2017-07-19

Apoptosis and Atherosclerosis: From Modeling Sudden Cardiac Death to Bioluminescence-Based Detection of Programmed Cell Death

Trajen Rex Head

University of Miami, trajenhead@gmail.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation
https://scholarlyrepository.miami.edu/oa_dissertations/1911

This Embargoed is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
UNIVERSITY OF MIAMI

APOPTOSIS AND ATHEROSCLEROSIS:
FROM MODELING SUDDEN CARDIAC DEATH TO
BOLUMINESCENCE-BASED DETECTION OF PROGRAMMED CELL DEATH

By

Trajen Rex Head

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2017
UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

APOPTOSIS AND ATHEROSCLEROSIS: FROM MODELING SUDDEN CARDIAC DEATH TO BIOLUMINESCENCE-BASED DETECTION OF PROGRAMMED CELL DEATH

Trajen Rex Head

Approved:

Sylvia Daunert, Ph.D.  Sapna K. Deo, Ph.D.
Professor and Lucille P. Markey Chair  Associate Professor of
Biochemistry and Molecular Biology  Biochemistry and Molecular
Associate Director, Dr. JT Macdonald  Biology
Biomedical Nanotechnology Institute

Pascal J. Goldschmidt-Clermont, M.D.  Guillermo Prado, Ph.D.
Professor and Dean Emeritus,  Dean of the Graduate School
Leonard M. Miller School of Medicine

Roberto I. Vazquez-Padron, Ph.D.
Associate Professor of Surgery
Cardiovascular disease (CVD) is the leading cause of mortality globally, responsible for over 17.3 million deaths annually (~31% of all deaths). Over the next decade, this value is projected to increase, surpassing 23.6 million annual deaths by 2030. The vast majority of these deaths are caused by complications resulting from progressive thickening of the arterial wall, and the formation and growth of arterial plaques within the vessel lumen. Subsequent plaque rupture may potentially lead to blood clot (thrombus) formation. The combined effect of these processes is the reduction of vessel diameter (stenosis) and subsequent reduction in blood flow (ischemia), while complete occlusion of a vessel via thrombus formation may precipitate a heart attack or stroke. Utilizing discrete event simulation (DES) modeling, the various processes of lesion formation and growth were examined, and it was demonstrated that an increased number of raised lesions in an individual at an early age represents a reliable biomarker for increased risk of subsequent sudden cardiac death (SCD). As a result of this finding, the levels of known CVD risk factors for a group of ~3,000 young people (ages 15-34) were investigated. Using generalized linear regression modeling (GLM) the manner in which these factors contribute to lesion formation and growth was analyzed. From this analysis,
a subpopulation (~13%) exhibiting an increased number of raised lesions was identified. However, no correlation between accelerated atherosclerosis and measured levels of risk factors was observed in this group, suggesting one or more “hidden” risk factors. Thus, a biosensor capable of identifying this “high-risk” group presenting a significantly increased number of lesions would be extremely beneficial. Atherosclerotic plaques possess a large number of apoptotic and necrotic cells. Based on this, a fusion protein was designed and expressed that was capable of targeting and binding to these apoptotic cells (Annexin V), as well as generating a detectable bioluminescent signal (RLuc8). This Annexin-Renilla Fusion Protein (ArFP) was subsequently characterized and applied both \textit{in vitro} and \textit{in vivo}, demonstrating that apoptosis detection is possible in a number of samples including vascular tissue. Plaque identification through apoptosis detection with this unique sensor provides another step towards the early identification and possible intervention of those at the highest risk of CVD-related death.
for my Mother

You have always been an inspiration, showing me that regardless of the situation, there is always reason to smile, to trust, and to be at peace.

for my Father

Thank you for showing me the worth of hard work and perseverance. You taught me to constantly strive to be better tomorrow than I am today.

for my Sister

I cannot express how grateful I am that you have always been there for me, whether we were one mile or a thousand miles apart.

“Together forever, never apart. Maybe by distance, but never at heart”
“When you eliminate the impossible, whatever remains, however improbable, must be the truth.”

Sir Arthur Conan Doyle, 1859-1930

Captain Spock, USS Enterprise, 2230-2285

“Sometimes science is more art than science. A lot of people don’t get that.”

Rick Sanchez

“Science isn’t about why, it’s about why not.”

Cave Johnson, Aperture Laboratories

“Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines… Our posturings, our imagined self-importance, the delusion that we have some privileged position in the Universe, are challenged by this point of pale light. Our planet is a lonely speck in the great enveloping cosmic dark. In our obscurity, in all this vastness, there is no hint that help will come from elsewhere to save us from ourselves.”

Carl Sagan, 1934-1996
CONTENTS

LIST OF FIGURES ........................................................................................................ viii
LIST OF TABLES .......................................................................................................... xvii

1 Introduction | Cardiovascular Disease

1.1 Cardiovascular Disease: (Still) A Growing Epidemic ........................................... 1
  1.1.1 Atherosclerosis: Ancient and Inflammatory ....................................................... 2
  1.1.2 Formation of a Plaque ......................................................................................... 6
1.2 Imaging .................................................................................................................. 13
  1.2.1 Invasive Coronary Angiography ......................................................................... 13
  1.2.2 Magnetic Resonance Imaging ........................................................................... 14
  1.2.3 Noninvasive Coronary Computed Tomography Angiography ......................... 18
  1.2.4 Invasive Analyses for Plaque Morphology and Hemodynamics ....................... 22
  1.2.5 Molecular Imaging ............................................................................................ 30
  1.2.6 Hybrid Technology and Image-based Computational Models ......................... 34
1.3 Closing Remarks .................................................................................................... 43

2 Discrete Event Simulation Model of Sudden Cardiac Death

2.1 Perspective ............................................................................................................. 45
2.2 Results ................................................................................................................... 46
  2.2.1 Brief Model Description .................................................................................... 46
  2.2.2 Comparison of the Model to Pathological Data ............................................... 47
  2.2.3 Model Predictions of Cumulative and Annual SCD Risk .................................. 51
  2.2.4 Model Predictions of Coronary Screening and Treatment ............................... 53
  2.2.5 Model Predictions of Medications .................................................................... 58
  2.2.6 Model Predictions of Atheroma Distributions .................................................. 59
2.3 Discussion ............................................................................................................... 62
2.4 Methods ................................................................................................................ 65
  2.4.1 Age of First Atheroma Formation .................................................................... 65
  2.4.2 Atheroma Growth ............................................................................................. 66
  2.4.3 Likelihood of Arrhythmic Cardiac Death .......................................................... 69

3 Identification of Accelerated Atherosclerosis in Youth

3.1 Perspective ............................................................................................................. 71
3.2 Results ................................................................................................................... 74
  3.2.1 Identification and Analysis of the Discovery Cohort ........................................ 74
4 Bioluminescence-based Apoptosis Detection

4.1 Perspective ................................................................. 97
4.2 Results ........................................................................... 101
  4.2.1 Design, Expression, and Characterization of ArFP .............. 101
  4.2.2 Activity Characterization of ArFP .................................. 104
  4.2.3 Apoptosis Detection In Vitro ........................................ 105
  4.2.4 Apoptosis Detection Ex Vivo ....................................... 108
  4.2.5 Apoptosis Detection In Vivo and Disease Model ............... 109
4.3 Discussion......................................................................... 114
4.4 Methods ......................................................................... 115
  4.4.1 Reagents and Kits ........................................................ 115
  4.4.2 Fusion Gene Design and Transformation .......................... 117
  4.4.3 Protein Expression ...................................................... 118
  4.4.4 Verification and Characterization of ArFP ......................... 118
    4.4.4.1 Western Blot Analysis of Protein Expression .............. 118
    4.4.4.2 Bioluminescence Measurement and Characterization .... 119
    4.4.4.3 Circular Dichroism (CD) Spectroscopy Analysis .......... 119
    4.4.4.4 Isothermal Titration Calorimetry (ITC) ..................... 120
  4.4.5 Cell Culture ............................................................. 120
    4.4.5.1 Apoptosis Induction ............................................. 120
    4.4.5.2 Flow Cytometric Analysis of Apoptosis ...................... 121
    4.4.5.3 Western Blot Analysis of Caspase Activation ............. 121
  4.4.6 Animal Models of Apoptosis ......................................... 122
    4.4.6.1 Ex Vivo Rat Iliac Artery Apoptosis Induction ............ 122
    4.4.6.2 In Vivo Mouse Corneal Apoptosis Detection .............. 123
    4.4.6.3 In Vivo Mouse Retina Apoptosis Detection ............... 123
    4.4.6.4 Statistical Analyses of Animal Studies ....................... 124

5 Conclusions and Future Work

5.1 Perspective ................................................................... 125
5.2 Modeling the Future ....................................................... 126
5.3 Illuminating Potential Directions ...................................... 127
Appendix A Discrete Event Simulation Model of Sudden Cardiac Death Supporting Information

A.1 Structure of the DES Model .............................................................. 133
A.2 Parameters of the Model ............................................................... 133

Appendix B Identification of a Subpopulation in Youth with Accelerated Atherosclerosis Supporting Information

B.1 Univariate Analysis of CVD Risk Factors ........................................ 139
B.2 Generalized Linear Regression Modeling ...................................... 142
  B.2.1 Description of the Models ......................................................... 142
  B.2.2 Single Population GLM for the Discovery Cohort (739 Subjects) ... 144
  B.2.3 Two-Group Approach GLM for the Low-Risk Subset of the Discovery Cohort (646 Subjects) ........................................ 144
  B.2.4 Two-Group Approach GLM for the High-Risk Subset of the Discovery Cohort (93 Subjects) ........................................... 144
  B.2.5 Two-Group Approach GLM for the High-Risk Subset of the Verification Cohort (250 Subjects) ......................................... 145
  B.2.6 Two-Group Approach GLM for the High-Risk Subset of the Confirmation Cohort (250 Subjects) ...................................... 145
B.3 Tables of Correlation Coefficients and p-Values ............................. 145
B.4 Tables of GLM Estimated Coefficients and Statistics ...................... 152

Appendix C Bioluminescence-based Apoptosis Detection Supporting Information

C.1 Apoptosis and Its Detection ............................................................. 157
  C.1.1 Defining Apoptosis ................................................................. 157
  C.1.2 Targeting Apoptosis ............................................................... 158
C.2 Fusion Gene Design and Transformation ........................................ 162
C.3 Dynamic Light Scattering (DLS) of Lipid Vesicles ......................... 162
C.4 Cell Culture of Caco-2 Cells ........................................................ 162
  C.4.1 Apoptosis Induction ............................................................... 163
  C.4.2 Cell Metabolic Activity (MTS) Assay ....................................... 166

References ............................................................................................. 168
LIST OF FIGURES

1.1 Hemodynamic forces exerted on the arteries. (a) Trans-mural pressure exerted radially across the vessel wall. (b) Endothelial Shear Stress (ESS), determined as the product of blood viscosity (\(\eta\)) and its shear rate (\(\delta v/\delta y\)) such that ESS = \(\eta \times \left(\delta v/\delta y\right)\). (c) Typical sites of atherosclerotic plaque formation (lateral walls of bifurcations, inner wall of vessel curvature, and near branch points). Laminar flow in these regions is disrupted, causing oscillatory or reversed flow and subsequent low ESS. ........................................ 7

1.2 Formation and progression of an atherosclerotic plaque. As an inflammatory disease, the initial stages of atherosclerosis involve an inflammatory insult to the epithelial cells lining the artery lumen. In response, these cells begin to express adhesion molecules that facilitate the transmigration of monocytes into the vessel intima. These monocytes then differentiate into macrophages and begin engulfing lipid and lipid products, forming foam cells. As foam cells aggregate, they form the characteristic fatty streak, while many macrophages begin to undergo apoptosis. Inefficient clearance of apoptotic macrophages leads to secondary necrosis, resulting in a growing lipid-rich necrotic core. In response to the growing lesion, smooth muscle cells migrate to the intima, helping to form the overlying fibrous cap. Rupture of this cap can expose the necrotic core, leading to thrombus formation and subsequent acute cardiovascular events. ......................................................... 9

1.3 Identification of high-intensity plaques (HIPs) using Magnetic Resonance Imaging (MRI). (a) Magnetic Resonance Angiography (MRA) of coronary plaque. (b) Presence of high-intensity signal on T1-weighted imaging (T1WI) of same coronary plaque. (c) Coregistered images of non-contrast T1WI and MRA. Adapted from [77] with permission. ........................................ 17

1.4 Coronary atherosclerosis imaging with coronary computed tomography angiography (CCTA). (a) Proximal right coronary artery (RCA, arrows) with no detectable plaque. (b) Proximal RCA exhibiting calcified plaque (arrows). (c) Proximal RCA containing calcified (large arrows) and non-calcified (small arrow) plaques. (d) Proximal RCA with non-calcified plaques (arrows). Adapted from [99] with permission. ........................................ 21

1.5 Identification of plaque features via grayscale IVUS. (a) Image demonstrating attenuation (asterisk) indicating large necrotic core. (b) Calcified nodule (asterisk) and acoustic shadowing. (c) Small deposits or spotty calcification (arrow head). (d) Multi-layer appearance correlated with plaque progression and increased risk in heart transplant recipients. Adapted from [108] with permission. ......................................................... 24
1.6 Post-processing techniques for classification of lesions identified via IVUS. (a) Thin-cap fibroatheroma, (b) Thick-cap fibroatheroma, (c) fibrotic, and (d) fibrocalcific plaques identified with VH-IVUS. (e) Grayscale and (f) analysis using IB-IVUS. (g) Grayscale and (h) analysis of using iMAP-IVUS. Figures in panels a-d adapted from [112], and figures in e-f adapted from [113], with permissions.

1.7 Analysis of plaque morphologies using optical coherence tomography. (a) Fibrotic plaque. (b) Thin-cap fibroatheroma (arrow). (c) Fibrocalcific plaque. (d) Lipid-rich fibroatheroma. In each frame, the OCT guide wire is indicated with an asterisk (*). Figure adapted from [115] with permission.

1.8 Coronary atherosclerotic plaque analysis via near-infrared spectroscopy (NIRS) and a hybrid NIRS-IVUS probe. (a) 2-dimensional chemogram representing pullback position (in millimeters) of the probe on the x-axis and the circumferential position in degrees along the y-axis, with pixel color representing the probability of lipid present at the respective location. Colors range from red (low probability) to yellow (high probability). (b) Representative schematic of a combined NIRS-IVUS hybrid catheter tip. (c) Three cross-sections of pullback imaging superimposed with corresponding IVUS image. Figures in panels a and c adapted with permissions from [126]. Figure in panel b adapted with permissions from [109].

1.9 Determination of fractional flow reserve (FFR) in (a) a healthy artery and (b) across an atherosclerotic lesion. Pressure is measured both distal to the lesion (P_d) as well as in the aorta (P_a). The FFR is calculated as the ratio of these values (P_d/P_a), and provides a means to determine flow reduction caused by the lesion. FFR values lower than 0.75-0.8 typically considered representative of significant stenosis.

1.10 Molecular imaging of coronary atherosclerosis using PET. (a) Invasive coronary angiogram of patient with acute ST-segment elevation myocardial infarction showing proximal occlusion (red arrow) of the LAD. (b) ^18^F-fluoride uptake (red arrow) at the site of the atherosclerotic plaque. (c) ^18^FDG imaging of the same segment, demonstrating high background signal from both the myocardium (yellow arrow) and the esophagus (blue arrow). Figure adapted with permissions from [138].

1.11 Hybrid IVUS-OCT imaging. (a) IVUS imaging indicating presence of calcifications and acoustic shadowing (yellow arcs). (b) OCT imaging indicating fibrous cap (blue arrow) and underlying detail. Adapted with permission from [151].

1.12 Stages of computational model development for analysis of a coronary artery. (a) Angiogram of the RCA. (b) Segmentation of the RCA and
identification of vessel walls. (c) Surface and volumetric meshes fit to the segmented vessel. (d) Physiologic parameters utilized for the generation of boundary conditions such as blood pressure and flow rates. (e) Iterative simulation of the model and results output demonstrating in silico determination of FFR. (f) Validation of model results to standard or known measurements. Adapted from [166] with permission.  

2.1 Using SCD-Predict, cohorts of 100 men (blue) and 100 women (red) were simulated and analyzed. Visualization of the findings include (a) simulated age at time of death, (b) total number of atheroma, (c) total vulnerable atheroma, and (d) number of arrhythmias. 

2.2 Graphical representations of the (a) cumulative and (b) annual risk of SCD for both men (blue) and women (red). Each point in the cumulative risk curves was generated by running SCD-Predict for the cohort of 1,000 simulated persons with fixed simulation stop times. Annual risk curves were generated by calculating the derivative of the cumulative risk. Vertical lines indicate the cumulative or annual risk of SCD at 80 years of age. 

2.3 Survival curves illustrating the impact of various screening strategies with parameters of both scan age and threshold of atheromas varied. (a) Impacts of screening strategies showing survival curves for identified male subpopulations both with (solid lines) and without (dashed lines) screening. (b) Effects of screening strategies showing the survival curves for the entire male population. (c) Focus region of the survival curves presented for the entire male population; dashed line indicates survival of half of the simulated cohort. (d) Impacts of screening strategies showing survival curves for identified female subpopulations both with (solid) lines and without (dashed lines) screening. (e) Effects of screening strategies showing the survival curves for the entire female population. (f) Focus region of the survival curves presented for the entire female population; dashed line indicates survival of half of the simulated cohort. 

2.4 Comparison of the effects of various screening strategies with variations in scan age as well as thresholds for number of atheromas. Examination of average increase in treated (a) men and (b) women, as well as the total number of life years added to the male population of (c) men and (d) women. 

2.5 Comparisons of untreated and treated male (blue) and female (red) populations with various medication treatment parameters (medications #1 and #2 reducing atheroma and clot growth rates, respectively). Data for each gender normalized to the case of no treatment. Simulated cohorts containing 5,000 individuals for each condition provided information on
2.6 Average numbers of thin-cap, medium-cap, thick-cap, and total atheromas as well as the average number of observed ruptures in various populations including: (a) Men having died of non-SCD causes, (b) Women having died of non-SCD causes, (c) Men having died of SCD, (d) Women having died of SCD. (e) Comparison of the total number of atheromas for men having died of either SCD or non-SCD causes. (f) Comparison of the total number of atheromas for women having died of either SCD or non-SCD causes.

