this analysis assumed a constant yearly follow-up cost of \$16,582 per incident CVD. The initial treatment costs and follow-up expenses of the false negatives associated with the non-LB and individual CVD risk factor approaches were depended on the year in which each event occurred.

To calculate level III expenses (CUSE), the false negative cases in each year were multiplied by the discounted average cost of initial management of a CVD event (i.e.  $\Sigma \text{cost} = \sum_{t=1}^{12} [\text{False negative cases} * \$16,981 / 1.03^t]$ ) and then the discounted average yearly follow-up costs were added for every subsequent follow-up year, through year 12 (i.e.  $\Sigma = \sum_{t=1}^{12} [\text{False negative cases} * \$16,582 / 1.03^t]$ ). For instance, the non-LB absolute CVD

risk approach was associated with 3 false negatives in year one. The discounted initial costs for year one was:  $3 \text{ individuals} * \$16,981/1.03^1 = \$49,459.22$ . The follow-up costs for these events was the sum of the discounted follow-up costs for each of the subsequent years of follow-up (see Table 21).

Table 21: Initial & follow-up costs for treating false negative cases by prevention strategy

	Individual CVD risk factors program		non-LB absolute CVD risk program		LB absolute CVD risk program	
Year	Discounted initial costs	Discounted f/up costs	Discounted initial costs	Discounted f/up costs	Discounted initial costs	Discounted f/up costs
Yr 1	\$65,946	\$660,229	\$49,459	\$495,172	\$49,459	\$495,172
Yr 2	\$80,031	\$744,791	\$32,012	\$297,917	\$48,019	\$446,875
Yr 3	\$46,620	\$399,984	\$62,160	\$533,313	\$77,700	\$666,641
Yr 4	\$75,437	\$590,766	\$75,437	\$590,766	\$60,350	\$472,613
Yr 5	\$117,184	\$827,363	\$58,592	\$413,681	\$87,888	\$620,522
Yr 6	\$85,328	\$534,700	\$42,664	\$267,350	\$71,107	\$445,583
Yr 7	\$27,614	\$150,459	\$69,036	\$376,147	\$69,036	\$376,147
Yr 8	\$160,859	\$740,961	\$160,859	\$740,961	\$174,264	\$802,708
Yr 9	\$91,102	\$340,597	\$65,073	\$243,284	\$104,116	\$389,254
Yr 10	\$138,990	\$395,428	\$88,448	\$251,636	\$113,719	\$323,532
Yr 11	\$159,477	\$306,923	\$147,209	\$283,314	\$147,209	\$283,314
Yr 12	\$142,922	\$139,563	\$119,101	\$116,303	\$226,293	\$220,975
Total	\$1,191,509	\$5,831,766	\$970,051	\$4,609,843	\$1,229,159	\$5,543,336

Discounted initial costs have been calculated by multiplying false negative cases with the discounted average cost of initial management of CVD ( $\$16,981/1.03^1$ ) published by Chapman et al. (2011).

Discounted follow-up (f/up) costs have been calculated by multiplying false negative cases each year with the discounted annual follow-up cost in subsequent years using the average cost of the first year of CVD follow-up ( $\$16,582/1.03^1$ ) published by Chapman et al. (2011).

False negatives are calculated from ARIC data based on the sensitivity/specificity of the screening method used described in Tables 17-19.

### *Summary of level III expenses*

The total costs of treating false negative cases in the individual CVD risk factors program ( $CUSE_{rjt}$ ) were calculated by summing up the initial and follow-up costs in Table 21.

$$CUSE_{rjt} = \$1,191,509 + \$5,831,766 = \$7,023,275$$

The total costs of treating false negative cases in the non-LB absolute CVD risk program ( $CUSE_{ajt}$ ) were calculated by summing up the initial and follow-up costs in Table 21.

$$CUSE_{ajt} = \$970,051 + \$4,609,843 = \$5,579,894$$

The total costs of treating false negative cases in the LB absolute CVD risk program ( $CUSE_{ljt}$ ) were calculated by summing up the initial and follow-up costs in Table 21.

$$CUSE_{ljt} = \$1,229,159 + \$5,543,336 = \$6,772,495$$

### *Outcomes associated with the three prevention programs*

Ideally, the effectiveness of a disease prevention program depends on the accuracy of the screening method used, and the potency of the prescribed preventive interventions. However, due to data limitations, this analysis assumed that the preventive interventions prescribed under the individual risk factors and absolute CVD risk approaches to CVD prevention were equally potent in preventing the true CVD cases identified by the respective screening methods. Therefore, the number of prevented CVD events for each preventive strategy were dependent on the sensitivity/specificity of its screening algorithm at its optimal risk threshold.

In the context of this analysis, sensitivity denotes the proportion of the observed (true) CVD cases that were correctly identified as positive (high risk) by a screening algorithm at baseline. On the other hand, specificity is the proportion of true CVD free cases that were correctly identified as negative (low risk) by a screening algorithm at baseline.

The individual risk factors approach focusing on screening for diabetes and/or hypertension had its optimal balance of sensitivity/specificity at the high risk (presence of diabetes and/or hypertension) threshold where sensitivity is 78.05% and specificity 45.65% as detailed in Table 16. This implies that when individuals with diabetes and/or hypertension are considered as positive cases, about 78% of all individuals who would end up experiencing CVD events during follow-up were identified at baseline. At the same time, about 56% if individuals who did not end up developing CVD events are misclassified as positive and hence unnecessarily put on preventive interventions.

The non-LB Framingham algorithm had its optimal balance of sensitivity/specificity at the moderate risk (10-20%) threshold where sensitivity was 82.04% and specificity 47.79% as detailed in Table 17. This implies that when individuals with  $\geq 10\%$  absolute CVD risk score were considered as positive cases, about 82% of all individuals who would end up experiencing CVD events during follow-up were correctly identified at baseline. At the same time, about 52% if individuals who did not end up developing CVD events were misclassified as positive cases.

The LB Framingham algorithm had its optimal balance of sensitivity/specificity at the moderate risk (10-20%) threshold where sensitivity was 77.06% and specificity 56.44% as detailed in Table 19. This implies that when individuals with  $\geq 10\%$  absolute

CVD risk score were considered as positive cases, about 77% of all individuals who would end up experiencing CVD events during follow-up were correctly identified at baseline. At the same time, almost 44% if individuals who did not end up developing CVD events were misclassified as positive cases.

True positive cases that occurred later during follow-up were weighted less than those occurring early in the follow-up using the annual discounting rate of 3% discussed under the analysis framework. Table 22 summarizes the discounted true positives (predicted true CVD events) for each program.