2.7 Comparison of the average number of atheromas for men, women, and total simulated population between individuals simulated to have died of SCD and individuals simulated to have died of non-SCD causes. In each case, the average number of atheromas is significantly higher (p-value <10^{-6}) for the SCD group compared to the corresponding non-SCD group.

3.1 Identification and analysis of the Discovery Cohort. (a) Demographic comparison of the entire PDAY study population (n = 2,651) and the identified Discovery Cohort (n = 739). (b) Age distribution of the subjects in the Discovery Cohort. (c) Distribution of the number of observed raised lesions in the RCA of subjects belonging to the Discovery Cohort. (d) Identification of 93 subjects exhibiting an elevated number of coronary raised lesions (red circles). (e) Demographic analysis of the 93 subjects belonging to the high-risk subset of the Discovery Cohort. (f) Comparison of the ages (left) and observed number of coronary raised lesions (right) between the 646 subjects belonging to the main population of the Discovery Cohort (low-risk) and the identified high-risk subset of 93 subjects.

3.2 Generalized linear models of the number of raised lesions in the RCA. (a) Comparison of the observed number of raised lesions with predictions from the two approaches used for GLM analysis (single population model and two-group approach). (b) Predicted number of raised lesions for the 646 subjects belonging to the general population of the discovery (low-risk subset) varying only levels of LDL. Additional predictions for the low-risk subset of the Discovery Cohort demonstrating effects varying multiple risk factor levels, including varying serum thiocyanate levels at the (c) 50th percentile of LDL and (d) 75th percentile LDL measured in the Discovery Cohort. Dashed line indicates predicted number of lesions at age 25.

3.3 Identification and analysis of the Verification Cohort. (a) Demographic comparison of the entire PDAY study population (n = 2,651) and the identified Verification Cohort (n = 1,912). (b) Age distribution of the subjects in the Verification Cohort. (c) Distribution of the number of
observed raised lesions in the RCA of subjects belonging to the Verification Cohort. (d) Identification of 250 subjects exhibiting an elevated number of coronary raised lesions (red circles). (e) Demographic analysis of the 250 subjects belonging to the high-risk subset of the Verification Cohort. (f) Comparison of the ages (left) and observed number of coronary raised lesions (right) between the 1,662 subjects belonging to the main population of the Verification Cohort (low-risk) and the identified high-risk subset of 250 subjects. .................................................................

3.4 Identification and analysis of the Confirmation Cohort. (a) Demographic comparison of the entire PDAY study population (n = 2,651) and the identified Confirmation Cohort (n = 751). (b) Age distribution of the subjects in the Confirmation Cohort. (c) Distribution of the number of observed raised lesions in the RCA of subjects belonging to the Confirmation Cohort. (d) Identification of 84 subjects exhibiting an elevated number of coronary raised lesions (red circles). (e) Demographic analysis of the 84 subjects belonging to the high-risk subset of the Confirmation Cohort. (f) Comparison of the ages (left) and observed number of coronary raised lesions (right) between the 667 subjects belonging to the main population of the Confirmation Cohort (low-risk) and the identified high-risk subset of 84 subjects. .................................................................

3.5 (a) Comparison of predictions for the number of coronary raised lesions from the Discovery, Verification, and Confirmation Cohorts as a function of age. (b) Predictions for the number of raised coronary lesions in subjects of the Confirmation Cohort at various levels of LDL to account for its minor contribution to the equation describing this cohort. Dashed line indicates predicted number of lesions for each model at age 25. ......................

3.6 Summary of cohort PDAY study population and cohort identification and analysis. The PDAY study population (white) consists of 2,651 subjects. The Discovery Cohort (green) of 739 was identified based on availability of demographic and risk factor measurements. From the Discovery Cohort, a high-risk subset (green, red outline) of 93 subjects (12.6%) was identified. The Verification Cohort (light blue) consisted of 1,912 subjects not included in the Discovery Cohort. A similar high-risk subset from the verification cohort (light blue, red outline) was identified and contained 250 subjects (13.1%). The Confirmation Cohort (dark blue) consisted of 751 subjects from within the Verification Cohort based on availability of risk factor measurements. The high-risk subset of this cohort (dark blue, red outline) included 84 subjects (11.2%). .......................................................

4.1 Pathways of apoptosis (extrinsic – blue box, and intrinsic – orange box) and detection concept. Both pathways converge to the activation of caspases that are principally responsible for morphological and biochemical changes associated with apoptosis including
phosphatidylserine (PS) exposure. Although PS is sequestered to the inner-face of the plasma membrane under healthy conditions, the ArFP construct exhibits the ability to bind PS at surface of apoptotic cells and generate a bioluminescent signal in the presence of substrate coelenterazine.

4.2  (a) SDS-PAGE analysis of components and fusion protein (lane 1: Annexin V, lane 2: RLuc8, lane 3: ArFP, lane 4: Precision Plus Protein Ladder (Bio-Rad)). (b) Western blot analysis of components and fusion protein identified using anti-Annexin V antibody (lane 1: Precision Plus Protein Ladder, lane 2: Annexin V, lane 3: RLuc8, lane 4: ArFPP) or anti-Renilla luciferase antibody (lane 5: Precision Plus Protein Ladder, lane 6: Annexin V, lane 7: RLuc8, lane 8: ArFP). (c) Far-UV CD spectra of ArFP (purple), Annexin V (teal), and RLuc8 (blue). (d) Protein $^2\theta$ structure assignment (CDSSTR analysis via Dichroweb) for ArFP. (e) Thermostability analysis (ellipticity at 222 nm) of ArFP. (f) Comparison of Annexin V and ArFP thermostability (upper) and first derivative analysis of curves (lower). (g) Comparison of RLuc8 and ArFP thermostability (upper) and first derivative analysis of curves (lower).

4.3  (a) Top: ITC data from the titration of 30 µM ArFP with lipid vesicles composed of 60:40 mixture of POPC:POPS in the presence of 0.75 mM Ca$^{2+}$ at 18 °C. Middle: Heat of dilution (control/baseline). Bottom: Integrated heats of binding as a function of ligand:protein mole ratio. (b) Bioluminescence spectra from 400 nm to 700 nm resulting from addition of coelenterazine substrate to ArFP (purple) or RLuc8 (blue). (c) Kinetic bioluminescence data resulting from injection of coelenterazine substrate to ArFP (purple) or RLuc8 (blue) (250 integrations at 0.24 s/each). (d) Concentration-dependent bioluminescence output comparison of ArFP (purple) or RLuc8 (blue) after injection of coelenterazine substrate (10 s integration). All points are the mean of three measurements ± one standard deviation. Error bars that are not visible are obstructed by the data point. ...
measurements ± one standard deviation. Error bars that are not visible are obstructed by the data point.

4.5 (a) Diagram of operational procedure, illustrating position of treated and collected tissues (teal: right iliac/treated condition, purple: left iliac/reperfusion control, blue: carotid/healthy control). (b) Luminescence comparisons of treated (teal) and untreated (purple) iliac arteries as well as the carotid (blue) artery as a healthy control, performed in triplicate. (c) IVIS image of the carotid (1), left iliac (2), and right iliac (3) arteries after treatment. (d) Histological analysis of treated vascular tissue indicating the emergence of apoptotic cells following treatment.

4.6 (a) IVIS images comparing mice from groups Control 1 (addition of protein with no injury or substrate), Control 2 (addition of substrate with no injury or protein), Untreated (addition of protein and substrate with no injury), and Treated (addition of protein and substrate following corneal injury). (b) Comparison of bioluminescence data from the total cohort (n = 13) of mice receiving ethanol treatment (Treated, n = 8) or PBS (Untreated, n = 5). (c) TUNEL analysis of corneal sections. Whole mount (top) and cross section (bottom) images of corneal tissues from untreated (PBS, left) and treated (ethanol, right) mice indicating the presence of apoptotic (green) and non-apoptotic (blue) nuclei.

4.7 (a) IVIS analysis of bioluminescence comparing control (Saline injection) and retinal degeneration model (NaIO3 injection) mice. (b) Histological analysis of retinal sections from control (A1, PBS injection) and retinal degeneration model (A2 and A3, NaIO3 injection) mice. Left: H&E stain of retinal cross sections. Right: TUNEL analysis of retinal cross sections indicating presence of apoptotic (green) and non-apoptotic (blue) nuclei.

5.1 Comparison of bioluminescence recorded from ArFP incubated with aorta samples from either wtC57BL/6 mouse (left) or ApoE−/− mouse (right). To ensure valid comparison of data, the signals were normalized to the mass of the collected tissue.

5.2 Expression of additional Annexin V fusion proteins (AFPs). (a) SDS-PAGE analysis of three AFPs: Annexin V-Renilla luciferase (ArFP), Annexin V-aequorin (AqFP), and Annexin V-Gaussia luciferase (AgFP). Calculated molecular weights for each component and fusion protein provided. (b) Western blot analysis of ArFP and components using anti-Annexin V monoclonal mouse antibody (left) or anti-RLuc rabbit polyclonal antibody (right). (c) Western blot analysis of AqFP and components using anti-Annexin V monoclonal mouse antibody (left) or anti-aequorin rabbit polyclonal antibody (right). (d) Western blot analysis of AgFP and components using anti-Annexin V monoclonal mouse antibody (left) or anti-GLuc rabbit polyclonal antibody (right).
pPROEX HTb/Annexin V, respectively. (b) Visualization on 1% agarose gel of ArFP gene resulting from OE-PCR of prepared Annexin V and RLuc8 genes with appropriate terminal sequences coding for inclusion into pET-30 Xa/LIC. Expected gene sizes: RLuc8 (993 nt), Annexin V (969 nt), and ArFP (1944 nt).

C3 Schematic of OE-PCR. 1) Annealing of primers with appropriate LIC (orange) and linker (green) regions to template. 2) Extension of primers. 3) Gene product containing LIC or linker extension. 4) Annealing of primers to product from (3) while template strands are used for additional cycles. 5) Extension of primers. 6) Generation of complete gene with LIC and linker sequences. 7) Annealing of linker sequences and primers. 8) Extension of primers. 9) Generation of complete fusion gene. 10) Amplification.

C4 pET-30/ArFP vector expression system plasmid map.

C5 Dynamic light scattering (DLS) measurements of prepared 60:40 POPC:POPS vesicles. Vesicles were prepared as multilamellar vesicles, followed by downsizing via extrusion through a 0.2 µm polycarbonate filter for a minimum of 21 times. The resulting vesicles exhibited a mean diameter of less than 200 nm.

C6 (a) Bioluminescence response of ArFP incubated with Caco-2 cells after treatment with 1 µM staurosporine for various times. (b) Absorbance at 490 nm of wells containing cells incubated with MTS reagent for 4 hours following treatment with 1 µM staurosporine for various times.
LIST OF TABLES

2.1 Comparison of the mean numbers of ruptured and non-ruptured fibrous cap atheromas simulated with the SCD-Predict model (sim) and observed in the pathological study from which data used to fit the model was obtained [185]. Weighted average (w.a.) totals were calculated as 0.64(value for men) + 0.36(value for women) to reflect the gender distribution observed in the pathological study. ................................................................. 49

2.2 Comparison of the percentage of ruptures, thrombi, and absence of thrombi in atherosclerotic plaques of coronary arteries simulated with SCD-Predict and observations of an additional pathological study [183]. ......................... 50

2.3 Increases in average life expectancy for total simulated populations as well as subpopulations identified using three different screening strategies. ...... 56

3.1 Number of raised lesions. Comparison of the observed and modeled distributions and correlation coefficients. For this analysis, race (white = 1, black = 2) and sex (female = 1, male = 2) were given quantifiable values. ... 80

3.2 Number of raised lesions. Comparison of the observed and modeled distributions and correlation coefficients. ................................................................. 90

4.1 Summary of Protein Characteristics. ................................................................. 104

A1 Parameters of the DES Model. ................................................................. 138

B1 Correlation Coefficients. Discovery Cohort, all 739 Subjects. Significant Correlations (p<0.05) Highlighted in Yellow. ......................... 145

B2 Correlation Coefficients. Discovery Cohort, 340 Subjects Aged 15-24 Significant Correlations (p<0.05) Highlighted in Yellow. ......................... 146

B3 Correlation Coefficients. Discovery Cohort, 399 Subjects Aged 25-34 Significant Correlations (p<0.05) Highlighted in Yellow. ......................... 147

B4 Correlation Coefficients. Discovery Cohort, 646 Subjects of the Low-Risk Subset. Significant Correlations (p<0.05) Highlighted in Yellow. ............... 148

B5 Correlation Coefficients. Discovery Cohort, 93 Subjects of the High-Risk Subset. Significant Correlations (p<0.05) Highlighted in Yellow. ............... 149

B6 p-Values. Discovery Cohort, all 739 Subjects. ................................................................. 150
| B7 | p-Values. Discovery Cohort, 93 Subjects of the High-Risk Subset. .......... | 151 |
| B8 | Estimated Coefficients and Model Statistics, Single Population Model of the Discovery Cohort (739 Subjects). .......................................................... | 152 |
| B9 | Estimated Coefficients and Model Statistics, Two-Group Model of the Discovery Cohort, Low-Risk Subset (646 Subjects). ................................. | 153 |
| B10 | Estimated Coefficients and Model Statistics, Two-Group Model of the Discovery Cohort, High-Risk Subset (93 Subjects). ....................... | 154 |
| B11 | Estimated Coefficients and Model Statistics, Two-Group Model of the Verification Cohort, High-Risk Subset (250 Subjects). .................... | 155 |
| B12 | Estimated Coefficients and Model Statistics, Two-Group Model of the Confirmation Cohort, High-Risk Subset (84 Subjects). ...................... | 155 |
Chapter 1

Introduction | Cardiovascular Disease

1.1 Cardiovascular Disease: (Still) A Growing Epidemic

At the beginning of the 20th century, deaths attributed to cardiovascular disease (CVD) accounted for less than 10% of global totals [1, 2]. Today, however, this metric has grown to over 30%, representing over 17.5 million deaths annually and ranking CVD as the primary cause of mortality worldwide [3]. CVD encompasses a large number of diseases affecting either the heart or blood vessels, with clinical manifestations of the disease including angina pectoris, arrhythmia, myocardial infarction, heart failure, sudden cardiac death, and stroke among others. Not only are these events extremely prevalent, they are often devastating: sudden cardiac arrest arising from CVD results in a survival rate of less than 1% worldwide [4]. However, the health burden associated with CVD is not equivalently distributed across all countries and populations. Specifically, this burden is exceedingly prominent in lower- and middle-income countries (LIMIC) and those identified as newly industrialized countries (NIC) such as Mexico, China, and India [5-11]. While the incidence rates of CVD in the United States and other developed nations have steadily decreased since the 1960s to their current level of approximately 10%, over 80% of all CVD-related deaths occur in LIMIC and NIC [5-7]. As a result, although the impact of this disease may be receding in several developed nations, it continues to be a critical global threat. Consequently, a considerable volume of research over the last decade has been focused on techniques and technologies to understand the fundamental
mechanisms for the progression of atherosclerosis – the underlying cause for the vast majority of all CVD-related complications.

In this chapter, the pathological and biochemical processes of atherosclerotic plaque formation and growth will first be examined. This will be followed by an overview of imaging techniques (both invasive and non-invasive), briefly discussing important progress in hybrid imaging techniques and image-based computational methods that provide a greater understanding of this progressive disease.

1.1.1 Atherosclerosis: Ancient and Inflammatory

Atherosclerosis is defined by the formation and growth of plaques within arterial lumen with concurrent loss of vascular elasticity. Eventually, this condition can lead to consequent blood flow reduction through the affected vessel, and has long been associated with CVD-related death. Around the year 1505, Leonardo da Vinci recorded his observations of a “parched and shrunk and withered … artery that feeds the heart,” attributing this as the cause of death of an elderly man who passed away “without any movement or sign of anything amiss” [12]. However, CVD remained a relatively uncommon cause of death until the dramatic surge in CVD mortality in the past 100 years, due to a number of important factors including a staggering reduction in communicable diseases resulting in significant life expectancy increases. Coupled with the augmentation of numerous lifestyle-associated CVD risk factors during this time period, many began to believe that atherosclerosis is an exclusively modern disease. Recent reports challenging this view demonstrate that mummies from various ancient societies – from as much as 5,300 years ago and spanning 4,000 years of human history – exhibit a high prevalence of vascular calcification; a condition pathognomonic with
atherosclerosis [13-16]. Additionally, a single nucleotide polymorphism (SNP) known to be a genetic risk factor of CVD (minor allele of rs10757274 on chromosome 9p21) was identified in at least one of these individuals, despite the fact that it was initially believed that this mutation arose much later in history [15]. Although the diets and lifestyles of these various peoples and societies were fundamentally different from our modern equivalent, risk factors such as excessive smoke inhalation from indoor cooking fires, as well as frequent and chronic inflammation caused by infection, contributed to widespread prevalence of atherosclerosis [13, 14]. In contrast, risk factors most commonly associated with disease in modern populations include modifiable risks such as high cholesterol, physical inactivity, or smoking as well as non-modifiable risks including gender, numerous identified SNPs, and race [17, 18]. Algorithms designed to interpret these factors provide relative individual risk scores (Framingham [19], QRISK®-2-2016 [20], ACC/AHA ASCVD [21]), and indicate that the strongest risk factor is simply age; increasing age while holding all remaining risk variables constant results in a marked increase in risk [17, 22-24]. The average annual rates of first CVD events for men increase over 20-fold between the age groups 35-44 (incidence of 3 per 1000) and 85-94 (incidence of 74 per 1000), while similar rates are observed in women 10 years later in life [25]. Similarly, the average age of myocardial infarction is 65 and 72 for men and women respectively [3]. These acute clinical events occurring at advanced ages often manifest in the absence of associated symptoms, and are frequently fatal [26]. However, atherosclerosis itself – as the principal origin for the majority of these events – is a chronic, inflammatory, progressive disease with initial manifestations beginning at young
ages [27-29]. Evidence of the disease has been reported in multiple arterial regions in children under the age of 10, and in the aortas of individuals less than a year old [28, 29].

Historically, atherosclerosis was believed to be exclusively a result of dyslipidemia, attributing an increase in low-density lipoprotein (LDL) from lifestyle and diet as the principal cause of disease. However, elevated circulating LDL alone is generally insufficient to account for the observed extent of disease, and is responsible for at most half of arterial lesions with atherosclerosis also found in the complete absence of hyperlipidemia [30, 31]. Near the end of the 20th century, the notion that atherosclerosis was a fundamentally inflammatory disease began to gain popularity, catalyzed largely by Russel Ross’ pioneering report titled “Atherosclerosis—An Inflammatory Disease” [32]. The formation and progression of atherosclerotic plaques as we understand today largely support Ross’ own “response-to-injury” hypothesis, implicating plaque formation as a consequence of focal endothelial cell (EC) injury and inflammation [30, 33]. Although many believed that oxidized lipoprotein products served as this initial endothelial insult, evidence demonstrating this effect in humans is limited and – at best – tenuous. Meanwhile, alternative mechanisms continue to become more robust. Specifically, as more of the underlying mechanisms and pathways controlling atherosclerosis become identified, it becomes clear that this disease does not simply follow a single, forward direction of irreversible progression. In contrast, it is evident now that atherosclerosis is the result of a myriad of independent pathways and their complex interactions [34]. Recently, it has been shown that genetic expression profiles and epigenetic modifications exert complex control and regulatory mechanisms for atherosclerosis [23, 35]. Perhaps most critical for the understanding of this disease, recent findings have established that
the most determinant factors for atherosclerosis progression are repair mechanisms, slowing or preventing the progression of disease. Specifically, a growing body of evidence indicates that vascular repair effected by bone marrow (BM)-derived vascular progenitor cells serves to counter the effects of endothelial injury and dysfunction that would otherwise lead to atherosclerosis progression [22-24]. Using a well-established mouse model of atherosclerosis it was shown that intravenous injection of BM-derived progenitor cells from young donors could “rescue” aged recipients from disease [22]. Additionally, this work demonstrated that the ability of these BM-derived progenitors to elicit reparative effects was lost with age. Other work has identified the accumulation of senescent cells that produce pro-inflammatory cytokines and matrix metalloproteinases (MMPs) within atherosclerotic sites, but not in adjacent surrounding tissues [36]. These senescent cells have been shown to propagate atherosclerosis, while data suggests that BM-derived progenitor cells may protect from these effects through reduction of endothelial senescence [22, 36, 37]. More recently, a novel role of aging as a means to modulate cardiovascular disease risk was identified in hematopoietic progenitor cells. With aging, these cells have been shown to acquire somatic DNA mutations that provide competitive growth advantages and lead to an expanded pool of mutated clones (a process known as clonal hematopoiesis of intermediate potential, or CHIP) [38]. A recent study examined over 8,000 individuals from case-control studies, and revealed that carriers of CHIP had nearly double the risk of coronary heart disease when compared to non-carriers [39]. Moreover, this study demonstrated that the loss of function mutation of Tet2 (the second most commonly mutated gene in CHIP) led to accelerated atherosclerosis in mice, likely as a result of transcriptional modifications of associated
macrophages [39]. As our understanding of this complex disease advances, it has become clear that there are numerous competing forces of disease progression and repair. It appears likely therefore that, in this case, disease progression is the result of both the accumulation of signals and systems to advance the formation of atherosclerotic plaques, and the decline in the efficiency of repair mechanisms from aging and exposure to traditional risk factors [17, 22]. To better understand current advances in both the imaging of atherosclerosis as well as computational modeling of the disease, it is important to review the progressive process by which arterial plaques form, and the importance of inflammation and repair throughout.

1.1.2 Formation of a Plaque

Although risk factors for atherosclerotic disease systemically affect all regions of the vasculature, plaque formation is not uniformly distributed. Instead, atherosclerotic plaques typically occur at a number of reproducible focal sites such as the inner wall of vessel curvature, the lateral walls of bifurcations, and near branch points. An early theory to explain this plaque distribution disparity relied on the relative amount of shear stress exerted on the endothelial cells (ECs) in contact with the blood as it traverses the dynamic lumen geometry. There are two dominant forces exerted by blood on the inner lining of ECs within the vasculature. These hemodynamic forces include transmural pressure (a force exerted radially across the vessel wall, Figure 1.1a) and shear stress (the frictional force applied to the ECs by blood tangential to the direction of flow, Figure 1.1b). The value of endothelial shear stress (ESS, sometimes wall shear stress or WSS) is determined as a function of blood viscosity (η) and its shear rate (\(\frac{\delta v}{\delta y}\), a measure of the rate that blood velocity increases from the endothelium toward the center of the vessel,
Figure 1.1b). Blood flow in regions of relatively straight, uniform diameter vessels exhibits nearly laminar flow, resulting in ESS values of 20-70 dyne/cm$^2$ (2-7 Pa) \([40, 41]\). However, at the aforementioned locations prone to atherosclerotic plaque formation, the laminar flow characteristics are disrupted, causing regions of oscillatory or reversed flow and a subsequent reduction in ESS (< 10 dyne/cm$^2$, Figure 1.1c). In these regions of disturbed flow, a number of events and responses to low shear stress occur that begin the process of atherosclerotic plaque formation.

Following the “response-to-injury” hypothesis of atherosclerosis formation and progression, conditions of low overall ESS have been shown to facilitate additional insults to the endothelium leading to dysfunction. One such path initiated by low ESS is
the sustained activation of sterol regulatory elements binding proteins (SREBPs). Once activated, these endoplasmic reticulum-bound transcription factors contribute to a strong inflammatory response and increase the expression of genes that encode for the LDL receptor (LDLR) [42, 43]. Yet another change associated with low ESS is the shift in EC morphology from tightly packed cells aligned with the direction of blood flow to a damaged morphology with cuboidal cells demonstrating no observed preferential alignment, and resulting in leakier junctions between cells [44-46]. This disorganized cell morphology is compounded by the loss of repair mechanisms effected by progenitor cells. As discussed previously, BM-derived progenitor cells play a critical role in protecting against atherosclerosis. One reparative role played by these cells is the restoration of damaged endothelial tissue or engraftment into these damaged regions to reverse endothelial injury [24]. The net result of this upregulation of LDLR and increased cell permeability is a dramatic increase in levels of LDL infiltration in regions of low ESS (Figure 1.2, Stage I) [41]. Additional pro-inflammatory pathways activated under conditions of low ESS include the mitogen-activated protein (MAP) kinase and nuclear factor κB (NF-κB) signaling pathways, while levels of miR-10a and its regulatory effects on the NF-κB pathway are attenuated under these conditions [47, 48].

The combined net result of declining vascular repair mechanisms coupled with endothelial inflammation and the activation of ECs is the capture of circulating monocytes. This process, known as the leukocyte adhesion cascade, begins with the interaction of monocytes and displayed adhesion molecules in the regions of inflammation, leading to the stepwise rolling, firm adhesion, and transmigration of the monocyte into the vascular intima (Figure 1.2, Stage II) [49, 50]. This process of
Figure 1.2: Formation and progression of an atherosclerotic plaque. As an inflammatory disease, the initial stages of atherosclerosis involve an inflammatory insult to the epithelial cells lining the artery lumen. In response, these cells begin to express adhesion molecules that facilitate the transmigration of monocytes into the vessel intima. These monocytes then differentiate into macrophages and begin engulfing lipid and lipid products, forming foam cells. As foam cells aggregate, they form the characteristic fatty streak, while many macrophages begin to undergo apoptosis. Inefficient clearance of apoptotic macrophages leads to secondary necrosis, resulting in a growing lipid-rich necrotic core. In response to the growing lesion, smooth muscle cells migrate to the intima, helping to form the overlying fibrous cap. Rupture of this cap can expose the necrotic core, leading to thrombus formation and subsequent acute cardiovascular events.
monocyte recruitment is largely directed by the presence of EC-derived chemokines (including CCL5, CCL2, and CX3CL1) and their respective receptors as well as adhesion molecules such as VCAM-1 and ICAM-1 [33, 51]. Following transmigration from the lumen into the intima, these monocytes undergo differentiation into macrophages capable of engulfing the growing pool of apoB-containing lipoprotein such as LDL and its modified variants [52]. These differentiated macrophages identify and engulf LDL through expression of scavenger receptors such as oxidized LDL receptor 1 (LOX-1), scavenger receptor A (SRA), and CD36 (Figure 1.2, Stage II) [52, 53]. Interestingly, through the use of CD36/SRA double knockouts, it has been shown that lipid-laden macrophages will still accumulate in atherosclerotic vessel walls, suggesting additional routes of lipid uptake [54]. Other proposed mechanisms include the capability of macrophages to accumulate LDL via receptor-independent pinocytosis, and the hydrolysis of free cholesterol from LDL aggregates [55, 56]. Cholesterol liberated from lipoproteins taken up by these pathways is transported within the macrophage to the endoplasmic reticulum (ER) where it is esterified to cholesteryl fatty acid esters by acetyl-coenzyme A:cholesterol acetyltransferase 1 (ACAT1) and neutral cholesterol ester hydrolase (nCEH), while lipid transport out of the cell is controlled by ATP-binding cassette (ABC) transporters [52]. Under atherosclerotic conditions, macrophages demonstrate both an increase in lipid uptake and inefficient efflux of cholesteryl esters, generating characteristic lipid-laden foam cells (Figure 1.2, Stage II) [57]. It has been shown that such defects in lipid metabolism can cause ER stress and dysfunction, ultimately leading to apoptosis of the cell [58]. Compounding this problem, defects in cholesterol efflux also inhibit efficient efferocytosis (clearance of apoptotic cells by
nearby macrophages), leading to an accumulation of apoptotic cells in these regions. The growing population of apoptotic macrophages and foam cells generate the characteristic fatty streak associated with early atherosclerotic plaques (Figure 1.2, Stage III/IV) [59, 60]. Apoptotic cells not cleared through phagocytosis then undergo secondary necrosis, releasing cellular debris along with additional oxidized lipids and pro-inflammatory cytokines that contribute to continued inflammation and a growing necrotic core of the atherosclerotic lesion (Figure 1.2, Stage IV) [56, 59].

Atherogenesis often ceases here, with many lesions progressing no further than the fatty streak, due in large part to various reparative mechanisms [17]. However, some lesions undergo additional changes within the vascular intima causing a pronounced growth of the plaque and remodeling of the vessel itself. Specifically, macrophages and foam cells within the fatty streak begin releasing inflammatory cytokines (interleukin (IL)-1, IL-4, and tumor necrosis factor (TNF)-α) that interact with various growth factors (platelet-derived growth factor (PDGF), CD-40) to activate matrix metalloproteinases (MMPs) [61]. Once activated, these MMPs are responsible for remodeling the extracellular matrix (ECM) surrounding the vascular smooth muscle cells (VSMC) of the vessel, while growth factors released by components of the fatty streak promote migration of these VSMCs to the growing lesion [62]. These VSMCs, exhibiting a non-contractile or synthetic phenotype, then begin to generate collagen-rich fibrous ECM, eventually representing a major component of this newly formed fibroatheroma (Figure 1.2, Stage V) [56]. Complicating this process further, neither elevated levels of cholesterol nor extensive macrophage involvement are strictly required for this VSMC proliferation. In addition, deposition and growth of calcium granules may also occur, and
has been shown to exhibit both stabilizing and destabilizing effects to the vulnerability of
the fibrous cap depending on their extent and location [63]. Modulations of the ECM and
proliferation of VSMCs are also required for arterial remodeling, which can be either
constrictive or expansive (as a compensatory mechanism to counter atherosclerotic
plaque growth) [56, 64]. These advanced plaques also possess the potential to rupture,
exposing the thrombogenic core to the blood flow of the vessel lumen (Figure 1.2, Stage
VI). A common trait of plaques vulnerable to rupture is a thin covering of ECs, VSMCs,
and ECM – termed thin-cap fibroatheromas (TFCAs) and possessing a cap of <65 µm in
thickness [65]. Activation of MMPs from abundant macrophages present in the plaque
may cause degradation of the ECM near the cap, destabilizing endothelial cell adhesion
and allowing for exposure of the necrotic core, leading to subsequent thrombus formation
in the lumen of the vessel. Although this is a rare fate for atherosclerotic plaques (most
either do not progress to advanced stages or become stable plaques with caps >65 µm),
the subsequent formation of a thrombus from plaque rupture can cause complete
occlusion of the local vessel. The thrombus may also embolize and travel to a secondary
vessel causing ischemia, myocardial infarction, or sudden coronary death [56, 66].
Alternatively, repeated cycles of clinically silent rupture of the cap followed by healing
may contribute to overall plaque growth and luminal stenosis. Another potent source of
plaque growth is intraplaque hemorrhage. As the atherosclerotic plaque grows, the region
becomes progressively more inflamed as well as hypoxic, inducing angiogenesis of
vessels in the vasa vasorum into the plaque itself. These neovessels have been observed
to be fragile and leaky, and are an important source of hemorrhage within the wall of the
artery, contributing not only to further inflammation, but also plaque growth and vessel
stenosis [67]. In this way, tissue ischemia caused by luminal stenosis can occur even in the absence of clinically significant rupture [68].

Atherosclerosis is the most significant health problem globally. We know today that the disease does not follow a simple, unidirectional progression, and is determined by a myriad of pathways, control mechanisms, and repair processes. Ultimately, however, the clinical result for an immensely large number of individuals is the formation and growth of vascular lesions with the potential to rupture, leading to life-threatening conditions. Because of the critical nature of these vascular lesions, it is imperative that methods are developed to both predict and detect their formation, progression, and clinical status. In the next two sections, recent advances in imaging techniques for plaque detection in individuals and populations as well as computational models to predict and understand the progression of disease will be investigated. These techniques will primarily focus on atherosclerosis within the major coronary artery regions, as these commonly generate clinically significant and potentially fatal events, and present unique challenges for numerous imaging techniques due to their size and constant motion.

1.2 Imaging

1.2.1 Invasive Coronary Angiography

For many, the “gold standard” of imaging for coronary artery disease (CAD) is invasive coronary angiography, a technique widely employed and largely unchanged since its accidental discovery in the late 1950s. In the United States alone, over one million patients undergo invasive coronary angiography annually [25]. Through the administration of a radiopaque contrast agent to the major coronary arteries (left anterior descending (LAD), left circumflex (LCX), and right coronary artery (RCA)), this
technique allows for highly accurate imaging of luminal stenosis. Historically, the 2D images generated through this technique would be visually analyzed to identify potential sites of stenosis. More recently, the technique of quantitative coronary angiography (QCA) has emerged, utilizing computer generated algorithms to more reproducibly and reliably determine the presence and severity of stenosis. This technique has been used extensively to predict potential future clinical events, and is used to guide intervention techniques such as balloon angioplasty and stenting. However, this technique is not without its detractors. In addition to being an invasive procedure requiring catheterization, the angiograms provide little to no information on plaque composition or the hemodynamic parameters of the local environment. Additionally, because traditional coronary angiography only visualizes the luminal diameter, positive vascular remodeling may reduce accuracy of this technique by producing angiographic results of seemingly normal luminal diameter despite the presence of substantial disease. Although invasive coronary angiography remains fundamental for analyzing stenotic burden, recent technological advances have fostered the development of other imaging techniques with unique advantages. The two major non-invasive techniques that have arisen to challenge invasive coronary angiography in the detection and quantification of luminal stenoses are magnetic resonance angiography (MRA) and coronary computed tomography angiography (CCTA).

1.2.2 Magnetic Resonance Imaging

Unlike invasive coronary angiography (as well as CCTA discussed later), the use of high-resolution magnetic resonance imaging (MRI) for angiographic analysis is an ionization-free imaging technique. Instead of relying on the differential absorption of X-rays
between tissues (low) and contrast agents (high), magnetic resonance angiography (MRA) generates images of the vasculature using radiofrequency pulses in the presence of a strong magnetic field. When in the presence of such a magnetic field, the protons of water molecules are polarized to the direction of the field. An intense radiofrequency pulse is then applied to these protons, causing the spin direction of the protons to adjust. Following the pulse, the protons undergo relaxation to re-align with the magnetic field, and variances in both concentrations and relaxation times (T1 and T2) of protons in different tissues provide the contrast between these tissues. With this technique, MRA is able to provide information on both the distribution of plaques within the vessel wall and the severity of stenosis. By utilizing different pulse strategies (T1 weighted, T2 weighted, or proton density weighted), different MRI techniques can be employed to identify numerous important aspects of atherosclerotic disease in addition to luminal stenosis including plaque size, morphology, and composition [69]. To increase the contrast of images obtained, the use of gadolinium-based contrast agents is common. These compounds are typically injected intravenously, and decrease the T1 relaxation time of blood, enhancing the signal from the vessel lumen compared to the surrounding tissues.