Table 22: True positive cases stratified by CVD prevention program

			Individual CVD risk factors program		non-LB absolute CVD risk program		LB absolute CVD risk program	
	Observed CVD events		Predicted true CVD events		Predicted true CVD events		Predicted true CVD events	
Year	actual	discounted	actual	discounted	actual	discounted	actual	discounted
Yr 1	20	19	16	16	17	17	17	17
Yr 2	23	22	18	17	21	20	20	19
Yr 3	30	27	27	25	26	24	25	23
Yr 4	31	28	26	23	26	23	27	24
Yr 5	33	28	25	22	29	25	27	23
Yr 6	33	28	27	23	30	25	28	23
Yr 7	35	28	33	27	30	24	30	24
Yr 8	48	38	36	28	36	28	35	28
Yr 9	36	28	29	22	31	24	28	21
Yr 10	32	24	21	16	25	19	23	17
Yr 11	37	27	24	17	25	18	25	18
Yr 12	43	30	31	22	33	23	24	17
Total	401	327	313	257	329	270	309	254

Observed events are the CVD events that occurred each year in the at risk ARIC cohort

Predicted true CVD events are the true positive cases calculated from ARIC data based on the sensitivity/specificity of the screening method used

Discounted cases are the true positive cases multiplied by annual discounting rate  $(1.03)^{-t}$  where  $t$  = year 1 through 12

### Average and incremental cost effectiveness

#### *The average cost-effectiveness ratio*

The average cost-effectiveness of each CVD prevention approach is a function of the net costs divided by net benefits associated with each program. These equations are described in equations 22-24.

Equation 22: The individual CVD risk factors program.

$$\frac{\Delta TC_r}{\Delta E_r} = \frac{\Sigma[\$3,531,166 + \$7,023,275]}{[257]} = \frac{\$10,554,441}{257 \text{ true positives}} = \frac{\$41,068}{1 \text{ true positive}}$$

Equation 23: The non-LB absolute CVD risk program.

$$\frac{\Delta TC_a}{\Delta E_a} = \frac{\Sigma[\$2,762,607 + \$5,579,894]}{[270]} = \frac{\$8,342,501}{270 \text{ true positives}} = \frac{\$30,898}{1 \text{ true positive}}$$

Equation 24: The LB absolute CVD risk program.

$$\frac{\Delta TC_l}{\Delta E_l} = \frac{\Sigma[\$2,494,805 + \$6,772,495]}{[254]} = \frac{\$9,267,300}{254 \text{ true positives}} = \frac{\$36,485}{1 \text{ true positive}}$$

Overall, the individual risk factors and LB absolute CVD risk programs had their cost-effectiveness ratios higher by 25% and 14% respectively compared to the non-LB absolute CVD risk program. Compared to the non-LB absolute CVD risk program, the 12-year discounted costs were 21% and 9% greater in the individual risk factors and LB absolute CVD risk programs respectively. Both programs identified 5% and 6% fewer cases respectively compared to the non-LB absolute CVD risk program.



*The incremental cost-effectiveness ratio*

The incremental cost-effectiveness ratio was calculated by populating the ICER model with the costs and outcomes discussed in the above sections. For each CVD prevention strategy, the costs were calculated by adding the costs of screening, initial visit, discounted follow-up visits, and the discounted cost of preventive treatments prescribed (CRx), and the cost of treating false negatives (CUSE). The outcomes for each CVD prevention strategy were calculated as the discounted true positive cases and constitutes CVD events which could be prevented through early detection of risk depending on the screening algorithm used.

The incremental cost effectiveness ratio (ICER) of the non-LB absolute CVD risk approach versus the individual CVD risk factor approach is calculated in equation 25.

Equation 25: ICER=

$$\frac{\Delta TC_{a-r}}{\Delta E_{a-r}} = \frac{\sum_{j=1}^2 [\$2,762,607 + \$5,579,894 - \$3,531,166 - \$7,023,275]}{[270 - 257]} = \frac{-\$2,211,9401}{+13 \text{ true positives cases}}$$

Interpretation: The non-LB absolute CVD risk approach would cost \$2 million less over 12 years to identify 13 more actual CVD cases than the individual risk factors approach. For every extra case that the non-LB approach identifies, it saves \$170,000. Hence, the non-LB approach completely dominates the individual risk factors approach in both costs and predictive ability.

The incremental cost effectiveness ratio (ICER) of the non-LB absolute CVD risk approach versus the LB absolute CVD risk approach is calculated in equation 26.

Equation 26: ICER=

$$\frac{\Delta TC_{a-l}}{\Delta E_{a-l}} = \frac{\sum_{j=1}^2 [\$2,762,607 + \$5,579,894 - \$2,494,805 - \$6,772,495]}{[270 - 254]} = \frac{-\$924,799}{+16 \text{ true positive cases}}$$

Interpretation: The non-LB absolute CVD risk approach would cost \$900,000 less over 12 years to identify 16 more actual CVD cases than the LB absolute CVD risk. For every extra case that the non-LB approach identifies, it saves \$58,000. Hence, the non-LB approach completely dominates the LB approach in both costs and predictive ability.

## CHAPTER 5

### DISCUSSION

The overarching goal of this study was to externally validate and determine the cost-effectiveness of the non-LB Framingham algorithm in the multi-racial ARIC dataset. Validating and costing the non-LB Framingham algorithm in a multiracial sample with high representation of individuals who self-report black race was postulated to be an important step in availing a risk assessment tool that could guide CVD prevention in resource constrained settings. Important results discussed in this chapter include; the influence of social determinants of health on CVD risk assessment, and comparative predictive performance and cost-effectiveness of non-LB Framingham Algorithm.

#### Social determinants of health and CVD risk assessment

The conditions in which individuals are born, grow, live, work and age are known to play an important role in the evolution of many diseases including CVD (Will et al., 2011). These social determinants of health are particularly influential in attenuating or exacerbating manifestation of the modifiable CVD risk factors included in this study. Risk assessment algorithms, such as the non-LB Framingham, that incorporate risk factors which are shaped by social determinants of health require an evaluation of relevance and validity before they are generalized across populations.

The organizing framework used in this study appreciates the modulating impact of the social determinants of health on CVD risk and provides a basis for examining the

comparability of effects of CVD risk factors in the white versus black US population. Since the social determinants of health are in part driven by distribution of resources, the framework also sets the stage for evaluating the feasibility of various CVD risk assessment strategies in resource constrained settings.

#### Incidence of CVD by sex and race

The statistical differences between the eligible and ineligible samples are expected in this kind of study. Since over 81% of the ineligible sample was excluded due to prevalent CVD at baseline, it is expected that CVD risk factors in this group was significantly higher than the eligible sample which was free of CVD at baseline. The higher percentage of blacks in the ineligible sample (36%) compared to the eligible sample (23%) is consistent with the known higher burden of CVD in this population.