To date, the application of MRI has largely been focused on imaging atherosclerotic disease in the carotid, aortic, and peripheral artery regions. For example, it was shown early on that MRI could be used to calculate the area of lesions within the carotid artery. This technique was validated by comparing the plaque size calculated via MRI images of the vessel wall to measurements of the same plaque ex vivo following carotid endarterectomy [70]. Additionally, using T1- or T2-weighted relaxation measurements, important morphological characteristics could be identified within carotid plaques such as
lipid composition of the core, fibrous tissue and calcification within the cap, and intraplaque hemorrhage. However, the application of MRI (and by extension, MRA) has lagged behind other techniques for the imaging of coronary arteries and atherosclerotic disease in these vessels. This is due in large part to the relatively small caliber of these vessels (luminal diameters of approximately 3 to 4 mm) and their constant motion; the position of the coronary arteries is affected by both the cardiac cycle as well as respiration, making accurate imaging challenging [71]. To address these challenges, MRI of the coronary arteries can be electrocardiogram (ECG)-gated either prospectively (imaging only during mid-diastole or end-systole when limited motion occurs) or retrospectively (imaging throughout the cardiac cycle, but discarding data gathered during heart motion). To account for respiratory motion, imaging is conducted either during a breath hold (limiting total image acquisition duration) or employs a navigator sequence that tracks the position of the diaphragm, allowing image acquisition only at specific positions [72]. With these advances, coronary MRI techniques for the detection of atherosclerotic disease within the coronary arteries have improved [73]. A recent analysis reported an 86% specificity and a 91% sensitivity for detection via MRA when compared to QCA, demonstrating the reliability of this technique [74]. A benefit of these improvements has been the ability of this imaging modality to detect both thrombus formation and intraplaque hemorrhages. The compound methemoglobin differs from the oxygen transport molecule hemoglobin via oxidation of the heme group from the ferrous (Fe$^{2+}$) to the ferric (Fe$^{3+}$) state. The accumulation of this oxidation product has been associated with thrombus formation as well as intraplaque hemorrhage, and exhibits short T1 relaxation times [75, 76]. These short relaxation times in turn, generate a high signal
in non-contrast T1-weighted MRI, allowing for the identification of these so-called high-intensity plaques (HIPs, Figure 1.3). In a study of 568 individuals with known or suspected CAD, 26% of individuals possessing HIPs exhibiting this and other vulnerable features (positive remodeling, large necrotic core, etc.) as identified via MRI had an adverse coronary event, compared to only 3% of individuals without HIPs [77]. Additionally, because MRI is an ionization-free technique, it provides an opportunity for serial imaging in longitudinal studies and imaging of young patients to potentially generate information on the earliest stages of the disease. However, even with these

Figure 1.3: Identification of high-intensity plaques (HIPs) using Magnetic Resonance Imaging (MRI). (a) Magnetic Resonance Angiography (MRA) of coronary plaque. (b) Presence of high-intensity signal on T1-weighted imaging (T1WI) of same coronary plaque. (c) Coregistered images of non-contrast T1WI and MRA. Adapted from [77] with permission.
benefits and important potential applications, the diagnostic efficiency of MRA in the coronary arteries still falls short of other techniques such as coronary computed tomography angiography (CCTA), and is thus generally not used for clinical applications [78]. In contrast, the use of MRI for the analysis and characterization of both ischemic and non-ischemic myocardial disease has been shown to be critically important. As mentioned, gadolinium (Gd)-based contrast agents are commonly used to enhance the signal-to-noise ratios of MRI data. Importantly, it has been shown that these Gd-based agents are excluded from normal, healthy myocardial cells. However, in the event of a myocardial infarction, the distribution of Gd-based contrast agents is altered. Specifically, the membranes of myocardial cells in the acute infarcted region become disrupted, allowing for uptake of these contrast agents. These regions can then be identified via MRI using a technique known as delayed-contrast enhancement [79]. Moreover, increased uptake in both the peripheral region of the infarcted area and in fibrotic (scar) tissue of chronic infarctions allows for highly descriptive analysis of these conditions. Recently, it has also been shown that various cardiomyopathies can also be identified through this delayed-contrast enhancement methodology [80]. Using these characteristics of uptake and exclusion, delayed-contrast enhancement cardiac MRI can also be used to guide attempts at revascularization, demonstrating another aspect of this versatile technique [79, 81].

1.2.3 Non-invasive Coronary Computed Tomography Angiography

The use of computed tomography (CT) is another non-invasive imaging technique capable of determining the presence of vessel stenosis, plaque characteristics, and disease burden in multiple vessel regions. Notably, coronary CT angiography (CCTA), has found
wider and more rapid acceptance for the clinical detection and analysis of atherosclerosis than corresponding MR techniques. Similar to invasive coronary angiography, CCTA utilizes iodine-based radiopaque contrast agents to visualize vessel lumen. At the time of its introduction over 15 years ago, CCTA scans were strongly hindered by cardiac and respiratory motion as previously described for application of coronary MRA [82]. However, rapid advances in technology have allowed dramatically reduced data acquisition time, significantly improving temporal resolution. Current generation instruments incorporate multidetector systems containing between 64 and 320 detection rows, providing effective temporal resolutions ranging from 175 ms down to 27 ms [83, 84]. Additionally, as equipment continues to improve, the spatial resolution of CT has improved, with current technology providing resolutions of 240 to 300 µm [85]. Not only do these improvements mean that the duration (or necessity) of breath-holds is reduced, but also that the acceptable range of heart rates is expanded. To eliminate cardiac motion artifacts, CCTA techniques will often rely on a system of prospective electrocardiogram (ECG)-gated imaging such that imaging occurs during a window ~15% of the R-R cycle during which the heart remains relatively motionless. At a rate of 60 heartbeats per minute, this corresponds to an imaging window of 150 ms. Although the temporal resolution of current generation CCTA instruments surpass this requirement, the number of evaluable artery segments decreases proportionately as heart rate increases. Thus, the use of beta-blockers to suppress accelerated heart rates (> 80 bpm) is still common to reduce these rates to the target of 60-65 bpm [85, 86]. A common concern regarding the use of CCTA is the requirement of radiation exposure, although advancements in scan protocols and equipment have helped ameliorate this issue as well. Utilizing the latest
scanners (dual 128-slice or 320-slice scanners) with reduced tube voltage and novel acquisition modalities combined with new image reconstruction techniques, the standard radiation dose for CCTA scans has been reduced from ~10 mSv to as low as < 1 mSv [85, 87-89]. Although these advances translate to a wide variability in equipment and techniques, the use of CCTA for the detection of significant (≥ 50%) luminal stenosis demonstrated a median sensitivity of 97.8% and specificity of 89.6% at the patient level across all published meta-analyses of CCTA utilizing at least 64-slice detector systems [90]. Perhaps the greatest strength of luminal stenosis analysis via CCTA, however, is its powerful negative predictive value, regularly determined to be as high as 95-99% [91, 92]. In one 18-month follow-up study of 1265 patients examined via CCTA, it was shown that of the 802 individuals lacking evidence of coronary artery stenosis, merely 0.5% experienced an acute coronary event compared to 4.8% in those with identified obstructive coronary lesions [93]. A five year follow-up study by these same investigators demonstrated that in patients for which coronary stenoses had been ruled out by CCTA, only 3.3% experienced hard cardiac events compared to 10.4% of those with ≥ 1 stenotic lesion [94]. Finally, utilizing data from the CONFIRM registry, a study of over 24,700 patients revealed that for those lacking stenosis identified by CCTA the mean annual death rate was only 0.28% [95]. Thus, although some have argued against the use of CCTA for patients considered to be at moderate- to high-risk, there seems to be consensus that CCTA has emerged as an excellent tool for the non-invasive analysis of coronary stenoses [96, 97].

In addition to its excellent demonstrated capability at identifying coronary stenoses (more specifically, the lack thereof), the use of coronary CT has also been
applied to the analysis of atherosclerotic plaques directly. As previously indicated, the presence of vascular calcification has been shown to serve as a surrogate for overall atherosclerotic plaque burden [98]. Macroscopic deposits of calcium within coronary arteries can be detected using CT imaging without the need for contrast agent (Figure 1.4) [99]. Although these macroscopic calcium features have been shown to confer

**Figure 1.4:** Coronary atherosclerosis imaging with coronary computed tomography angiography (CCTA). (a) Proximal right coronary artery (RCA, arrows) with no detectable plaque. (b) Proximal RCA exhibiting calcified plaque (arrows). (c) Proximal RCA containing calcified (large arrows) and non-calcified (small arrow) plaques. (d) Proximal RCA with non-calcified plaques (arrows). Adapted from [99] with permission.
stabilizing effects to individual plaques, quantification of the total amount of coronary artery calcification (CAC) is associated with overall plaque burden. Additionally, CAC scoring is also used in conjunction with the Framingham Risk Score to improve prognostic value of coronary events [100]. Additionally, because CT exhibits exceptional sensitivity for calcium, a scan revealing complete lack of calcification (CAC = 0) is sufficient evidence to reclassify patients into lower risk categories [101, 102]. However, the sensitivity of CT towards calcium can also be detrimental, with calcium blooming artifacts in acquired images resulting in significantly lowered specificity of CCTA [103]. Thus, it is noteworthy that effective CCTA cannot be performed in patients with heavy coronary calcification (CAC ≥ 400) [97, 103]. The presence and extent of calcification is thus the most common feature of individual plaque morphology identified by coronary CT, but other characteristics have also been identified including potential positive remodeling and the presence of a necrotic core (associated with low attenuation, < 30 Hounsfield units) [85, 104]. Although coronary CT imaging has made tremendous advances in the last decade, limitations to the technique still exist. Specifically, due to the limited current spatial resolution, assessment of atherosclerotic plaque morphological features via CT varies substantially between investigators with low reproducibility. Thus, invasive techniques such as intravascular ultrasound (IVUS) and optical coherence tomography (OCT) remain the “gold standard” in these applications.

1.2.4 Invasive Analyses for Plaque Morphology and Hemodynamics

A variety of invasive techniques are utilized in addition to the imaging modalities discussed thus far for the assessment of various atherosclerotic plaque characteristics. Many of these visualization tools are well established in the interrogation of arterial
plaques, while others have only just recently begun to be explored for their beneficial insights into the understanding of this disease. Each of these techniques, however, provides fundamental information on individual plaque morphology, hemodynamic significance, or disease progression. The following section will briefly discuss the applications and limitations of four of the major intravascular modalities for coronary atherosclerosis assessment; intravascular ultrasound (IVUS), optical coherence tomography (OCT), near infrared spectroscopy (NIRS), and fractional flow reserve (FFR).

For over two decades, intravascular ultrasound (IVUS) has been recognized as being an extremely valuable complementary tool for the investigation of coronary atherosclerosis [105]. To generate images, IVUS utilizes a miniaturized piezoelectric transducer at the tip of a catheter that generates sound waves between 20 to 60 MHz. In grayscale IVUS, the time-delayed reflection of these sound waves is then recorded, producing cross-section images of the vessel with an axial resolution between 150-200 µm, a lateral resolution of ~250 µm, and tissue penetration up to 10 mm [106, 107]. The intensity of the echo has been shown to be associated with various plaque features; very low attenuation is representative of lipid pools and necrotic cores, while bright reflections indicate the presence of calcification (Figure 1.5) [108]. Of note, these calcific deposits have been shown to obstruct further penetration of the ultrasound signal (also referred to as acoustic shadowing), hindering compositional analysis in regions of heavy calcification [109]. Although grayscale IVUS has limited capacity to differentiate lesion components, this basic analysis is capable of providing a quantification of plaque burden, a known predictor of disease progression [110]. To improve the plaque characterization
of IVUS, multiple approaches for the post-processing of collected data have been
designed. Most common, virtual histology (VH)-IVUS uses an advanced algorithm to
analyze the ultrasound backscatter radiofrequency (RF) signal [111]. As different
components generate distinguishable backscatter patterns, VH-IVUS is capable of
identifying fibrous, fibrofatty, necrotic core, and dense calcium components (Figure 1.6).
With the ability to identify these components, VH-IVUS has been applied to categorize
atherosclerotic plaques as either thin-cap fibroatheromas (TCFA, with confluent necrotic
core > 10% of cross section and in contact with the lumen, Figure 1.6a), thick-cap
fibroatheromas (ThCFA, large volume of extracellular lipid and debris surrounded by a
thick fibrous cap, Figure 1.6b), fibrotic (consisting mostly of fibrous tissue, Figure 1.6c),
or fibrocalcific (limited necrotic component with dense calcium, figure 1.6d) [112]. Other
more recent post-processing techniques include iMAP-IVUS, which uses pattern
recognition of backscatter from 40 MHz IVUS (Figure 1.6e,f), and integrated backscatter

\textbf{Figure 1.5:} Identification of plaque features via grayscale IVUS. (a) Image
demonstrating attenuation (asterisk) indicating large necrotic core. (b) Calcified
nodule (asterisk) and acoustic shadowing. (c) Small deposits or spotty
calcification (arrow head). (d) Multi-layer appearance correlated with plaque
progression and increased risk in heart transplant recipients. Adapted from
[108] with permission.
IVUS, a technique that applies a fast Fourier transform of the measured backscatter (Figure 1.6g,h) [113]. However, a major limiting factor for each of these techniques is the limited spatial resolution of IVUS, insufficient to accurately detect the thin fibrous cap of the so-called vulnerable plaque [114].

Optical coherence tomography (OCT) is another catheter-based imaging technique that utilizes infrared light (1280-1350 nm wavelengths) to determine coronary plaque characteristics. Due to the fact that red blood cells strongly absorb light at these wavelengths, imaging with OCT requires flushing of the lumen prior to data acquisition.
Acting as an optical analogue to IVUS, OCT measures backscatter of the IR light from different tissues and plaque components. However, unlike the sound waves used to generate a signal for IVUS, the signal of OCT travels at the speed of light. As a result, time-delayed detection of the “echo” cannot be utilized, instead employing interferometric techniques to analyze the backscattered light. Most commonly, this is achieved by sweeping a variable frequency light source across the vessel and processing the backscattered light with fast Fourier transform. OCT offers dramatically improved resolutions in both the axial (10-15 µm) and lateral (20-90 µm) directions when compared to IVUS [107, 115]. This allows for the highly detailed structural analysis of the superficial layers of atherosclerotic plaque that are beyond the resolution of IVUS.

Various features of plaque morphology have been identified using OCT, including fibrous cap thickness, calcification (characterized as regions of low reflectivity defined by sharp edges), and the presence of a lipid pool (regions of little or no backscatter and poorly defined edges) (Figure 1.7) [108, 116]. Critically, because of the high spatial resolution of OCT it has been demonstrated to have the ability to identify multiple features of plaque instability and rupture; fibrous cap erosion and ulceration can be visualized, while identification and classification of thrombus formation is also possible [117, 118]. Additionally, accurate measurement of plaque thickness is possible with OCT, providing a means to not only determine plaque vulnerability, but also allowing for disease progression monitoring and therapy efficacy analysis [119]. Like IVUS, a number of post-processing techniques have also been applied to OCT data to improve plaque characterization. In one example, the variance of the OCT signal is measured using normalized standard deviation, providing a means for the detection of macrophages
Granulometry is another technique that is capable of identifying specific structures from OCT images, and when combined with normalized standard deviation measurements, can distinguish the presence of foam cells within the vascular lesions [118]. Beyond imaging of plaque characteristics, OCT has also been utilized as a guide for percutaneous coronary intervention (PCI) by providing guidance for stent deployment, and has been shown to rival both IVUS and angiography in its efficacy [118, 120-122]. Emerging micro-OCT (µOCT) technology is pushing the resolution of OCT even further, providing spatial resolution down to 1 µM and providing imaging of cellular and subcellular structures associated with atherosclerosis progression [123]. However, the excellent spatial resolution of OCT comes at the expense of significantly decreased tissue penetration of only 1-3 mm [107]. Thus, while OCT allows for the highly sensitive analysis of the superficial plaque layers, it is unable to examine deeper components of the lesion. Additionally, the presence of superficial calcium deposits has been shown to interfere with the accurate determination of lipid content, generating a “shadowing” effect similar to that observed with IVUS [124].

Complementing the capability of IVUS and OCT to examine both the composition and highly-detailed surface characteristics of atherosclerotic plaques, recent
developments in near-infrared spectroscopy (NIRS) have provided yet another means for the identification of coronary atherosclerosis. Specifically, this relatively new imaging technique provides the determination of the chemical composition of plaques, generating important information on total lipid content [125, 126]. The technique of NIRS has been used extensively throughout the physical sciences as a means to determine a substance’s chemical composition based on its characteristic pattern of absorption of near-infrared (800-2500 nm) light. When applied to the detection of atherosclerosis, NIRS is utilized to detect the presence of lipid-rich areas within the vessel wall. The stark difference in chemical composition of these lipid-rich regions and the surrounding tissue are represented by a color-coded map referred to as a chemogram (Figure 1.8a) [108]. This two-dimensional map is a grid of pixels that represents the pullback position of the NIRS detector in millimeters on the x-axis and the circumferential position at that location in degrees along the y-axis, while the color of each pixel is representative of the probability of lipid present at that coordinate location [125, 127]. For a given segment of investigated artery, a parameter referred to as the lipid core burden index (LCBI) can be determined as the fraction of lipid-positive pixels on the chemogram of that segment multiplied by 1000 [125, 128]. This technique of identifying lipid-rich zones of artery segments has been paired with IVUS in a single catheter-based probe capable of measuring both plaque and lipid burden, improving overall plaque detection (Figure 1.8b,c) [129]. Although the ability of this hybrid IVUS-NIRS system is still under investigation, preliminary results suggest that it may be capable of identifying culprit lesions following ST-segment elevation myocardial infarction (STEMI) and that LCBI could be used to identify patients at an increased risk of periprocedural myocardial infarction [130, 131].
Finally, unlike the previous three techniques, fractional flow reserve (FFR) isn’t a genuine imaging technique. Instead, this assessment of atherosclerosis seeks to reveal the hemodynamic effect and physiological significance of plaque growth within a vessel. The basic principle behind FFR measurements is described by the hydraulic analogy of Ohm’s Law: when resistance is constant, changes in pressure are directly proportional to changes in flow [132]. Using coronary catheterization, a pressure sensor mounted on a 0.014 in. guidewire is directed to a distal location of a vessel containing the lesion of interest. Intravascular blood pressure is then measured on both the distal side ($P_d$) of the coronary plaque and in the aorta ($P_a$). FFR is then calculated as the ratio of these two pressures ($P_d/P_a$). Typically a value of 0.75-0.80 (a reduction of 20-25% blood flow) is

![Figure 1.8: Coronary atherosclerotic plaque analysis via near-infrared spectroscopy (NIRS) and a hybrid NIRS-IVUS probe. (a) 2-dimensional chemogram representing pullback position (in millimeters) of the probe on the x-axis and the circumferential position in degrees along the y-axis, with pixel color representing the probability of lipid present at the respective location. Colors range from red (low probability) to yellow (high probability). (b) Representative schematic of a combined NIRS-IVUS hybrid catheter tip. (c) Three cross-sections of pullback imaging superimposed with corresponding IVUS image. Figures in panels a and c adapted with permissions from [126]. Figure in panel b adapted with permissions from [109].](image)
considered to represent significant stenosis (Figure 1.9). To obtain reproducible values, these measurements are performed during hyperemia, typically via the injection of adenosine. This sort of functional assessment of plaque-induced ischemia has been shown to be beneficial in guiding revascularization decisions, ensuring that only hemodynamically significant lesions are treated [133]. Although FFR is necessarily an invasive technique, other non-invasive have been developed for functional imaging including single-photon emission computed tomography (SPECT) and positron emission tomography (PET) [69].

1.2.5 Molecular Imaging

Most of the imaging techniques discussed so far have focused on the identification and determination of anatomical or structural detail as it pertains to the progression of atherosclerosis and subsequent cardiovascular disease. However, as with the analysis of fractional flow reserve, there is a growing interest in modalities capable of evaluating physiologic and metabolic characteristics, including techniques such as positron emission

**Figure 1.9:** Determination of fractional flow reserve (FFR) in (a) a healthy artery and (b) across an atherosclerotic lesion. Pressure is measured both distal to the lesion (P_d) as well as in the aorta (P_a). The FFR is calculated as the ratio of these values (P_d/P_a), and provides a means to determine flow reduction caused by the lesion. FFR values lower than 0.75-0.8 typically considered representative of significant stenosis.
tomography (PET) and single-photon emission computed tomography (SPECT). In contrast to the classic angiographic techniques that focus on the late stages of atherosclerosis progression, molecular imaging techniques such as these provide the opportunity to investigate the very earliest changes associated with the disease. Both PET and SPECT rely on specifically designed molecular probes that have been labeled with a radiotracer. The primary difference in their method of detection is that tracers used in SPECT (common examples include $^{99}$Tc, $^{111}$In, and $^{67}$Ga) emit gamma ($\gamma$) radiation that is detected directly, whereas PET tracers (commonly $^{18}$F, $^{82}$Rb, $^{13}$NH$_3$, and $^{15}$OH$_2$) emit positrons that annihilate with nearby electrons and subsequently release $\gamma$-radiation. Because of this difference, the effective spatial resolutions of these two techniques are disparate; PET imaging has reported spatial resolutions of approximately 1-2 mm, while the spatial resolution obtained by SPECT has been reported between 2-10 mm [69, 134]. Despite these low spatial resolutions, the overwhelming advantage of these systems arises from their dramatically improved sensitivity to molecular signals, capable of detecting even picomolar concentrations of reporters [134]. These imaging systems have allowed for the direct observation of inflammation, microcalcification, hypoxia, and neovessel angiogenesis, providing insights on the individual patient level that traditional structural imaging provides on the population level.

One of the most widely explored applications of nuclear molecular imaging in the detection of atherosclerosis has been through the use of the glucose analogue 2-(18F)-fluoro-2-deoxy-d-glucose (18FDG). Because of its similarity in structure to glucose, 18FDG is readily taken up by metabolically active cells through glucose transporters such as GLUT1 [135]. Once present inside the cell, 18FDG is phosphorylated, but cannot be
processed further via the glycolytic pathway due to the lack of 2’-hydroxyl leading to an accumulation in cells proportionate to their metabolic activity. $^{18}$FDG accumulation has thus been shown to correlate with macrophage density in *ex vivo* histological samples due to the elevated metabolic demand of these cells compared to surrounding tissues. Additionally, this relation has been utilized for the *in vivo* identification of carotid atherosclerotic plaques (Figure 1.10) [136, 137]. A major challenge for $^{18}$FDG imaging in the coronary arteries is that the myocardium is also an extremely metabolically active tissue, and thus generates a high background that can obscure signal produced within the small overlaying arteries [138]. This background interference has limited much of the work with $^{18}$FDG to a focus on the carotid arteries. As a result, a variety of other radiotracers have been explored for their ability to specifically identify atherosclerotic plaques within the coronary arteries. One such tracer is $^{18}$F-sodium fluoride, which has previously been utilized to identify bone metastases based on its ability to be deposited on hydroxyapatite. PET imaging of $^{18}$F-sodium fluoride has demonstrated that this tracer

![Figure 1.10](image1.png)

*Figure 1.10*: Molecular imaging of coronary atherosclerosis using PET. (a) Invasive coronary angiogram of patient with acute ST-segment elevation myocardial infarction showing proximal occlusion (red arrow) of the LAD. (b) $^{18}$F-fluoride uptake (red arrow) at the site of the atherosclerotic plaque. (c) $^{18}$FDG imaging of the same segment, demonstrating high background signal from both the myocardium (yellow arrow) and the esophagus (blue arrow). Figure adapted with permissions from [138].
colocalizes to microcalcification in atherosclerotic plaques, and can be used as a means to identify high-risk plaques [139]. Use of this detection concept is currently under investigation in the PREFFIR study, and aims to determine if $^{18}$F-sodium fluoride imaging confers prognostic value to patient outcome [140, 141]. Other tracers used in PET molecular imaging include $^{68}$Ga-DOTATATE targeting somatostatin receptors expressed by activated macrophages, and associated with plaque vulnerability; $^{11}$C-PK11195 targeting mitochondrial receptors involved in cholesterol transport and overexpressed in activated macrophages; and $^{18}$F-labeled peptides with an affinity for the integrin $\alpha_v\beta_3$ receptor as a means to target neoangiogenesis [142]. As additional novel tracers become available, the capabilities for PET imaging to improve atherosclerosis prognostication will inevitably improve. However, this technique presents inherent unique challenges, including the relatively short half-life of most of its positron-emitting tracers ($^{18}$F ~110 minutes, $^{11}$C ~20 minutes). Consequently, PET imaging requires an on-site cyclotron and trained specialists for on-demand tracer production. In light of the challenges presented by PET, SPECT has also been utilized for the detection and imaging of atherosclerotic plaques. Like PET, tracers targeting the integrin $\alpha_v\beta_3$ receptor have been labeled with the SPECT reporter $^{125}$I for imaging of neoangiogenesis [134]. SPECT has also been used for detection of healthy macrophages, VCAM-1 expression, and apoptotic macrophages; each of these representing important components of the progressing atherosclerotic plaque [69]. Importantly, tracers used in SPECT have markedly longer half-lives than those used in PET ($^{99}$Tc ~6 hours, $^{67}$Ga ~3.3 days, $^{125}$I ~59 days), meaning the additional on-site resources for tracer production required by PET are not needed. Current-generation technology incorporating cadmium-zinc-telluride
detectors has also improved the spatial resolution of SPECT, not only improving the image quality generated by this technique but also providing the potential to reduce the amount of tracer required, and subsequently the overall required radiation exposure [143].

1.2.6 Hybrid Technology and Image-based Computational Models

Every imaging technique for the assessment of atherosclerotic plaque progression presents unique strengths and weaknesses; invasive vs. non-invasive procedures, the requirement and dose of radiation exposure, and limitations on spatial or temporal resolution among others. Thus, it is unsurprising that a number of hybrid technologies have been developed to take advantage of the unique traits of multiple complementary imaging modalities. For example, although the molecular imaging techniques of PET and SPECT provide unique individual insights into the underlying biochemistry of atherosclerotic plaque progression, their limited spatial resolution requires that in current clinical applications their scans be coregistered with an additional imaging modality to accurately localize the detected signal to specific anatomic locations. Most commonly, these hybrid imaging platforms utilize CT, although recent advances have introduced hybrid imaging platforms of both PET and SPECT with MRI as well. PET-CT has been well established in clinical research, serving as a major imaging modality for the evaluation of atherosclerosis in the carotid arteries and aorta utilizing $^{18}$FDG. This hybrid imaging system has demonstrated the capability to identify culprit lesions following stroke and can detect increased signal following myocardial infarction [137, 144]. In another application, PET was utilized to determine myocardial perfusion based on the uptake of $^{18}$FDG, and these images were subsequently combined with stenosis assessment
provided by CT [145]. In this manner, the ischemic effect of the individually identified plaques could be determined, an outcome neither modality could provide individually. Hybrid techniques have also found abundant use in small animal models of atherosclerosis, where the underlying biochemical and metabolic processes of plaque formation can be investigated. Using a microSPECT-CT imaging system, MMPs involved in vascular remodeling were imaged, providing insight into the mechanism by which the vasculature responds to atherosclerotic insult [146]. As each of these techniques require ionizing radiation, a growing concern surrounding both PET-CT and SPECT-CT involves the increased radiation dosages required for combined imaging analysis. However, technologies and imaging parameters are perpetually under revision to reduce both radiation requirements and exposure. More recently, hybrid imaging systems utilizing the non-ionizing modality of MRI have emerged. Although these technologies are predominantly pre-clinical, PET-MRI has been utilized successfully to identify carotid atherosclerotic plaque with good agreement with PET-CT imaging of the same patients [147]. Numerous invasive hybrid techniques (such as IVUS-NIRS discussed previously) have also arisen in an attempt to better identify and characterize atherosclerotic plaques. Early work with hybrid IVUS imaging utilized x-ray angiography, and has since given rise to the ability to generate 3-dimensional models of coronary artery lumen upon which IVUS imaging can be superimposed [148]. Expanding on this technique, IVUS-NIRS has also been combined with angiography, providing models with the potential to investigate hemodynamic effects on plaque composition and geometry [149]. One of the more promising advances in intravascular hybrid imaging has been the combination of IVUS and OCT due to their complementary strengths. The
superior penetration (but limited spatial resolution) of IVUS allows for examination of positive remodeling and analysis of the necrotic core, whereas the excellent resolution (but drastically reduced penetration) of OCT provides a means to investigate the overlying fibrous cap thickness, presence of thrombus, and detection of rupture or erosion (Figure 1.11) [150, 151]. Although this IVUS-OCT hybrid technique has not yet found adoption in clinical practice due to a number of limiting factors including low image quality, long image times, and large catheter sizes, there will undoubtedly be continued interest in improving this technique. Various other IVUS-hybrid techniques (including combinations of near-infrared fluorescence spectroscopy [NIRF], time-resolved fluorescence spectroscopy [TRFS], and intravascular photoacoustic [IVPA] imaging) are also in development, each providing its own unique advantages for plaque identification and assessment [152-154].

Finally, an important field of research providing insights to the understanding of coronary atherosclerosis progression both in populations as well as individuals is that of computational modeling. There exists an immensely diverse range of computational models for the analysis of atherosclerosis, each varying in approach as well as desired outcomes. Purely mathematical models, such as those dependent on partial differential

![Figure 1.11: Hybrid IVUS-OCT imaging. (a) IVUS imaging indicating presence of calcifications and acoustic shadowing (yellow arcs). (b) OCT imaging indicating fibrous cap (blue arrow) and underlying detail. Adapted with permission from [151].](image-url)
equations and process algebras for complex pathway systems, are utilized to investigate various important processes of plaque formation such as LDL or macrophage infiltration into the intima, or the interaction of variable forces such as ESS with plaque caps to investigate stability and rupture [155-157]. In other approaches, predictive population models may be developed, utilizing data obtained in large clinical studies to generate simulated populations that exhibit the observed rates of formation and progression of atherosclerotic disease. Through manipulation of the governing equations describing disease incidence and progression rates, the efficacy of various therapeutic or screening strategies can be explored rapidly and in an iterative manner to aide in clinical decision making for a given population [158]. Additionally, predictive models of this nature also possess the capability of determining previously unexamined interactions between known or suspected risk factors of disease to provide improved risk stratification of the population being modeled based on their known demographics and disease risk characteristics. Largely driven by improvements to previously discussed imaging technologies, image-based computational models have also seen complementary improvements in their ability to simulate various aspects of atherosclerotic disease. These platforms, utilizing the generated geometric and morphologic information of the vessel and plaque, supplement existing imaging modalities by modeling and calculating the mechanical stresses within and around the lesion: a functionality that medical imaging is incapable of providing alone.

In order to generate appropriate computational models based on acquired image data, a number of considerations must be made concerning (a) fluid behavior, (b) structural characteristics, and (c) the interaction of these two parameters. It was already
noted that the shear force exerted on the wall of a vessel (the ESS) by flowing blood is an important factor in determining the formation and progression of atherosclerotic plaques. To accurately model blood flow, computational fluid dynamics (CFD) is utilized, providing a means to solve the governing equations of fluid behavior for the system. In the case of blood flow, most computational models will operate with the assumption that blood is a homogenous, incompressible fluid, allowing for the use of the Navier-Stokes equations and approximating blood as a Newtonian fluid (although its granularity and shear-thinning characteristics are known to deviate from this ideal case) [155]. As with many operations involved in the development of computational models, the underlying mathematics involved are beyond the scope of the presented review, but are well established and detailed elsewhere [159, 160]. The next step in the design of the model is to define the structures (vessel wall, lumen, plaque components, thrombus, etc.) that will be examined. Based on the strengths of the various imaging techniques discussed, 2-dimensional (2D) or three-dimensional (3D) geometries can be obtained from imaging using invasive (IVUS, OCT, etc.) or non-invasive (CCTA, MRI, etc.) techniques (Figure 1.12a). Ideally, the structures would be identified with high-resolution and substantial contrast between various tissues of interest. This contrast serves to allow for segmentation of the structures, identifying boundaries between different components (Figure 1.12b). For example, in models examining blood flow dynamics, boundaries for the lumen of the vessel would need to be identified, while examinations of plaque stability would need to be able to clearly identify components such as the fibrous cap, necrotic core, and the presence of calcification. These segmentation methods can either
be manual, or they can rely on a various number of algorithms to determine the desired boundaries based on preset contrast values or ‘training’ with representative images [161]. Additional considerations must be made for models requiring either one wall condition (such as the inner-wall of the vessel for examining blood flow) or multiple wall boundaries (for examination of maximum plaque wall stress and rupture predictions) [162, 163]. The final step in preparing the structural/fluid data for the computational model is to generate what is known as a “mesh” of the segmented geometries (Figure 1.12c). These meshes are generally composed of simple geometric shapes, with triangles

**Figure 1.12**: Stages of computational model development for analysis of a coronary artery. (a) Angiogram of the RCA. (b) Segmentation of the RCA and identification of vessel walls. (c) Surface and volumetric meshes fit to the segmented vessel. (d) Physiologic parameters utilized for the generation of boundary conditions such as blood pressure and flow rates. (e) Iterative simulation of the model and results output demonstrating *in silico* determination of FFR. (f) Validation of model results to standard or known measurements. Adapted from [166] with permission.
and tetrahedrons common for 2D and 3D meshes, respectively. Again, the mathematics and algorithms used to determine the process by which the mesh is generated are beyond the scope of this review, but are varied and diverse [164]. However, it must be noted that an increase in complexity of the mesh (generally speaking, the total number of mesh elements) is largely determinant for the model accuracy, but also increases required processing time. Thus, depending on the goal of the model, reductions in complexity may serve to improve overall analysis. For example, both 2D and 3D models of arterial bifurcations as well as linear vessels have been used successfully for the analysis of plaque stability and luminal stenosis caused by an atherosclerotic plaque, although overall model complexity varies [165]. Finally, information regarding boundary conditions must be established. Because no model is capable of examining the entire cardiovascular system at once, parameters for the hemodynamic or physiologic conditions must be set for the inlet/outlet of the model, as well as at the vessel walls (Figure 1.12d) [166]. Although highly accurate boundary conditions previously required invasive catheter-based measurements of blood flow, recent work has shown that mathematical scaling of these values based on vessel diameter generate similar values [167]. At this point, simulation of the model will be performed over a number of successive iterations to calculate the desired solution (hemodynamic flow parameters, fluid wall stress, intraplaque stress, etc., Figure 12.e) that can be further validated through comparison to standard or known values (Figure 12.f). However, a number of practical considerations must be made to generate meaningful data based on clinical imaging of atherosclerosis. Specifically, if both fluid motion and wall parameters are considered, a comprehensive Fluid-Solid Interaction (FSI) model is utilized. In this type of analysis,
boundary conditions dictate the direction and velocity of the blood flow through the two-or three-dimensional model, and are generally more complex than those of CFD models. Subsequently, FSI models will also require greater computational cost; exceeding a processing time more than ten-fold over CFD models in some situations [166]. Additionally, because many computational models examining atherosclerosis are focused on determining wall and intraplaque stresses, considerations must be made within the model for the residual stress present in the system. Blood flow through the vasculature is pulsatile in nature due to the cardiac cycle. Although luminal pressures vary with time, vessel tissues (and plaque components) are always under some level of physical stress from sources such as blood pressure and vessel tethering. To determine the magnitudes of these residual stresses within the system, algorithms have been developed to first shrink the measured vessel geometry to a stress-free state, and then determine the relative amounts of stretch and pressure required to return the geometry to its original conformation [168]. In comparison, an inverse – or backward incremental – method has been developed which requires no manual input to determine residual strain values and closely resembles the geometry determined via imaging [169]. Despite these seemingly stringent requirements, computational modeling continues to be a valuable resource for the analysis of atherosclerotic disease, providing a means to determine processes such as LDL infiltration and oxidation, macrophage recruitment, vascular smooth muscle cell proliferation, as well as the prediction of sites of plaque rupture based on patient-specific imaging data [161, 165, 168, 170].

As technology continues to move forward, improved diagnostic and imaging techniques will inevitably give rise to better modeling capabilities. Conversely, as patient
specific modeling becomes more robust, the prognostic value of various imaging
techniques will subsequently improve. Already, imaging techniques for the direct
measurement of hemodynamic flow conditions are being employed. For example, CCTA
has been shown to be capable of determining FFR across an atherosclerotic plaque in a
technique referred to as FFR$_{CT}$ [171, 172]. In this approach, the process of computational
fluid dynamics is applied to CCTA images, providing hemodynamic parameters of
identified stenoses in a completely non-invasive manner. Moreover, it has been shown
that the sensitivity and specificity of this technique to identify FFR $\leq 0.8$ (as confirmed
by invasive FFR) were 90% and 54%, respectively [173]. More importantly, through the
use of this technique, it was shown that the number of patients undergoing invasive
coronary angiography with non-obstructive disease was reduced, identifying individuals
for whom this invasive procedure simply wasn’t required [174]. Another imaging
technique investigating hemodynamic parameters is the so-called ‘time-resolved three-
dimensional phase contrast MRI with three directional velocity encoding,’ or simply ‘4D
flow MRI’ [175]. This technique is, in essence, an extension of an older MRI approach
known as phase contrast (PC) MRI. By applying a bipolar gradient in a single direction of
flow (x, y, or z in the three-dimensional coordinate plane), the velocity of the observed
contrast agent can be determined as a function of its phase change following the gradient;
stationary nuclear spins will exhibit no net phase change, while nuclei in motion will
result in a detected phase change that is proportional to their velocity. Using ECG-gating
for image acquisition and bipolar gradients applied in all three planes, the time-resolved
motion of the contrast agent can be determined through post-image processing. 4D flow
MRI has found application in a number of arterial regions, providing information on
intracranial hemodynamics, assessment of ESS in the thoracic aorta, and imaging of blood flow in the heart. While determination of coronary FFR isn’t yet possible with this technique, experimental protocols are investigating this possibility [175-177]. These novel approaches of atherosclerosis analysis demonstrate how advances in imaging technology and techniques may provide a more complete understanding of disease, resulting in the formation of more comprehensive models for in silico analysis.