In the entire ARIC cohort, the incidence rate of CVD was lower but comparable to the Framingham cohort (11.1 versus 11.5 cases per 1000 person-years respectively). white women in the ARIC dataset had a significantly lower incidence rate compared to their counterparts in the Framingham cohort (6.5 versus 8.4 cases per 1000 person-years respectively). On the other hand, white men in the ARIC cohort had a higher incidence rate compared to men in the Framingham cohort (15.5 versus 15.1 cases per 1000 person-years respectively).

The factors associated with the disparate incidence rates between the two cohorts could be multifaceted. One potential factor may be the manner in which the CVD variable is operationalized in the two studies. In the Framingham dataset the CVD variable included CHD, heart failure, stroke and peripheral vascular disease, however the latter variable was not included in the ARIC incidence data provided by NHLBI. As a

result, the CVD variable in ARIC does not capture incident peripheral vascular disease. If the incidence rate of peripheral vascular disease in the Framingham dataset (1.2 and 2.2 cases per 1000 person-years in women and men respectively) were to be applied in the ARIC cohort, the white women's CVD incidence rate would increase slightly and be comparable to the rate in Framingham cohort, while white men would have a significantly higher CVD incident rate.

The different incident rates among white men in ARIC versus Framingham may also be as a result of variation in sample characteristics. Since ethnicity was not reported in the ARIC dataset, there may be unaccounted differences in CVD incidence by ethnicity. Correspondence with NHLBI clarified that Hispanics in the ARIC sample were coded as black or white depending on their self-reported racial group.

The similarity between the incidence proportion of CVD among white and black men in ARIC (18.6% and 18.8% respectively,  $p=0.900$ ) is atypical since black men have been reported to have a higher incidence of CVD (R. Cooper et al., 2000; Mozaffarian et al., 2015). If there were a significant number of Hispanic individuals in the ARIC cohort, this could have a negative effect on the discrimination and calibration since the Framingham algorithms are generally known to perform poorly among Hispanics (Beswick et al., 2008).

#### Performance of the non-LB Framingham CVD risk assessment algorithm

After confirming that all covariates included in the non-LB and LB Framingham models met the proportionality of hazard assumption, and there was no significant multicollinearity between variables, sex specific Cox regression was used to test the mathematical performance of the Framingham models in the ARIC dataset. Use of sex

specific models was critical because CVD risk factors are known to have different effect sizes among women and men. Just as reported in the Framingham dataset, the unadjusted non-LB and LB Framingham models were comparable in discrimination and calibration, and their performance was superior in women compared to men in the ARIC dataset.

Among women, both the non-LB and LB Framingham models performed well, but among men, discrimination was low and calibration was poor for both models. The comparable performance of the non-LB and LB Framingham models in both sexes suggest that BMI could be an adequate proxy for HDL and total cholesterol in both sexes. The similar performance of both the non-LB and LB Framingham models in the white and black cohorts collaborates previous studies which have reported optimal performance of Framingham algorithms in the US black population.

The discrimination of risk among women in ARIC was within the confidence interval reported in the Framingham dataset for the published non-LB (95% CI, 0.76-0.81) and LB Framingham models (95% CI, 0.77-0.81). However, risk discrimination among men was significantly lower. The underperformance of the models among men in the ARIC dataset suggest that the independent variables used to predict CVD risk in the Framingham dataset do not capture the full extent of risk among men in the ARIC dataset. This phenomenon necessitates further analysis to examine the effect of the Framingham derived risk factors in the ARIC cohort, and additional or alternative variables which may improve the performance of the non-LB Framingham models.

Evaluation of the effect and impact of the Framingham derived risk factors in the ARIC cohort was done by comparing the regression coefficients from the models generated in the ARIC dataset to those reported from the Framingham dataset using the z

score formulae. Overall, all the covariates in the non-LB and LB models had similar effect and impact in the black cohort, but smoking and diabetes had a different effect in the white cohort. The impact of smoking was higher among white women but lower among white men, while diabetes had a higher impact among white men in ARIC compared to the Framingham dataset. These differences pinpoint potential causes of the low performance of the non-LB and LB Framingham models among men. Since white men constituted 79% of the male sample in the ARIC cohort, the comparative low impact of smoking on CVD risk in this group may have affected the overall performance of the models among men. The differences also suggest that additional or alternative variable(s) may have a stronger explanatory power on CVD risk compared to smoking.

The equivalence of effect of the Framingham generated risk factors in ARIC's black cohort suggests that the general CVD Framingham algorithms are applicable to the US black population. This applicability mirrors what has been reported with earlier versions of Framingham algorithms which focused on hard coronary events and were validated in black datasets (D'Agostino RB, Grundy, Sullivan, Wilson, & CHD Risk Prediction Group, 2001a). The semblance of effect of the traditional CVD risk factors in the black ARIC dataset and the white Framingham dataset also supports the widely accepted premise that traditional CVD risk factors have a fairly similar effect and impact across populations (Yusuf et al., 2004). These results contribute to the body of knowledge pertaining to CVD epidemiology in the black dataset and provide an evidence based foundation upon which research on novel risk factors hypothesized to have a unique impact in the black population could be added and tested.

When more CVD risk factors were added to the non-LB and LB Framingham models, there was no significant improvement in discrimination or calibration. In fact, adding more risk factors to the non-LB model (family history of premature CHD and waist hip ratio) and the LB model (BMI, waist hip ratio, family history of premature CHD, and apolipoprotein A and B) tended to increase the confidence interval of their C statistic without meaningful improvement in discrimination or calibration. The lack of significance suggests that the impact of these additional risk factors is mediated by the traditional risk factors already included in the published Framingham models.

It is worth noting that some Framingham risk factors lost their statistical significance when additional risk factors were included. For instance, waist hip ratio replaced BMI as the significant variable in the adjusted non-LB model for all cohorts except among white women. Similarly, waist hip ratio replaced total cholesterol as the significant variable in the adjusted LB model among black women and men, and was significant among white men without affecting the significance of total cholesterol.

The differential effect of body mass across populations has been reported previously. For instance, the association between high BMI and CVD mortality has been reported as stronger in white women than in black women (Abell et al., 2007). These racial differences in the effect of BMI were apparent in this analysis. Among the non-LB Framingham risk factors, BMI had a marginally lower effect on black women and men in the ARIC ( $\beta=0.668$  and  $0.631$  respectively) compared to white women and men ( $\beta=0.797$  and  $0.793$  respectively) in the ARIC dataset.