1.3 Closing Remarks

Given the complexity of the formation and progression of atherosclerotic plaques, it is unsurprising that an enormous diversity of techniques has been – and will continue to be – developed for the imaging of this disease. It follows that the manner in which data generated by these imaging techniques will also continue to grow in its intricacy and level of comprehensive analysis of atherosclerotic plaque initiation, growth, and clinically relevant features such as stenosis and rupture. An important consideration as both researchers and clinicians move forward in their understanding of the underlying mechanics and causes of atherosclerosis will be the development of a broad spectrum of additional techniques and approaches built upon these technologies that are capable of providing meaningful insight. Attempts must be made for the identification of important pathways and biochemical reactions governing the opposing forces of disease progression and lesion repair. There exists room for improvements in the specificity and selectivity of every imaging modality, as well as their respective parameters of spatial and temporal resolution, early disease identification, and prognostic value. Above all, these individual improvements must not be considered in vacuum, but must necessarily be considered as single components of a complex whole, joining imaging techniques, biochemical
pathway elucidation, and computational modeling (both image-based and purely mathematical) to provide better treatment for those currently exhibiting clinical disease and better prevention for future generations. In the following chapters, examples of additional components that contribute to this whole will be described. Specifically, two distinct mathematical models examining sudden cardiac death and CVD risk factors will be discussed, followed by the description of a novel biosensor for the detection of apoptosis and its potential applications to the detection of atherosclerotic plaques.
Chapter 2

Discrete Event Simulation Model of Sudden Cardiac Death

2.1 Perspective

The formation and growth of atherosclerotic material within the walls of arterial vessels is a common and prevalent cause of cardiovascular disease (CVD), and accounts for 80% of sudden cardiac death (SCD) [178]. Annually, the cost of CVD exceeds $286 billion in the United States, while SCD alone is accountable for 180,000-450,000 deaths [26, 179]. Though numerous approaches have been implemented to prevent SCD through the treatment of CVD, the use of simulation modeling has been limited to applications such as the cost of anti-CVD treatment [180]. In this chapter, a discrete event simulation (DES) model is described that has been designed to predict the occurrence of SCD within a simulated population, allowing for predictive means of intervention as well as treatment to reduce both the incidence and cost associated with SCD. The use of this DES approach allows for the modeling of the complicated chains of random events within individuals or within a population. This is, thus, beneficial for the description of disease evolution. Additionally, the extensive available data from epidemiological [26, 178-182] and pathological [65, 183-185] studies allow for highly accurate refinement and verification of the model results. The model parameters chosen to fit these data, including frequency and distribution of atherosclerotic lesions (atheromas), were drawn from a pathological study [185], while data on the age of individuals at the time of first atheroma formation were drawn from a comprehensive epidemiological study [182]. Predicted estimations of annual SCD risk generated by the model were then compared with the available
epidemiological data. Finally, predictors of SCD were investigated, and the impact on life expectancy was modeled for simulated patients provided two types of interventions: (i) non-invasive methodologies for highly specific and selective identification of coronary atheromas, coupled with atherosclerosis treatment in a certain ‘at risk’ group of the population, and (ii) highly efficient medications which decrease the growth rate of atheromas and clots, administered to the whole population after a certain age, and predicted to reduce the number of SCDs up to 8-fold. An important advantage of this model is its ability to predict the optimal time window of the intervention. Another advantage of the model is its ability to integrate data from pathological, clinical, and epidemiological studies. If implemented in a clinical setting, predictions on treatment and intervention from this model would not only reduce CVD-associated costs, they could also increase the average life spans of given populations and, more critically, save lives.

2.2 Results

2.2.1 Brief Model Description

Following the data from the pathological studies and the simplified scheme for classification of atherosclerotic lesions [65, 184, 185], it was assumed that SCD caused by coronary artery disease (CAD) could occur due to three major chains of events: (i) thin-cap fibrous atheroma formation, growth, rupture, clot formation, clot growth, thrombosis, arrhythmia, SCD; (ii) medium-cap fibrous atheroma formation, growth, erosion, clot formation, clot growth, thrombosis, arrhythmia, SCD; (iii) thick-cap atheroma formation, growth, stenosis, arrhythmia, SCD.

The developed software, *SCD-Predict*, simulates the progression of atherosclerosis in an individual, starting with the formation of the first atheroma and
ending either with SCD or with the simulation stop time. This final event can represent either the end of the clinical study or a non-cardiac death, depending on the objective of the simulation. To model a population or a cohort of such individuals, *SCD-Predict* is queried multiple times from the parent program specifying the properties of the population.

### 2.2.2 Comparison of the Model to Pathological Data

*SCD-Predict* was fit with the data drawn from the pathological study of 50 whole human hearts as previously noted [185]. Patients from which these hearts were obtained had died of cardiovascular (n = 33), non-cardiovascular (n = 13), and unknown (n = 4) causes. This study reported the number of ruptured and non-ruptured thin-, medium-, and thick-cap atheromas found in the coronary arteries of these hearts. Table 2.1 presents comparisons of the simulation results with these observations. Exact values of all simulation parameters can be found in Table A1 in Appendix A. To obtain this comparison, *SCD-Predict* analyzed 100 simulated men and 100 ‘simulated women. Individuals in this simulated cohort were either modeled to have died of SCD or determined to have died of non-SCD by reaching the simulation stop time. This endpoint was modeled as a normal distribution with an average 72.8±8 years, simulating the age distribution observed in the pathological study with which the simulated results are compared [185]. As seen in Table 2.1, the simulated numbers of ruptured, non-ruptured, thin-cap, medium-cap, and thick-cap atheromas are in good agreement with the numbers observed for both men (n = 32) and women (n = 18), as well as the total group in the study (root mean squared error of 13.5% for men, 21.8% for women, and 8.6% for the total group). It is, however, important to note that while some parameter fitting was
performed, minimization of the mean error was not considered a major goal. While designing the model, excessive fitting of parameters was avoided, instead keeping parameter values of initial atheroma size, time to rupture, and degree of stenosis consistent with the body of publications [65, 182, 184-192]. Detailed explanations on the rationale of parameter selection can be found in Section 2.3 (Methods). As an illustration of this aversion to over-fitting, the parameters of the model were not changed when simulating men and women, instead using the same parameters in Table A1 for both cohorts. As a result, the only difference in modeling male and female cohorts was found in the normal age distributions of first atheroma formation, with a mean age of 55±12 years for men and 61±12 years for women. The rationale for changing only initial distributions and preserving the model parameters is that the protection against atherosclerosis is known to be lost within a 5-year period post-menopause [182]. Thus, a model with identical parameters and a time shift of 6 years should provide an accurate simulation for both men and women, as demonstrated by Table 2.1.

Additionally, Table 2.1 contains the results for both men and women who died of SCD and non-SCD cases separately. Here, the comparison with the findings from the pathological analysis on which the model is based becomes complicated. In the pathological study, the group of individuals having died from cardiovascular causes includes all cases of CVD-related deaths such as myocardial infarction or cardiomyopathy. In comparison, the SCD group from SCD-Predict contained only those who had died of SCD caused by CAD with subsequent arrhythmias. This explains why the mean numbers of atheromas in the simulated SCD-Predict group are higher than those observed in the corresponding group from the pathological study. However, there
remains good agreement between the numbers of atherosclerotic lesions and their morphologies for the two groups (the corresponding numbers for the non-SCD and non-CVD groups have an even better correlation).

**Table 2.1**: Comparison of the mean numbers of ruptured and non-ruptured fibrous cap atheromas simulated with the *SCD-Predict* model (sim) and observed in the pathological study from which data used to fit the model was obtained [185]. Weighted average (w.a.) totals were calculated as 0.64(value for men) + 0.36(value for women) to reflect the gender distribution observed in the pathological study.

<table>
<thead>
<tr>
<th>Patients</th>
<th>Atheroma</th>
<th>Ruptured</th>
<th>Thin</th>
<th>Mid</th>
<th>Thick</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (sim)</td>
<td>100</td>
<td>0.43</td>
<td>0.65</td>
<td>0.50</td>
<td>2.52</td>
<td>4.10</td>
</tr>
<tr>
<td>Men [185]</td>
<td>32</td>
<td>0.44</td>
<td>0.63</td>
<td>0.47</td>
<td>3.59</td>
<td>5.13</td>
</tr>
<tr>
<td>Women (sim)</td>
<td>100</td>
<td>0.27</td>
<td>0.24</td>
<td>0.35</td>
<td>1.9</td>
<td>2.76</td>
</tr>
<tr>
<td>Women [185]</td>
<td>18</td>
<td>0.28</td>
<td>0.17</td>
<td>0.44</td>
<td>1.78</td>
<td>2.67</td>
</tr>
<tr>
<td>Total (w.a.)</td>
<td>--</td>
<td>0.37</td>
<td>0.50</td>
<td>0.44</td>
<td>2.27</td>
<td>3.58</td>
</tr>
<tr>
<td>Total [185]</td>
<td>50</td>
<td>0.38</td>
<td>0.46</td>
<td>0.46</td>
<td>2.94</td>
<td>4.24</td>
</tr>
<tr>
<td>Men SCD (sim)</td>
<td>14</td>
<td>1.00</td>
<td>0.71</td>
<td>1.14</td>
<td>3.79</td>
<td>6.64</td>
</tr>
<tr>
<td>Men nonSCD (sim)</td>
<td>86</td>
<td>0.34</td>
<td>0.64</td>
<td>0.40</td>
<td>2.31</td>
<td>3.69</td>
</tr>
<tr>
<td>Women SCD (sim)</td>
<td>8</td>
<td>0.88</td>
<td>0.36</td>
<td>1.00</td>
<td>3.62</td>
<td>5.86</td>
</tr>
<tr>
<td>Women nonSCD (sim)</td>
<td>92</td>
<td>0.22</td>
<td>0.23</td>
<td>0.29</td>
<td>1.75</td>
<td>2.49</td>
</tr>
<tr>
<td>SCD (w.a.)</td>
<td>--</td>
<td>0.95</td>
<td>0.58</td>
<td>1.08</td>
<td>3.68</td>
<td>6.29</td>
</tr>
<tr>
<td>CV [185]</td>
<td>33</td>
<td>0.45</td>
<td>0.55</td>
<td>0.55</td>
<td>3.12</td>
<td>4.67</td>
</tr>
<tr>
<td>nonSCD (w.a.)</td>
<td>--</td>
<td>0.29</td>
<td>0.49</td>
<td>0.36</td>
<td>2.08</td>
<td>3.22</td>
</tr>
<tr>
<td>nonCV [185]</td>
<td>13</td>
<td>0.31</td>
<td>0.31</td>
<td>0.23</td>
<td>2.23</td>
<td>3.08</td>
</tr>
</tbody>
</table>

Morphological characteristics of simulated plaques from *SCD-Predict* were also compared to an additional pathologic study [183] in which it was observed that 73% of
SCD cases involved plaque ruptures associated with thrombosis, 8% involved plaque fissures with intra-plaque fibrin deposition and hemorrhage, and 19% possessed no evidence of thrombus formation. Table 2.2 compares these results with those of SCD-
Predict, demonstrating that the percentages observed in this pathological study were represented closely to the average between the men and women simulated by our model. As shown, the average percentage of ruptured plaques predicted by SCD-Predict is also 73%, while the percentage of cases possessing thrombus with no rupture is 11%, and the absence of thrombus predicted in 16% of cases. Importantly, while there is a high similarity between the simulated and observed percentages, no parameter fitting to the data of this additional pathological study was performed. This similarity of the independently generated and observed data thus confirms the validity of the model.

Table 2.2: Comparison of the percentage of ruptures, thrombi, and absence of thrombi in atherosclerotic plaques of coronary arteries simulated with SCD-Predict and observations of an additional pathological study [183].

<table>
<thead>
<tr>
<th></th>
<th>Rupture Present</th>
<th>Thrombus, No Rupture</th>
<th>No Thrombus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (sim)</td>
<td>71%</td>
<td>22%</td>
<td>7%</td>
</tr>
<tr>
<td>Women (sim)</td>
<td>75%</td>
<td>0</td>
<td>25%</td>
</tr>
<tr>
<td>Total [183]</td>
<td>73%</td>
<td>8%</td>
<td>19%</td>
</tr>
</tbody>
</table>

Figure 2.1 provides a graphical summary of the simulated patient data via histograms of the distributions of total atheromas, vulnerable atheromas, arrhythmias, and age at death for the aforementioned simulated cohorts. While the maximum number of atheromas (10), vulnerable atheromas (4), and arrhythmias (3) are the same in both
simulated men and simulated women, the mean numbers are substantially higher for men than women, as expected and reflected by both Table 2.1 and Figure 2.1

Figure 2.1: Using SCD-Predict, cohorts of 100 men (blue) and 100 women (red) were simulated and analyzed. Visualization of the findings include (a) simulated age at time of death, (b) total number of atheroma, (c) total vulnerable atheroma, and (d) number of arrhythmias.

2.2.3 Model Predictions of Cumulative and Annual SCD Risk

Because of the nature of the SCD-Predict model and its demonstrated coherence to known pathological and epidemiological data, it is possible to use it as a predictive tool. First, the model was used to predict the age-dependent cumulative risk of SCD in the general population. To accomplish this, multiple cohorts of 1000 simulated persons were simulated using a simulation stop time that was set at various points between 45 and 95 years of age. The number of persons dead due to SCD in each simulated cohort divided by the total number of persons yielded the simulated cumulative risk of SCD (Figure 2.2a). After curve fitting these data points (MATLAB, Curve Fitting Tool), the time derivatives of the fits were calculated. This first derivative of the cumulative risk represents the annual (instantaneous) risk of SCD (Figure 2.2b). Interestingly, the peaks
of annual risk for SCD at age 80 are predicted by *SCD-Predict* to be 3.3% and 1.8% for men and women, respectively. This finding is in agreement with the published observations of the age-dependent incidence of SCD [181], where the peak incidence of SCD is observed for the 75-84 year old age group. The model predictions are also in agreement with the findings of annual incidence of death from coronary heart disease (CHD) in the Framingham study, where 30 CHD deaths per 1,000 men (3% annual risk) and 15 CHD deaths per 1,000 women (1.5% annual risk) were reported for the same age group [193]. Importantly, none of the epidemiological data from either the Framingham study or from the aforementioned epidemiological study was used in the development of the *SCD-Predict* model or in the parameter fitting process. This convergence of independent analysis from simulation and observation provides additional confirmation of the validity of our model.

![Figure 2.2:](image)

**Figure 2.2:** Graphical representations of the (a) cumulative and (b) annual risk of SCD for both men (blue) and women (red). Each point in the cumulative risk curves was generated by running *SCD-Predict* for the cohort of 1,000 simulated persons with fixed simulation stop times. Annual risk curves were generated by calculating the derivative of the cumulative risk. Vertical lines indicate the cumulative or annual risk of SCD at 80 years of age.

The detailed quantitative comparison, however, is complicated by the fact that these epidemiological studies present the combined data for SCD from the entire spectrum of etiologies: CAD (~80%), cardiomyopathies (10-15%), and others including congenital cardiac conditions, Brugada, and long QT syndromes (5-10%) [181]. While
the current model only represents the most common etiology of SCD, i.e. CAD, other etiologies could be introduced in future generations of SCD-Predict.

2.2.4 Model Predictions of Coronary Screening and Treatment

Due to the strong adherence of the simulated data to values observed in large-scale pathological and epidemiological studies on the rates of formation and progression of atherosclerotic plaques and the incidence and prevalence of disease, it was determined that the model could be applied to estimate the impact of early diagnosis and treatment of atherosclerosis. To accomplish this, an initial cohort of 1,000 men and 1,000 women were simulated to age 45 and 55 respectively. At this time, a screening strategy was implemented such that individuals presenting more than five atheromas were treated and saved from SCD. This strategy of screening at a young age for those individuals with an elevated number of atherosclerotic plaques provided dramatic results for the individuals identified. Eight men were identified who would have otherwise died between the ages of 45.2 and 66.1, whereas 11 women were identified who would have otherwise died between the ages of 55.2 and 77.8.\(^1\) Although this screening strategy resulted in the ability to add over 20 years of life to the identified individuals, the percentage of the population identified by this approach remained marginal (a total of 21 individuals met the screening criteria, or only 1.05% of the simulated population). Thus, the effect of varying the screening parameters for both scan age and threshold for number of atherosclerotic plaques was investigated. For both men and women, three intervention strategies were employed: scan at a young age (45 for men, 55 for women) with a low (1 atheroma) or high (4 atheroma) number of atherosclerotic plaques, and scan at an

---

\(^1\) Note that two additional women were identified using this screening strategy for a total of 13 identified women. However, these two would not have died of SCD regardless of intervention, and thus were not considered to be saved.
increased age (60 for men, 65 for women) with an intermediate number of atheroma (3 for men, 2 for women). Figure 2.3 presents survival curves representing the impact of these various screening strategies. Looking only at the population of men (Figure 2.3a) and women (Figure 2.3d) meeting the intervention parameters, a number of trends can be observed. For each of these strategies, survival curves are plotted for the 1,000-member cohort both with (solid line) and without (dashed line) intervention. As demonstrated, the average increase in life duration due to intervention is maximized for both genders in the case of a young scan age and a high threshold of atheromas. However, as shown previously, the total number of individuals identified by this strategy represents only about 2-5% of the population (between 20-50 identified individuals). Intuitively, utilizing a lower threshold for the required number of atheroma or an increased scan age would identify a greater percentage of the population, as is observed for both men and women. Although the average increase in life per individual (horizontal distance between the curves) is reduced utilizing this strategy, the total number of years added to the identified subpopulation (total area between the curves) is increased compared to a young scan age and high threshold of atheroma. This effect is most apparent when comparing the effects of the various intervention screening strategies on the survival curves of the entire simulated populations. For both men (Figure 2.3b) and women (Figure 2.3e), using an older scan age or a lower threshold for the number of atheromas results in a substantial increase in survival for the total population. Table 2.3 provides a summary of the impact from the three screening strategies on the increases in average life expectancy for the total population as well as identified subpopulations for both men and women.
Figure 2.3: Survival curves illustrating the impact of various screening strategies with parameters of both scan age and threshold of atheromas varied. (a) Impacts of screening strategies showing survival curves for identified male subpopulations both with (solid lines) and without (dashed lines) screening. (b) Effects of screening strategies showing the survival curves for the entire male population. (c) Focus region of the survival curves presented for the entire male population; dashed line indicates survival of half of the simulated cohort. (d) Impacts of screening strategies showing survival curves for identified female subpopulations both with (solid lines) and without (dashed lines) screening. (e) Effects of screening strategies showing the survival curves for the entire female population. (f) Focus region of the survival curves presented for the entire female population; dashed line indicates survival of half of the simulated cohort.
Table 2.3: Increases in average life expectancy for total simulated populations as well as subpopulations identified using three different screening strategies.