Stepwise regression analysis revealed that other covariates in the non-LB model (diabetes, systolic blood pressure, antihypertensive therapy, smoking and age) had



adequate discrimination among black women ( $C=0.7426$ ) and BMI added no significant improvement ( $C=0.7452$ ) to the CVD prediction model. Waist hip ratio had a slightly better effect than BMI among black women ( $C=0.7481$ ). The waning effect of BMI when other traditional risk factors are held constant is consistent with the findings reported by Abell and colleagues. The study demonstrated that the association between obesity and CVD mortality was no longer significant among black women when hypertension, total cholesterol, diabetes, age and smoking status were controlled (Abell et al., 2007).

These findings add to the body of literature suggesting that BMI may not be an optimal CVD risk indicator, and its effect could be mediated in part by other related risk factors such as hypertension and diabetes. Although BMI continues to be widely used as the metric for diagnosing overweight and obesity, there is a growing body of literature describing its limitations. Whereas BMI is an indicator total body fat, the metric does not take into consideration how the fat is distributed within the body (Simon, 2009). From a cardiovascular standpoint, abdominal fat is more dangerous than any other fat in the body.

The slightly better performance of the waist hip ratio in the non-LB and LB adjusted models also adds to the body of evidence suggesting that measures of central adiposity may be more relevant in predicting CVD compared to measures of body mass. The strong predictive power of central obesity has been reported by many studies including the INTERHEART study where investigators demonstrated a strong link between increased waist hip ratio and the risk of heart attack even after controlling for the traditional CVD risk factors. In the same study, BMI lost its modest association with myocardial infarction after adjusting for traditional risk factors (Yusuf et al., 2005).

Despite the growing evidence supporting the significant role of central adiposity in predicting CVD, most of the existing non-LB algorithms have not included it as a covariate. Possible reasons for this may include the reported difficulties associated with assessing and reproducing the waist hip ratio measure (Simon, 2009). Over the last decade, waist circumference has been suggested as an alternative measure of central adiposity that still has a strong link to the risk of myocardial infarction, but is relatively easier to reproduce (Simon, 2009; Yusuf et al., 2005).

The strong effect of central adiposity on CVD risk in the black population, and the superiority of waist hip ratio and circumference in predicting CVD provide insights on potential pathways to improve the non-LB Framingham algorithm. Replacing BMI with a measure of central adiposity such as waist circumference and testing the model in a large homogeneous black cohort could provide important data on the additional value of the measure in predicting absolute CVD risk. Other measures which could be helpful include the ankle brachial index which is used to diagnose peripheral vascular disease. Testing this measure in a homogeneous black cohort with peripheral vascular disease included in the CVD variable may provide insights as to whether it has any extra value in optimizing risk prediction of the non-LB Framingham algorithm. Evaluating the effect and impact of ankle branchial index was not possible in this study since peripheral vascular disease was not included in the CVD variable.

#### *Comparative performance in risk stratification*

The comparable performance of the published non-LB and LB Framingham algorithms in actual risk stratification of the ARIC dataset complements the similarities observed in the mathematical performance. The high overall agreement (92.76%) and

substantial kappa statistic (0.76) suggest that the two algorithms are very comparable in stratification of risk. Stratification of the cohort in the high and very high risk categories was essentially the same for both algorithms, but the LB algorithm placed slightly more individuals in the low risk category. Both algorithms placed a greater proportion of men and blacks in higher risk categories, a trend which is consistent with current published CVD epidemiology (Mozaffarian et al., 2015).

The agreement between the non-LB and LB versions of the recalibrated and the adjusted algorithms derived in ARIC were also high. However, the recalibrated and adjusted algorithms placed a significantly greater percentage of the ARIC population in the low risk category compared to the published Framingham non-LB algorithm. This discordance in risk stratification necessitated sensitivity and specificity analysis to determine the clinical usefulness of the non-LB Framingham algorithm, and its comparison with the alternative algorithms discussed above.

The comparability of the non-LB and LB Framingham algorithm AUROC curves (0.706 vs 0.71 respectively) further adds to the evidence that HDL and total cholesterol may not add significant marginal value to CVD risk prediction especially in the black population. Since AUROC curves usually depict the percentage of randomly selected pairs for which the test correctly classifies as normal or abnormal, the AUROC of the non-LB algorithm manifest better performance in the ARIC dataset compared to other alternatives. The higher AUROC of the published non-LB Framingham algorithm compared to those of the recalibrated and adjusted algorithms suggest that the non-LB Framingham algorithm could be ready for use ‘as is’ in the US black population.

The suboptimal performance of the calibrated non-LB (AUROC=0.67) and LB (AUROC=0.69) based algorithms indicate that ARIC derived survival and mean of risk factors are not better substitutes for those generated in Framingham. This could be as a result of the unmeasured confounders discussed earlier, which may also have contributed to the low performance of the adjusted non-LB (AUROC= 0.68) and LB (AUROC= 0.69) based models. It is important to note that inclusion of up to 9 covariates in the adjusted algorithm did not make much difference in its predictive ability.

Sensitivity/specificity analysis also revealed that the non-LB Framingham algorithm had a slightly better sensitivity but poorer specificity compared to the LB algorithm. For instance, at the high risk threshold (20-30%), the non-LB Framingham algorithm had sensitivity/specificity ratios of 0.25/0.95 for women and 0.51/0.73 for men. The LB Framingham algorithm sensitivity/specificity ratios were 0.23/0.95 for women and 0.50/0.75 for men respectively.

In the Framingham dataset, the non-LB algorithm was both slightly less sensitive and specific compared to the LB algorithm. For example, at 20% risk threshold, sensitivity/specificity ratios were 0.58/0.83 for women and 0.48/85 for men versus 0.60/0.84 for women and 0.49/85 for men in the non-LB and LB algorithms respectfully. (D'Agostino RB et al., 2008). These ratios indicate that in the ARIC dataset, preventive interventions should be initiated at least at the moderate risk category because more individuals who will end up developing events (true positives) was misclassified as not at-risk (false negatives) if the high risk category was adopted as the treatment threshold.

Whereas the tradeoff between sensitivity and specificity are often delicate, a non-LB algorithm with slightly higher sensitivity could be very important in helping early

detection of risk in the black population that carries the highest burden of CVD in the US. In fact, depending on availability of resources and the risk/benefit tradeoff of preventive treatments, a risk threshold such as the moderate risk category (10-20%) may be more beneficial since the non-LB Framingham algorithm would have a better overall sensitivity/specificity ratio (0.77/0.57) if selected as the treatment threshold in the black population.