<table>
<thead>
<tr>
<th>Years Added</th>
<th>Young Scan Age, High Threshold</th>
<th>Young Scan Age, Low Threshold</th>
<th>Older Scan Age, Intermediate Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (Total)</td>
<td>0.5</td>
<td>3.2</td>
<td>3.3</td>
</tr>
<tr>
<td>Men (Identified)</td>
<td>23.4</td>
<td>19.1</td>
<td>16.2</td>
</tr>
<tr>
<td>Women (Total)</td>
<td>0.7</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Women (Identified)</td>
<td>17.1</td>
<td>13.6</td>
<td>12.5</td>
</tr>
</tbody>
</table>

Expanding the number of screening strategy parameters, a comprehensive examination of the effects of screening age and threshold number of atheroma can be performed. As previously stated, as the screening age increases (or the threshold of atheromas decrease) the percentage of the population identified increases. Consequently, although the average increase in life for an individual in the identified population decreases, the total life years added to the population increases. These effects are illustrated in Figure 2.4. Six screening ages with thresholds between one and five atheromas are shown for a total of 30 screening strategies for each gender. Plots for both men (Figure 2.4a) and women (Figure 2.4b) demonstrate the relationship of the screening parameters and the resulting average increase in life for the individuals identified within each screening strategy. Each of the six lines, representing the different scan ages, consists of five connected points. From left to right, these points represent a screening threshold of five, four, three, two, or one atheroma. Thus, it can be seen that for both men and women that the screening strategy that provides the greatest average increase in life is that of a young age and high number of atheromas, as stated previously. However,
from this representation, it can be seen that the percentage of the total population that meets these scan parameters is extremely limited (~0.2%). Comparatively, scan parameters for the highest age and lowest threshold for number of atheromas only provide an average increase of 4-6 years, but increase the percentage of the population meeting the scan criteria to between 68-76%. By multiplying the average increase in life by the identified percentage of the population for each set of scan parameters, the total number of life years added to the population can be determined. The results of these calculations again show that for both men (Figure 2.4c) and women (Figure 2.4d), scanning at an increase age and with a lower threshold of atheromas provides the greatest increase in life years added to the population. Here it must be noted that because each treatment strategy targets a unique subpopulation, numerous considerations must be made. A consequence of an older screening age is that younger people who would benefit

**Figure 2.4**: Comparison of the effects of various screening strategies with variations in scan age as well as thresholds for number of atheromas. Examination of average increase in treated (a) men and (b) women, as well as the total number of life years added to the male population of (c) men and (d) women.
from atherosclerosis treatment would be missed. Additionally, using a lower threshold of atheromas would identify a larger number of individuals, but would also increase the percentage of unnecessarily treated patients and could increase healthcare costs. However, it is apparent that using a model such as SCD-Predict can generate compelling data, allowing for the development of clinical practices to develop optimal screening strategies that would likely lead to the saving of lives and increased life spans for individual patients.

2.2.5 Model Predictions of Medications

Having demonstrated that SCD-Predict was a functional tool for establishing screening parameters to aid in the identification of potential patients, SCD-Predict was next utilized to model the effects of disease intervention via the administration of medications to reduce the growth rates of atherosclerotic plaques and clots. While numerous medications with variant effects are used in clinical practice today, simulations were performed using the model to investigate the generic effects of highly effective medications rather than the implication of specific compounds. Thus, a statin-like medication #1 was modeled to reduce the growth rate of all types of atheromas by a factor of two, while an aspirin-like medication #2 was modeled to reduce the growth of clots by a factor of two. The effect of medication #1 administered independently at a young age, in conjunction with medication #2 at a young age, or in conjunction with medication #2 at an increased age was investigated for both men and women, simulating cohorts of 5,000 individuals for each condition (Figure 2.5). Specifically, these results illustrate the change in the number of SCDs (as a percentage of the population), arrhythmias, and time to plaque rupture relative to the case of no treatment (values were normalized to the gender-appropriate
value corresponding to no treatment). Importantly, as a result of medication, the mean time to plaque rupture is increased by 50-60% when treatment is started at 45 years of age, and it is increased by 30-40% when treatment is started at age 65. Similarly, the mean number of arrhythmias is reduced by more than 3-fold in the case of medications #1 and #2 used in conjunction beginning at age 45. Most importantly, however, the number of sudden cardiac deaths is reduced 4-fold in both men and women when medication #1 is started at age 45. This effect is even more pronounced when combined with medication #2 starting at age 45, with a 9-fold reduction of SCD in men and an 8-fold reduction in women. Postponing administration of these medications until age 65 attenuates this effect, with a more pronounced difference in efficacy observed in men than in women.

**Figure 2.5:** Comparisons of untreated and treated male (blue) and female (red) populations with various medication treatment parameters (medications #1 and #2 reducing atheroma and clot growth rates, respectively). Data for each gender normalized to the case of no treatment. Simulated cohorts containing 5,000 individuals for each condition provided information on percentage of SCD within each population, average number of arrhythmias, and average time to plaque rupture.

### 2.2.6 Model Predictions of Atheroma Distributions

Thus far, it has been shown that *SCD-Predict* is a versatile tool for the development of intervention and treatment strategies for the entire observed population, resulting in
optimized screening parameters and potential reduction of SCD. Perhaps more importantly, the model also provided insightful comparisons between individuals modeled to have died from SCD and those modeled to have died from non-SCD causes. These two groups present strikingly different characteristics in the observed number of atheromas of various morphologies and number of observed ruptures as demonstrated in Figure 2.6. For both men (Figure 2.6a) and women (Figure 2.6b) modeled to have died of non-SCD causes, it can be seen that the average number of all types of modeled atherosclerotic plaques (thin-cap, medium-cap, and thick-cap) per individual as well as the incidence of plaque rupture are initially low, increasing rapidly with age. Comparatively, for men (Figure 2.6c) and women (Figure 2.6d) modeled to have died of SCD, these values for average numbers of atheromas and ruptures are already elevated at an earlier age. To emphasize this, the total number of atheromas observed in the simulated population of men (Figure 2.6e) and women (Figure 2.6f) at various ages was compared directly between SCD and non-SCD groups. As shown, the number of atheromas in the non-SCD group is initially low (~0.5 total atheromas), increasing as much as 12-fold by age 85-90. In contrast, the number of atheromas in the SCD group is already high by age 45-50 (~5 total atheromas) and increases only by about 1.5-fold by age 85-90. This trend is conserved when investigating the number of ruptures, thin-cap, medium-cap, and thick-cap atheromas, demonstrating a much lower age-dependence in the SCD subpopulation than in the non-SCD subpopulation. Building upon these observations, the distribution of numbers of atheromas for men, women, and the entire population were compared between individuals simulated to have died of SCD and those simulated to have died of non-SCD causes (Figure 2.7). Critically, it is shown for each
Figure 2.6: Average numbers of thin-cap, medium-cap, thick-cap, and total atheromas as well as the average number of observed ruptures in various populations including: (a) Men having died of non-SCD causes, (b) Women having died of non-SCD causes, (c) Men having died of SCD, (d) Women having died of SCD. (e) Comparison of the total number of atheromas for men having died of either SCD or non-SCD causes. (f) Comparison of the total number of atheromas for women having died of either SCD or non-SCD causes.

group (men, women, and total simulated population) that the average number of atheromas present in individuals simulated to have died of SCD is significantly higher (p-value <10^{-6}) higher than the average number observed in the non-SCD population. Combined with the observation that the number of atheromas is relatively independent of
age for the SCD subpopulation, this suggests that a high number of atheromas at a young age is by itself capable of serving as a potential biomarker of SCD risk.

**Figure 2.7:** Comparison of the average number of atheromas for men, women, and total simulated population between individuals simulated to have died of SCD and individuals simulated to have died of non-SCD causes. In each case, the average number of atheromas is significantly higher (p-value <10^{-6}) for the SCD group compared to the corresponding non-SCD group.

### 2.3 Discussion

A discrete event simulation model of sudden cardiac death was generated by fitting pathological data for the number of ruptured and non-ruptured vulnerable atheromas, as well as the number of erosive and thick-cap atheromas from a study of 50 whole human hearts [185]. Application of this model to the general population predicted an age-dependent cumulative and annual risk of SCD that is in compliance with established epidemiological data, demonstrating its validity [181, 183, 193]. Additionally, as existing and emergent techniques for the non-invasive detection of atherosclerotic plaques continue to improve, screening of the general population given specific parameters and thresholds becomes a more viable possibility for identification of individuals at risk of
CVD and subsequent SCD. This model was demonstrated to be capable of providing a means to optimize screening strategies, and showed great promise in the application of efficient medications to reduce the progression of the underlying disease. The results from these investigations of screening and medical intervention suggest that early screening combined with simple treatments would substantially increase the life expectancy of all population groups. Most importantly, analysis of the data comparing the number of atheromas in the SCD and non-SCD subpopulations suggest the hypothesis that the total number of atheromas could serve as an early predictor, or biomarker, of SCD.

The important advantage of the DES approach is that it allows for the integration of data at multiple scales and levels: from microscopic to macroscopic levels, and from pathological, clinical, and epidemiological data. It allows for the combination of unique blocks describing processes at these various levels of detail, incorporating each aspect into a more holistic model that can grow more complex as information on additional factors and interactions of the modeled system become available. Based on these properties, SCD-Predict can be upgraded as new data on the mechanism of the incidence of SCD and its associated risks emerge. Several directions to further enhance the model may already be available. For example, chains of events representing other etiologies of SCD that disproportionately impact earlier age groups such as cardiomyopathies or inherited genetic conditions could be added to improve the predictive abilities of the model in these younger individuals. When modeling SCD in young adults, it is important to take into account the high prevalence of atherosclerosis in the age groups between 15-34 years [194]. Although the number of atherosclerotic plaques has been reported to be
several fold lower for these age groups, it was found that atherosclerotic disease was not uncommon, with many of these individuals presenting at least one lesion in the right coronary artery. This finding confirms the notion that repair mechanisms are more effective in young adults than in older individuals. As a result, blocks that model the healing and repair processes could be added to the current model, generating a more applicable simulation for this younger age group.

The strength of the DES model is its flexibility to incorporate numerous interacting chains of events. As more data become available, it may be possible to include chains of events representing disorders known to be comorbid with CVD such as depression, diabetes, and obesity. Another important avenue is to include blocks that represent pathways of cholesterol metabolism, a process known to be crucial for several of the conditions described. This interaction of diseases and comorbidities could, in turn, provide this model with the capability of assessing complex burden of disease for a given population. One of the major metrics by which the World Health Organization (WHO) analyzes burden of disease involves the measure of Disability-Adjusted Life Years (DALYs) [195]. This standard combines the Years of Life Lost (YLL) to disease with the Years Lived with Disability (YLD) in the simple formula DALY = YLL + YLD. Certain diseases, such as SCD, cause a significant amount of early death, while having a relatively low amount of resultant disability. In comparison, other diseases such as blindness or diabetes cause varying degrees of early death with relatively high amounts of resultant disability. Therefore, this DALY metric allows for the comparison of the impact of distinctly different diseases on a given population (one DALY effectively equals one year of healthy life lost). SCD-Predict, though not specifically designed at this
time to account for this, does provide information pertaining to the number of YLL and could be upgraded to include YLD (which can be extracted by comparing initial incidence of atherosclerotic plaque rupture/cardiovascular event and its associated effects with the onset of SCD). Thus, with these in place, interactions between the various chains of events would provide a more holistic representation of the steps leading to SCD as well as the impact of disease on both individuals and populations. Although these enhancements of the model would enable more personalized prediction of the risk of SCD for a given individual, they would more than likely not change the general conclusions of this work as it was based on average non-stratified data. Critically, the results of the model suggested that identification of those at greatest risk of SCD could be identified at a young age based solely on the number of atherosclerotic lesions present in their coronary arteries. The next logical step, therefore, would be to investigate ways in which these individuals can be identified, potentially creating a risk-factor profile for these high-risk individuals.

2.4 Methods

The SCD-Predict software was developed with the SimEvents simulation engine, which is part of the MATLAB 2011b (Mathworks, Natick, MA) family of software tools. The structure of the model is presented in Figures A1, A2, and A3 in Appendix A. Figure A4 in Appendix A provides the chain of events determining the likelihood of SCD for each arrhythmia.

2.4.1 Age of First Atheroma Formation

The age distribution of the first atheroma formation was estimated from the data gathered in an epidemiological study of 826 subjects in the age range 40-84 years [182]. The
cumulative risk of atherosclerosis observed in this study was approximated by the integrals of the normal distributions with mean of 55 years for men and 61 years for women, and standard deviation of 12 years. Because both the size of the detectable atheroma and the age at which it was first observed depend on the sensitivity of the utilized diagnostic procedure, it was important to use the data on the initial size and time of formation from the same study. In this way, errors due to the limited sensitivity in the initial size and time of formation distributions compensate for each other.

2.4.2 Atheroma Growth

Based on the data from the utilized epidemiological study, the size of the atheroma at formation was assumed to have the mean diameter of 1.3 mm for all types of lesions. The mean diameter of the coronary arteries was assumed to be 3.57±0.84 mm, according to electron beam computer tomographic angiography [192]. Therefore, the mean normalized initial size of atheroma (ratio of diameters of atheroma and coronary artery) was set to 0.36.

Both normalized initial sizes of atheromas and the time intervals (inter-generation time values) between the events of formation of atheromas of each type were simulated as random numbers with the probability density function described by the Weibull distribution [196]:

\[ f(x; \lambda, k) = \frac{k}{\lambda} \cdot (x / \lambda)^{k-1} \exp \left( -\left( \frac{x}{\lambda} \right)^k \right) \]  

(2.1)

This function describes the family of distributions starting from highly asymmetrical exponential type (0<k≤1) to Rayleigh (k = 2), to a nearly normal symmetrical probability density function (k = 5). The normalized initial atheroma size distribution was thus described by the Weibull function with \( \lambda = 0.36 \) and \( k = 4.5 \). This results in the standard
deviation ($\sigma/\lambda = 0.23$) consistent with the observed relative standard deviations for coronary artery diameter. The distributions of the inter-generation time values were modeled with the Weibull function with parameters $\lambda$ and $k$ that differed for different types of atheroma, and can be found in Table A1 in Appendix A.

A simple model was used to describe the growth of the lesion:

$$\frac{dx}{dt} = rx(1 - x/a)$$

(2.2)

In this function, $x$ represents the diameter of the lesion, $a$ represents the diameter of the blood vessel, and $r$ represents the growth rate. This logistic growth model is widely used within biology to model the growth of some generic object (e.g. population, tumor, lesion) that is being fed by an external resource while being constrained due to finite available space. As a result, the object must necessarily stop growing as $x$ approaches $a$. Given its widespread use to describe biological growth in the literature, this same growth model (2) was used for all atheromas and clots. However, the initial sizes, growth rates ($r$), and thresholds at which rupture of the thin-cap atheroma or erosion of the medium-cap atheroma occurred were allowed to vary. Average growth periods for atherosclerotic plaques have been observed to be approximately 15-20 years [157]. When the threshold for rupture or erosion is reached, a clot is formed and starts growing as described by equation (2). However, the growth rate of the clot in the case of a ruptured atheroma is much faster (months) than in the case of erosive atheroma (years). Arrhythmia is assumed to occur as soon as the diameter of an artery lumen is reduced to the threshold value due to the growth of the clot (thrombosis) or due to the growth of a stable thick-cap atheroma (stenosis).
The relationship of artery lumen narrowing and the occurrence of ventricular arrhythmia has been a topic of several studies and discussions. In one such study [184], it was concluded that neither core size nor cap thickness were statistically related. Moreover, it established that neither of these two characteristics of vulnerable plaques contributed to the size of the lesion or its degree of stenosis. In 71% of the patients investigated in that study, it was found that a thrombus occluding greater than 30% of the lumen cross-sectional area was present. Additionally, it has been noted that the presence of a cross-sectional luminal narrowing of at least 75%, even in the absence of a thrombus was a sufficient potential cause of SCD [186]. An additional study from the same group demonstrated that the mean percentage of stenosis in the arteries of 142 victims of SCD was associated with the number of prior plaque ruptures [187]. This mean varied form 56% to 83%, depending on the number of present healed ruptures. In a pathological study of the hearts from 72 men and 18 women who died of SCD, it was shown that the percentage of stenosis varied from 78±12 in the case of plaque rupture to 70±11 in the case of plaque erosion [188]. Research from this same group demonstrated that luminal stenosis was least in sections with thin-cap atheroma (59.6%), and greatest in regions of plaque rupture (73.3%) or healed plaque rupture (72.8%) [189, 190]. Finally, a study of SCD of adult hospital patients reported that 62% of these individuals had 75% or greater stenosis of one or more coronary arteries, supporting the conventional statement observed in previous reports [191]. Thus, as shown, the reported percentage of stenosis is assumed to vary from 30% to 83%, and depends on a multitude of factors including the morphology of the lesion. However, even for a given type of lesion, there is limited consensus between studies, likely due to numerous unreported factors involved in the
onset of arrhythmias. Nevertheless, the majority of these report luminal stenosis of approximately 50-75% in cases of SCD. Given the fact that not all arrhythmias are fatal, the threshold for arrhythmia is likely lower than the observed stenosis percentage in the postmortem analysis of SCD arteries. Based on these considerations, we used the threshold value of a 56% reduction in lumen cross section, corresponding to the radius of the lesion equal to 0.75 of the radius of the artery.