#### Cost-effectiveness of non-LB Framingham Algorithm

In the cost-effectiveness analysis where the individual CVD risk factors approach focusing on treating diabetes and/or hypertension was considered as the status quo, the non-LB absolute CVD risk approach helped detect more true CVD cases at a lower cost. The individual risk factors approach had a 25% higher average cost-effectiveness ratio. Over the 12year follow-up period, the discounted costs were 21% greater in the individual risk factors approach and 5% fewer CVD cases were identified.

Whereas the cost of screening and preventive interventions was higher in the non-LB approach, the cost of follow-up visits and treating CVD in false negative cases was very high in the individual CVD risk factor approach. The high follow-up costs are as a result of a higher number of true and false positives, and the ‘one size fits all’ approach taken by the individual CVD risk factor strategy. Individuals with hypertension and/or diabetes were scheduled for the minimum 2 visits per year recommended by the ADA guidelines. Ref The non-LB absolute CVD risk approach grades the number of follow-up visits based on the absolute risk score. Individuals with absolute CVD risk core below 20 are scheduled for a minimum 1 annual visit, while those with higher scores are scheduled

for 2 visits per year. The graded approach and the fewer true and false positive cases scheduled for follow-up by the non-LB approach makes the approach cost-effective.

The high number of discounted false negatives associated with the individual risk factors approach (13 more) further increased the costs associated with the approach due to the downstream expenditure of treating the resultant CVD cases. These downstream expenses make a strong case why using a risk assessment approach with high sensitivity and specificity is important. As illustrated in the methods section, clustering of multiple CVD risk factors is known to have an additive and synergistic effect that is not well captured by aggregating individual risk factors. The false negatives associated with the individual CVD risk factor approach occur early and are costed using lightly discounted treatment costs.

The higher expenditure for screening and primary prevention interventions associated with the non-LB approach manifest a strategic investment in prevention with an overall goal of avoiding expensive downstream costs of treating CVD in the false negative cases. The lower follow-up costs associated with the approach reflects how the absolute CVD risk scores enable directing intensive interventions to those who need them most as widely reported in the literature (Beswick et al., 2008; Wan et al., 2009).

In the cost-effectiveness analysis where the LB absolute CVD risk approach was considered as the status quo, the non-LB absolute CVD risk approach also helped detect more true CVD cases at a lower cost. The LB absolute CVD risk approach had a 14% higher average cost-effectiveness ratio. Over the 12year follow-up period, the discounted costs were 9% greater in the LB absolute CVD risk approach and 6% fewer CVD cases were identified.

The cost of screening using the LB approach was about three times higher per person compared to the non-LB approach, but the cost of preventive interventions was comparable for both strategies. The cost of follow-up visits was high in the non-LB absolute CVD risk approach since more individuals met the treatment threshold (moderate risk category) when screened using this approach.

The high number of discounted false negatives associated with LB approach (15 more) increased the downstream expenditure of treating the resultant CVD cases. These downstream expenses make a strong case why the high sensitivity associated with the non-LB Framingham algorithm makes it superior to the LB approach. Reducing the number of false negative cases cuts down expensive downstream costs. It is important to point out that although the false negative cases associated with the LB approach are higher than the individual CVD risk factors approach, they occur later and are hence costed using heavily discounted treatment costs.

The higher downstream costs also indicate that the costly screening associated with the LB approach was not matched with enhanced sensitivity. The slight improvement in specificity may be helpful in cases where false positive cases could be subjected to adverse therapies. However, from a cardiovascular standpoint, adverse effects from preventive therapies for the most part have less impact than unmitigated CVD risk. The lack of significant improvement in sensitivity/specificity with additional testing of lipids raises questions about the need for and relevance of these tests in predicting CVD. It is important to point out that while absolute CVD risk scores are recommended to guide treatment by major evidence based guidelines, caution is given

against ignoring individual risk factors (World Health Organization, 2007). For instance, screening of lipids and other biomarkers could be helpful based on clinician's discretion, especially in populations which have a tendency towards subclinical dyslipidemia without elevations in non-LB indicators such as BMI.

Implementing the non-LB rather than the LB approach would save about \$50,000 for every extra true CVD case detected. The negative incremental cost-effectiveness ratio denotes that the non-LB approach is both more effective, and less costly compared to the status quo.

#### Possible relevance to sub-Saharan Africa (SSA)

Although the global burden of disease statistics have been used to support a premise that SSA is exempt from the epidemic of CVD sweeping across developing countries, the limitations of the estimates have been detailed in various studies (R. S. Cooper, Osotimehin, Kaufman, & Forrester, 1998; Kariuki, Stuart-Shor, Leveille, & Hayman, 2015). The few rigorous studies focusing on CVD risk factors in the region suggest that the region may not be spared from the epidemiological transition as previously thought.

Results from the STEPwise approach to Surveillance (STEPS) surveys commissioned by the WHO indicate that more than 75% of all STEPS participants in sub-Saharan Africa have had at least one major risk factor for CVD. The most prevalent risk factors observed across the region include: high age-adjusted BMI especially in women, elevated systolic blood pressure, low consumption of fruits and vegetables, and increased levels of fasting blood glucose (Mensah, 2013).



A Malawian national representative survey conducted in 2009 using the STEPS approach reported that the age-adjusted prevalence of hypertension was 33.2% in participants aged between 25 to 64 years. Seventy-five percent of these participants reported never having their blood pressure checked previously, and over 94.9% of those with hypertension were not aware of their condition (Mtsyamboza, Kathyola, Dzowela, & Bowie, 2012). Similar observations have been made by other researchers in sub-Saharan Africa who have reported high rates of hypertension, sometimes exceeding those observed for the same age group in developed countries (Mathenge, Foster, & Kuper, 2010).

These data suggest that the epidemiology of CVD in SSA may not well understood or appreciated in the current global burden of disease statistics. Ignoring the problem, as it is currently happening, would lead to missed opportunities for primary prevention which eventually translates to high downstream costs of treating CVD as demonstrated in the cost-effectiveness analysis.

The equivalence of regression coefficients and comparability of performance between the blacks and whites in the US suggests that the validated non-LB Framingham algorithm may perform well even among groups with varying social determinants of health. Therefore, despite the well-known differences in social determinants of health, the validated non-LB Framingham algorithm may provide a beginning point for feasible CVD prevention in SSA pending validation studies.

The impressive benefits of primary prevention compared to no intervention has been simulated for SSA. For instance, pharmacotherapeutics primary prevention efforts targeting populations with more than 25% ten-year absolute risk of CVD were associated

with an incremental cost-effectiveness ratio of \$771 for each healthy year of life saved (QALY) (Gaziano, Opie, & Weinstein, 2006). Despite the current economic constraints in SSA, the estimated cost-effectiveness ratios for primary prevention are still considered feasible because they are below the WHO threshold which considers an intervention to be cost-effectiveness if it costs less than three times the gross national income per head to gain a QALY (Murray, Evans, Acharya, & Baltussen, 2000).

### Conclusion

Taken in total the results observed in this study demonstrate the validity and cost-effectiveness of the non-LB Framingham CVD risk assessment algorithm. The non-LB approach could provide a valuable and efficient alternative to the traditional LB-based approaches in the ongoing efforts to address the high burden of CVD in underserved communities especially the black population in the US. Due to lack of local dataset data or locally derived algorithms, the validated non-LB Framingham CVD risk assessment algorithm may provide a beginning point for initiating feasible CVD risk surveillance and guiding prevention programs in SSA.

## REFERENCES

- Abell, J. E., Egan, B. M., Wilson, P. W., Lipsitz, S., Woolson, R. F., & Lackland, D. T. (2007). Age and race impact the association between BMI and CVD mortality in women. *Public Health Reports (Washington, D.C.: 1974)*, 122(4), 507-512.
- American Diabetes Association. (2014). Standards of medical care in diabetes--2014. *Diabetes Care*, 37 Suppl 1, S14-80. doi:10.2337/dc14-S014 [doi]
- ARIC Investigators. (1989). The atherosclerosis risk in communities (ARIC) study: Design and objectives. the ARIC investigators. *American Journal of Epidemiology*, 129(4), 687-702.
- Barr, V. J., Robinson, S., Marin-Link, B., Underhill, L., Dotts, A., Ravensdale, D., & Salivaras, S. (2003). The expanded chronic care model: An integration of concepts and strategies from population health promotion and the chronic care model. *Hospital Quarterly*, 7(1), 73-82.
- Beswick, A. D., Brindle, P., Fahey, T., & Ebrahim, S. (2008). *A systematic review of risk scoring methods and clinical decision aids used in the primary prevention of coronary heart disease*. London: Royal College of General Practitioners. Retrieved from <http://www.ncbi.nlm.nih.gov/books/NBK55818/pdf/TOC.pdf>. (GRA)
- Bodenheimer, T., Wagner, E. H., & Grumbach, K. (2002). Improving primary care for patients with chronic illness. *JAMA: The Journal of the American Medical Association*, 288(14), 1775-1779.

- Bonow, R. O., Grant, A. O., & Jacobs, A. K. (2005). The cardiovascular state of the union: Confronting healthcare disparities. *Circulation*, *111*(10), 1205-1207.  
doi:111/10/1205 [pii]
- Brennan, R. L., Baker, E., & Metzler, M. (2008). *Promoting health equity: A resource to help communities address social determinants of health*. (). Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.
- Bureau of Labor Statistics. (2014). Occupational employment and wages: 29-1141 registered nurses. Retrieved from <http://www.bls.gov/oes/current/oes291141.htm>
- Centers for Medicare & Medicaid Services. (2015a). Clinical diagnostic laboratory fee schedule. Retrieved from <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ClinicalLabFeeSched/clinlab.html>
- Centers for Medicare & Medicaid Services. (2015b). CPT/HCPCS codes included in range 99201 - 99205. Retrieved from <https://www.cms.gov/medicare-coverage-database/staticpages/cpt-hcpcs-code-range.aspx?DocType=LCD&DocID=32007&Group=1&RangeStart=99201&RangeEnd=99205>
- Centers for Medicare & Medicaid Services. (2015c). Physician fee schedule search. Retrieved from <https://www.cms.gov/apps/physician-fee-schedule/search/search-criteria.aspx>

- Chapman, R. H., Liu, L. Z., Girase, P. G., & Straka, R. J. (2011). Determining initial and follow-up costs of cardiovascular events in a US managed care population. *BMC Cardiovascular Disorders*, 11, 11-2261-11-11. doi:10.1186/1471-2261-11-11 [doi]
- Cleves, M. A. (2008). *An introduction to survival analysis using Stata* (2nd ed.). College Station, Tex.: Stata Press. Retrieved from <http://www.loc.gov/catdir/enhancements/fy0917/2009277451-d.html>
- Cooney, M. T., Dudina, A. L., & Graham, I. M. (2009). Value and limitations of existing scores for the assessment of cardiovascular risk: A review for clinicians. *Journal of the American College of Cardiology*, 54(14), 1209-1227. doi:10.1016/j.jacc.2009.07.020
- Cooper, R., Cutler, J., Desvigne-Nickens, P., Fortmann, S. P., Friedman, L., Havlik, R., Thom, T. (2000). Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the united states: Findings of the national conference on cardiovascular disease prevention. *Circulation*, 102(25), 3137-3147.
- Cooper, R. S., Osotimehin, B., Kaufman, J. S., & Forrester, T. (1998). Disease burden in sub-Saharan Africa: What should we conclude in the absence of data? *The Lancet*, 351(9097), 208-210. doi:[http://dx.doi.org.ezproxy.lib.umb.edu/10.1016/S0140-6736\(97\)06512-4](http://dx.doi.org.ezproxy.lib.umb.edu/10.1016/S0140-6736(97)06512-4)

D'Agostino, R. B., & Pencina, M. J. (2016). Cardiovascular disease (10-year risk).

Retrieved from <https://www.framinghamheartstudy.org/risk-functions/cardiovascular-disease/10-year-risk.php#>

D'Agostino RB, S., Grundy, S., Sullivan, L. M., Wilson, P., & CHD Risk Prediction Group. (2001a). Validation of the Framingham coronary heart disease prediction scores: Results of a multiple ethnic groups investigation. *JAMA: The Journal of the American Medical Association*, 286(2), 180-187.

D'Agostino RB, S., Grundy, S., Sullivan, L. M., Wilson, P., & CHD Risk Prediction Group. (2001b). Validation of the Framingham coronary heart disease prediction scores: Results of a multiple ethnic groups investigation. *JAMA*, 286(2), 180-187.  
doi:joc10098 [pii]

D'Agostino RB, S., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care: The Framingham heart study. *Circulation*, 117(6), 743-753.  
doi:10.1161/CIRCULATIONAHA.107.699579

Dahlgren, G., & Whitehead, M. (1991). *Policies and strategies to promote social equity in health*. (). Stockholm: Institute for Future Studies. Retrieved from <http://www.core.ac.uk/download/pdf/6472456.pdf>

Dawber, T. R., Meadors, G. F., & Moore, F. E., Jr. (1951). Epidemiological approaches to heart disease: The Framingham study. *American Journal of Public Health and the Nation's Health*, 41(3), 279-281.

Drummond, M., & Drummond, M. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). Oxford; New York: Oxford University Press.  
Retrieved from <http://www.loc.gov/catdir/enhancements/fy0640/2005279427-d.html>; <http://www.loc.gov/catdir/enhancements/fy0640/2005279427-t.html>;  
<http://www.loc.gov/catdir/enhancements/fy0726/2005279427-b.html>

Escarce, J. J. (2007). *Racial and ethnic disparities in access to and quality of health care*.

Gaziano, T. A., Opie, L. H., & Weinstein, M. C. (2006). Cardiovascular disease prevention with a multidrug regimen in the developing world: A cost-effectiveness analysis. *Lancet*, 368(9536), 679-686. doi: S0140-6736(06)69252-0 [pii]

Gaziano, T. A., Young, C. R., Fitzmaurice, G., Atwood, S., & Gaziano, J. M. (2008). Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: The NHANES I follow-up study cohort. *Lancet*, 371(9616), 923-931. doi:10.1016/S0140-6736(08)60418-3

Golden, S. D., & Earp, J. A. (2012). Social ecological approaches to individuals and their contexts: Twenty years of health education & behavior health promotion interventions. *Health Education & Behavior: The Official Publication of the Society for Public Health Education*, 39(3), 364-372. doi:10.1177/1090198111418634; 10.1177/1090198111418634

Grundey, S. M., Pasternak, R., Greenland, P., Smith, S., Jr, & Fuster, V. (1999). Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: A statement for healthcare professionals from the American Heart Association and the American College of Cardiology. *Circulation*, 100(13), 1481-1492.

Hayman, L. L., Helden, L., Chyun, D. A., & Braun, L. T. (2011). A life course approach to cardiovascular disease prevention. *The Journal of Cardiovascular Nursing*, 26(4 Suppl), S22-34. doi:10.1097/JCN.0b013e318213ef7f; 10.1097/JCN.0b013e318213ef7f

IDF Clinical Guidelines Task Force. (2006). Global guideline for type 2 diabetes: Recommendations for standard, comprehensive, and minimal care. *Diabetic Medicine: A Journal of the British Diabetic Association*, 23(6), 579-593. doi:DME1918 [pii]

Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. (). Washington, D.C.: National Academy Press.



Institute of Medicine. (2003). Unequal treatment: Report of the institute of medicine on racial and ethnic disparities in healthcare. *The Annals of Thoracic Surgery*, 76(4), S1377-81. Retrieved from <https://www.iom.edu/~media/Files/Report%20Files/2003/Unequal-Treatment-Confronting-Racial-and-Ethnic-Disparities-in-Health-Care/Disparitieshcproviders8pgFINAL.pdf>

Jilcott, S. B., Keyserling, T. C., Samuel-Hodge, C. D., Johnston, L. F., Gross, M. D., & Ammerman, A. S. (2007). Validation of a brief dietary assessment to guide counseling for cardiovascular disease risk reduction in an underserved population. *Journal of the American Dietetic Association*, 107(2), 246-255.  
doi:10.1016/j.jada.2006.11.006

Kamble, S., & Boyd, A. S. (2008). Health disparities and social determinants of health among African-American women undergoing percutaneous coronary interventions (PCI). *Journal of Cultural Diversity*, 15(3), 132-142.

Kariuki, J. K., Stuart-Shor, E. M., Leveille, S. G., & Hayman, L. L. (2013). Evaluation of the performance of existing non-laboratory based cardiovascular risk assessment algorithms. *BMC Cardiovascular Disorders*, 13, 123-2261-13-123.  
doi:10.1186/1471-2261-13-123; 10.1186/1471-2261-13-123

Kariuki, J. K., Stuart-Shor, E. M., Leveille, S. G., & Hayman, L. L. (2015).

Methodological challenges in estimating trends and burden of cardiovascular disease in sub-Saharan Africa. *Cardiology Research and Practice*, 2015, 921021.

doi:10.1155/2015/921021 [doi]

Lloyd-Jones, D., Adams, R. J., Brown, T. M., Carnethon, M., Dai, S., De Simone, G., . . .

American Heart Association Statistics Committee and Stroke Statistics

Subcommittee. (2010). Heart disease and stroke statistics--2010 update: A report from the American Heart Association. *Circulation*, 121(7), e46-e215.

doi:10.1161/CIRCULATIONAHA.109.192667;

10.1161/CIRCULATIONAHA.109.192667

Lloyd-Jones, D. M. (2010a). Cardiovascular risk prediction: Basic concepts, current status, and future directions. *Circulation*, 121(15), 1768-1777.

doi:10.1161/CIRCULATIONAHA.109.849166;

10.1161/CIRCULATIONAHA.109.849166

Lloyd-Jones, D. M. (2010b). Cardiovascular risk prediction: Basic concepts, current status, and future directions. *Circulation*, 121(15), 1768-1777.

doi:10.1161/CIRCULATIONAHA.109.849166;

10.1161/CIRCULATIONAHA.109.849166

- Mathenge, W., Foster, A., & Kuper, H. (2010). Urbanization, ethnicity and cardiovascular risk in a population in transition in Nakuru, Kenya: A population-based survey. *BMC Public Health*, 10, 569-2458-10-569. doi:10.1186/1471-2458-10-569; 10.1186/1471-2458-10-569
- May, M., Lawlor, D. A., Brindle, P., Patel, R., & Ebrahim, S. (2006). Cardiovascular disease risk assessment in older women: Can we improve on Framingham? British women's heart and health prospective cohort study. *Heart (British Cardiac Society)*, 92(10), 1396-1401. doi:10.1136/hrt.2005.085381
- Medicaid.gov. (2016). Pharmacy drug pricing. Retrieved from <https://www.medicaid.gov/medicaid-chip-program-information/by-topics/benefits/prescription-drugs/pharmacy-pricing.html>
- Mendis, S., Lindholm, L. H., Anderson, S. G., Alwan, A., Koju, R., Onwubere, B. J., . . . Heagerty, A. (2011). Total cardiovascular risk approach to improve efficiency of cardiovascular prevention in resource constrain settings. *Journal of Clinical Epidemiology*, 64(12), 1451-1462. doi:10.1016/j.jclinepi.2011.02.001
- Mendis, S., Lindholm, L. H., Mancia, G., Whitworth, J., Alderman, M., Lim, S., & Heagerty, T. (2007). World health organization (WHO) and international society of hypertension (ISH) risk prediction charts: Assessment of cardiovascular risk for prevention and control of cardiovascular disease in low and middle-income countries. *Journal of Hypertension*, 25(8), 1578-1582. doi:10.1097/HJH.0b013e3282861fd3

- Mendis, S., Puska, P., Norrving, B., World Health Organization, World Heart Federation, & World Stroke Organization. (2011). *Global atlas on cardiovascular disease prevention and control*. Geneva: World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization.
- Mensah, G. A. (2008). Ischaemic heart disease in Africa. *Heart (British Cardiac Society)*, 94(7), 836-843. doi:10.1136/hrt.2007.136523; 10.1136/hrt.2007.136523
- Mensah, G. A. (2013). Descriptive epidemiology of cardiovascular risk factors and diabetes in sub-Saharan Africa. *Progress in Cardiovascular Diseases*, 56(3), 240-250. doi:10.1016/j.pcad.2013.10.014; 10.1016/j.pcad.2013.10.014
- Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., . . . American Heart Association Statistics Committee and Stroke Statistics Subcommittee. (2015). Heart disease and stroke statistics--2015 update: A report from the American Heart Association. *Circulation*, 131(4), e29-322. doi:10.1161/CIR.0000000000000152 [doi]
- Msyamboza, K. P., Kathyola, D., Dzowela, T., & Bowie, C. (2012). The burden of hypertension and its risk factors in Malawi: Nationwide population-based STEPS survey. *International Health*, 4(4), 246-252. doi:10.1016/j.inhe.2012.09.005; 10.1016/j.inhe.2012.09.005

- Murray, C. J., Evans, D. B., Acharya, A., & Baltussen, R. M. (2000). Development of WHO guidelines on generalized cost-effectiveness analysis. *Health Economics*, 9(3), 235-251. doi:10.1002/(SICI)1099-1050(200004)9:3<235::AID-HEC502>3.0.CO;2-O [pii]
- National Rural Health Association. (2013). Rural health foundation: Caring for the future of rural America. Retrieved from <http://www.ruralhealthweb.org/>
- Pandya, A., Weinstein, M. C., & Gaziano, T. A. (2011). A comparative assessment of non-laboratory-based versus commonly used laboratory-based cardiovascular disease risk scores in the NHANES III population. *PloS One*, 6(5), e20416. doi:10.1371/journal.pone.0020416; 10.1371/journal.pone.0020416
- Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, 36(4), 859-866. doi:10.1111/j.1745-9125.1998.tb01268.x
- Pearson, T. A., Blair, S. N., Daniels, S. R., Eckel, R. H., Fair, J. M., Fortmann, S. P., . . . Taubert, K. A. (2002). AHA guidelines for primary prevention of cardiovascular disease and stroke: 2002 update: Consensus panel guide to comprehensive risk reduction for adult patients without coronary or other atherosclerotic vascular diseases. American heart association science advisory and coordinating committee. *Circulation*, 106(3), 388-391.

- Selvarajah, S., Haniff, J., Kaur, G., Guat Hiong, T., Bujang, A., Chee Cheong, K., & Bots, M. L. (2013). Identification of effective screening strategies for cardiovascular disease prevention in a developing country: Using cardiovascular risk-estimation and risk-reduction tools for policy recommendations. *BMC Cardiovascular Disorders*, 13, 10-2261-13-10. doi:10.1186/1471-2261-13-10 [doi]
- Shillinglaw, B., Viera, A. J., Edwards, T., Simpson, R., & Sheridan, S. L. (2012). Use of global coronary heart disease risk assessment in practice: A cross-sectional survey of a sample of U.S. physicians. *BMC Health Services Research*, 12, 20-6963-12-20. doi:10.1186/1472-6963-12-20; 10.1186/1472-6963-12-20
- Simon, H. (2009). Waisted: Abdominal obesity and your health. *Harvard Men's Health Watch*, 13(6), 1-6.
- Smith, D. H., & Gravelle, H. (2001). The practice of discounting in economic evaluations of healthcare interventions. *International Journal of Technology Assessment in Health Care*, 17(2), 236-243.
- Stuart-Shor, E. M., Berra, K. A., Kamau, M. W., & Kumanyika, S. K. (2012). Behavioral strategies for cardiovascular risk reduction in diverse and underserved racial/ethnic groups. *Circulation*, 125(1), 171-184. doi:10.1161/CIRCULATIONAHA.110.968495

- University of North Carolina at Chapel Hill. (2013). Atherosclerosis Risk in Communities study: ARIC study design. Retrieved from [http://www2.csc.unc.edu/aric/study\\_design](http://www2.csc.unc.edu/aric/study_design)
- Wan, Q., Harris, M. F., Zwar, N., Campbell, T., Patel, A., Vagholkar, S., . . . Denney-Wilson, E. (2009). Study protocol for a randomized controlled trial: The feasibility and impact of cardiovascular absolute risk assessment in Australian general practice. *American Heart Journal*, 157(3), 436-441. doi:10.1016/j.ahj.2008.11.016
- Weber, M. A., Schiffrin, E. L., White, W. B., Mann, S., Lindholm, L. H., Kenerson, J. G., Harrap, S. B. (2014). Clinical practice guidelines for the management of hypertension in the community a statement by the American Society of Hypertension and the International Society of Hypertension. *Journal of Hypertension*, 32(1), 3-15. doi:10.1097/HJH.0000000000000065 [doi]
- Weinstein, M. C., Siegel, J. E., Gold, M. R., Kamlet, M. S., & Russell, L. B. (1996). Recommendations of the panel on cost-effectiveness in health and medicine. *JAMA*, 276(15), 1253-1258.
- Whitehead, M., & Dahlgren, G. (1991). What can be done about inequalities in health? *Lancet*, 338(8774), 1059-1063.
- WHO. (2015). Cardiovascular diseases fact sheet. Retrieved from <http://www.who.int/mediacentre/factsheets/fs317/en/#>

Will, R., Keydron, G., Cynthia, A., Luis, M., & Zachary, B. (2011). Social determinants of Health—Its place in Redefining medicine. *Association of American Medical Colleges Annual Meeting*, Denver, Colorado.

Wilson, R. (2013, 09/03/2013). Hispanics most likely to go without health insurance. *Washington Post*

World Health Organization, (2007). *Prevention of cardiovascular disease: Guidelines for assessment and management of cardiovascular risk*. Geneva: WHO.

Yusuf, S., Hawken, S., Ounpuu, S., Bautista, L., Franzosi, M. G., Commerford, P., . . . INTERHEART Study Investigators. (2005). Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: A case-control study. *Lancet (London, England)*, 366(9497), 1640-1649. doi:S0140-6736(05)67663-5 [pii]

Yusuf, S., Hawken, S., Ounpuu, S., Dans, T., Avezum, A., Lanas, F., . . . INTERHEART Study Investigators. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study. *Lancet*, 364(9438), 937-952. doi:10.1016/S0140-6736(04)17018-9