2.4.3 Likelihood of Arrhythmic Cardiac Death

Personal history of sudden cardiac arrest (SCA) and arrhythmias are among the risk factors of SCD. To account for this in the DES model, it was assumed that the likelihood to survive (L(n)) each successive arrhythmia is lower than the previous, and is described by the function:

$$L(n) = p^{\ln(n)}$$

(3)

where p represents the probability to survive the first arrhythmia and n represents the number of arrhythmias in a given patient. The existing data on arrhythmia survival rate are contradictory. While anticipating that survival rate depends on multiple factors including type of arrhythmia, age of the patient, and the promptness of advanced medical support [179, 197], this has not been introduced to the current iteration of the model. Instead, the compromise average of 22% for the probability (p) to survive the first arrhythmia has been used.
The three chains of events leading to SCD were assumed to be independent in the current version of SCD-Predict (Figures A1-A3). The only interaction between these chains of events was the competition for ending the simulation process. As soon as the condition of SCD was reached in any of the three chains, the simulation was halted and the data on age at death, number of arrhythmias, and number and morphology of atheromas for the given individual was returned to the parent program.
Chapter 3

Identification of Accelerated Atherosclerosis in Youth

3.1 Perspective

Cardiovascular disease (CVD) is the leading cause of death worldwide, and is predicted to remain in this position for at least the next 10 years. Moreover, CVD is a true global epidemic, affecting both developed and developing nations. While CVD is consistently responsible for approximately 38% of all deaths in high-income countries, the rates in lower- and middle-income countries (LIMIC) and newly industrialized countries (NIC) are varied: ranging from as low as 10% in sub-Saharan Africa to as high as 58% in Eastern Europe [198-200]. However, the overwhelming majority of these deaths across all socioeconomic regions are attributable to the formation and growth of atherosclerotic lesions. Due to the global impact of CVD, a substantial amount of effort has been made towards identifying potential risk factors associated with atherosclerosis. Currently, it is generally accepted that common risk factors of atherosclerosis and CVD include modifiable risk factors such as high cholesterol, smoking, and obesity as well as non-modifiable risk factors including genetics and family history, sex, and race [19, 21].

Numerous large population studies have examined CVD risk factors to determine their potential prognostic capabilities for future acute cardiovascular events. For example, the PROSPECT\(^2\) study of 697 individuals in the United States and Europe investigated the presence of coronary atherosclerosis following percutaneous coronary intervention (PCI) [201]. Using VH-IVUS (see Chapter 1), this study demonstrated that baseline

\(^2\) Providing Regional Observations to Study Predictors of Events in the Coronary Tree
demographics such as age, sex, and race presented modest correlations with lesion formation and subsequent acute coronary events. Additionally, the inclusion of clinical risk factors such as the history of diabetes and modifiable risk factors marginally increased this correlation [202]. Similar findings were observed in additional studies such as the ARIC³ study, which consisted of 15,792 individuals in the United States and included measures of subclinical atherosclerosis [203]. However, many of these studies placed focus on the enrollment of patients who had been admitted for acute coronary events and belonged to age groups between 50 and 65 years old. The INTERHEART study of over 26,000 cases and controls predominantly included individuals from these older age groups, and demonstrated that a small number of risk factors were responsible for over 90% of the population attributable risk of acute myocardial infarction for both men and women [18]. This finding suggested that modification of risk factors in younger populations might be capable of reducing or preventing future acute coronary events.

In contrast to the majority of other population studies, the PDAY⁴ study sought to determine the relationship of CVD risk factors to atherosclerosis in much younger subjects. In this study, data was gathered from nearly 3,000 subjects between the ages of 15 and 34 who had died from homicide, suicide, or accidental injury [29]. This data included analysis of the extent and prevalence of arterial raised lesions, as well as the postmortem measurement of numerous CVD risk factors. Although atherosclerotic plaque morphology can be extremely diverse, the collective term 'raised lesion' includes simple fibrous plaques in addition to complicated fibrous plaques (exhibiting calcification, hemorrhage, ulceration, or thrombosis) [27, 29].

³ Atherosclerosis Risk in Communities
⁴ Pathobiological Determinants of Atherosclerosis in Youth
demonstrated that even subjects in the youngest age groups possessed such raised lesions in their coronary arteries [27]. More importantly, this analysis demonstrated that the risk factors associated with coronary atherosclerosis in adults were also important in the progression of the disease in youth.

In the previous chapter, a discrete event simulation (DES) model for the analysis of sudden cardiac death (SCD) was described [158]. The design of this model had been based on information regarding the number of coronary raised lesions observed in the postmortem analysis of hearts belonging to individuals who had died of both SCD and non-SCD causes. Subsequently, the model could accurately simulate populations of individuals exhibiting various coronary plaque morphologies beginning at age 45 and progressing until a simulated death. Importantly, this model could identify those that were simulated to die of SCD and those simulated to die of non-SCD causes, allowing for the comparative analysis of these two groups. Results from this DES model demonstrated that the number of coronary raised lesions was independently sufficient to assign an individual to either the SCD or non-SCD group with a high degree of accuracy. Moreover, the results established that for individuals in the non-SCD group, the number of raised lesions was initially low and increased markedly with age. In contrast, individuals of the SCD group exhibited a high number of coronary raised lesions at a young age, and demonstrated less age-dependent increases in this population over the age of 45. Based on these findings, it was hypothesized that a high number of coronary raised lesions could serve as an early predictor of future SCD risk in a much younger population. In this chapter, the possibility of identifying these high-risk individuals is explored, using the data garnered by the PDAY study, and based on quantification of
coronary raised lesions. Analysis of risk factor data from the subsequent subpopulations using generalized linear regression modeling could then elucidate which risk factors – or combinations of risk factors – could account for inclusion into the identified high-risk group.

3.2 Results

3.2.1 Identification and Analysis of the Discovery Cohort

The dataset utilized for the reanalysis of the PDAY study data consisted of 2,651 subjects with corresponding demographic and risk factor information (see section 3.4 Methods). Risk factors measured in this study included serum lipoprotein cholesterol concentrations, serum thiocyanate levels as a measure of smoking, body mass index (BMI) as a measure of obesity, and glycosylated hemoglobin levels as a measure of emergent diabetes mellitus [194]. Importantly for the present analysis, the number of raised lesions in the right coronary artery (RCA) of subjects in the study was counted. For many of these subjects, values for one or more of the investigated risk factors were missing. Selecting subjects presenting complete demographic and risk factor profiles, a subset of 739 subjects was identified and will be referred to as the Discovery Cohort. Comparing the demographic data of the Discovery Cohort and the entire PDAY dataset, it can be seen that there are no significant differences in race or sex between these two groups (Figure 3.1a). The ages of subjects in the Discovery Cohort spanned the entire range included in the PDAY study, and were relatively uniformly distributed from 15 to 34 years old (Figure 3.1b). In contrast, the distribution of the number of raised lesions found in the RCA of subjects of the Discovery Cohort was highly asymmetric (Figure 3.1c). While the majority of these subjects possessed zero raised lesions in their RCA,
Figure 3.1: Identification and analysis of the Discovery Cohort. (a) Demographic comparison of the entire PDAY study population (n = 2,651) and the identified Discovery Cohort (n = 739). (b) Age distribution of the subjects in the Discovery Cohort. (c) Distribution of the number of observed raised lesions in the RCA of subjects belonging to the Discovery Cohort. (d) Identification of 93 subjects exhibiting an elevated number of coronary raised lesions (red circles). (e) Demographic analysis of the 93 subjects belonging to the high-risk subset of the Discovery Cohort. (f) Comparison of the ages (left) and observed number of coronary raised lesions (right) between the 646 subjects belonging to the main population of the Discovery Cohort (low-risk) and the identified high-risk subset of 93 subjects.
nearly 20% of Discovery Cohort individuals were found to have anywhere between one and seven lesions. To account for the wide range of ages within the Discovery Cohort, the number of observed raised lesions was re-examined as a function of age (Figure 3.1d). As anticipated, the median number of observed raised lesions increased with an increase in the age of the corresponding group of subjects. Additionally, this grouping also allowed for the identification of subjects presenting an elevated number of raised lesions compared to other individuals at their age (Figure 3.1d, red circles). A total of 93 subjects in the Discovery Cohort were identified using this approach, representing approximately 12.6% of the cohort population. This subset of subjects included a 17-year old subject possessing three raised lesions, and another subject with seven raised lesions at age 26. Based on the findings from the previously established DES model (Chapter 2), this subpopulation of the Discovery Cohort with accelerated atherosclerosis would subsequently be at a much greater risk of SCD at later stages of life. Analysis of this high-risk group reveals that they are similar to the remaining 646 subjects within the Discovery Cohort, exhibiting similar distributions of race, sex, and age (Figure 3.1e,f). Importantly, however, there is a dramatic difference between the average number of raised lesions between the groups, with the low-risk subpopulation exhibiting an average of 0.116 raised lesions compared to the average of 2.183 raised lesions for the high-risk subpopulation (Figure 3.1f). Consequently, the annual risk of raised lesion formation is 19-fold higher in the subjects of the high-risk subpopulation compared to the subjects in the remainder of the Discovery Cohort.

3.2.2 Generalized Linear Regression Modeling: Numbers of Raised Lesions in High- and Low-Risk Subjects
Having identified a subpopulation of the Discovery Cohort exhibiting accelerated atherosclerosis, it was important to examine potential associations of inclusion into this high-risk group with known CVD risk factors. As previously described, the CVD risk factors investigated in the PDAY study included age, race, sex, serum lipoprotein and cholesterol concentrations, serum thiocyanate levels, BMI, and glycosylated hemoglobin levels (see Section 3.4 Methods). Potential correlations of each of these risk factors with the observed distribution of raised lesions within the entire 739-subject Discovery Cohort were investigated. However, it is exhaustively demonstrated that individual risk factors represent poor predictors of overall CVD risk when analyzed independently [193, 204, 205]. The results from this initial analysis are consistent with these prior observations, and demonstrated that no single risk factor was capable of accounting for the marked differences in the prevalence of raised lesions between the high-risk and low-risk groups previously identified (Figure B1 in Appendix B). A complete summary of the methods and results of this univariate analysis, as well as complete tables of correlation coefficients and p-values for these correlation coefficients can be found in Appendix B for this and all subsequent analyses.

In contrast to independent risk factors, it has been shown that the cumulative interaction of multiple risk factors demonstrates great efficacy in determining an individual’s absolute risk of CVD [193]. Therefore, generalized linear regression modeling (GLM) was utilized to investigate the effects of all of the measured CVD risk factors as well as their potential interactions on the resulting number of coronary raised lesions. Initially, the Discovery Cohort was modeled as a single group, containing all 739 previously identified subjects. The solution to this single population model contained 27
unique terms (including both interactions and quadratic terms for each of the known CVD risk factors), but left 60% of the variance in the observed number of raised lesions unaccounted for (see Appendix B for details on this, and all subsequent GLM analyses). Additional analysis of this model for the entire Discovery Cohort revealed that the residuals (differences between the observed and predicted values) did not exhibit a normal distribution, instead possessing long tails (Figure B3 in appendix B). A group of 82 outliers were identified from the Discovery Cohort as having residuals of 0.75 or greater. Interestingly, there was a 79% overlap of these 82 subjects with the previously identified high-risk subset of the Discovery Cohort. Taken together, these results indicate that when the Discovery Cohort is modeled as a single population, no combination of measured risk factors is sufficient to predict the observed distribution of coronary raised lesions. Additionally, these results support the previous identification of a unique, potentially high-risk subset of subjects within the Discovery Cohort who exhibit accelerated atherosclerosis; an acceleration that is not accounted for by the measured traditional risk factors for CVD.

Based on the relatively poor predictive performance of the single population GLM for the Discovery Cohort, the high-risk subset of 93 subjects exhibiting accelerated atherosclerosis was analyzed independently of the remaining 646 subjects (referred to as the low-risk subset), generating two independent GLMs. In contrast to the single population model, predictions of this two-group model more accurately reflected the observed number of raised lesions from the Discovery Cohort. For example, the single population model overestimated the number of individuals possessing one raised lesion by nearly two fold, while underestimating the number of subjects with two or more raised
Figure 3.2: Generalized linear models of the number of raised lesions in the RCA. (a) Comparison of the observed number of raised lesions with predictions from the two approaches used for GLM analysis (single population model and two-group approach). (b) Predicted number of raised lesions for the 646 subjects belonging to the general population of the discovery (low-risk subset) varying only levels of LDL. Additional predictions for the low-risk subset of the Discovery Cohort demonstrating effects varying multiple risk factor levels, including varying serum thiocyanate levels at the (c) 50th percentile of LDL and (d) 75th percentile of LDL measured in the Discovery Cohort. Dashed line indicates predicted number of lesions at age 25.

lesions by three fold (Figure 3.2a). In contrast, the two-group model was accurate (within 23% error) when describing individuals with one, two, or more raised lesions. Table 3.1 provides a complete comparison of both approaches with data observed in the Discovery Cohort, and includes the correlation coefficients of the observed or predicted number of raised lesions with the measured CVD risk factors. Once again, while the single population GLM consistently overestimates these correlation coefficients, the two-group approach provided far more accurate values. Furthermore, these results support the
previous identification of the high-risk group of the Discovery Cohort as a unique and independent subset.

**Table 3.1:** Number of raised lesions. Comparison of the observed and modeled distributions and correlation coefficients. For this analysis, race (white = 1, black = 2) and sex (female = 1, male = 2) were given quantifiable values.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>‘Single Population’</th>
<th>‘Two Group’</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.376</td>
<td>0.376</td>
<td>0.352</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td>0.966</td>
<td>0.467</td>
<td>0.822</td>
</tr>
<tr>
<td>0 Lesions</td>
<td>595</td>
<td>570</td>
<td>598</td>
</tr>
<tr>
<td>1 Lesion</td>
<td>79</td>
<td>147</td>
<td>61</td>
</tr>
<tr>
<td>2 Lesions</td>
<td>29</td>
<td>14</td>
<td>48</td>
</tr>
<tr>
<td>3 Lesions</td>
<td>18</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>4+ Lesions</td>
<td>18</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

In addition to providing more accurate values compared to the single population model, the two-group approach allows for additional comparisons between the previously
identified high-risk subset of the Discovery Cohort and the remaining 646 subjects. Most notably, the two models differed dramatically in terms of complexity. The equation describing the main, low-risk group of the Discovery Cohort included nine of the ten measured risk factors in addition to interaction and quadratic terms (Equation B2 in Appendix B). Subsequently, for subjects belonging to this group (approximately 87% of the Discovery Cohort), changes in the levels of modifiable risk factors resulted in changes in the predicted numbers of coronary raised lesions. To demonstrate this effect, the equation generated by this GLM was plotted using LDL concentrations corresponding to the 25th, 50th, 75th, and 95th percentile observed in the Discovery Cohort (Figure 3.2b). As demonstrated, increasing the concentration of LDL at a given age results in an increase in the predicted number of raised lesions. Additionally, it should be noted that when the predicted number of raised lesions at a given age is less than one, this value is interpreted as the probability at that age for a raised lesion to form. Thus, for a subject described by this GLM at age 25, an increase in LDL concentration from the 25th percentile to the 95th percentile corresponds to an increase in probability to develop a raised lesion from 1.4% to 21% (Figure 3.2b, dashed line). Additionally, because the equation for the GLM describing this group was dependent on nearly all of the measured CVD risk factors, varying the levels of multiple factors simultaneously has widely variant effects on the predicted number of coronary raised lesions. For example, the predicted number of raised lesions at various thiocyanate levels and median LDL concentrations (Figure 3.2c) differ from the predictions corresponding to various thiocyanate levels and the 75th percentile LDL concentrations (Figure 3.2d). Under these two sets of conditions, the probability to develop a lesion at age 25 increases from 2.8% at median LDL and
thiocyanate level to 24.1% at the 75th percentile LDL and thiocyanate levels (Figure 3.2c and 3.2d, dashed lines).

In stark contrast to the relatively complex model describing these 646 subjects of the Discovery Cohort, the generated model for the 93-subject high-risk subset was dramatically simpler. For this model, the equation describing the predicted number of coronary raised lesions ($N$) relied solely on the non-modifiable risk factors of race, sex, and age:

$$\ln(N) = 0.564 + 0.077(age) - 1.327(race) - 1.025(sex) + 0.734(race)(sex)$$  \hspace{1cm} (3.1)

Importantly, although this equation includes terms for both race and sex as well as an interaction term for these factors, their inclusion only marginally improves the models percentage of explained variance, and their p-values were either above or just at the threshold of significance. As a result, and accounting for the known demographic composition of the high-risk subset of the Discovery Cohort, the predicted number of coronary raised lesions for this group can be determined by the following equation:

$$\ln(N) = -1.34 + 0.077(age)$$  \hspace{1cm} (3.2)

The simplicity of this equation reveals an important distinction for the 93 subjects of the identified high-risk subset of the Discovery Cohort. Specifically, this equation suggests that the formation of coronary raised lesions within this high-risk subset of the Discovery Cohort is not detectably dependent from the levels of modifiable risk factors of CVD. For the low-risk subset of the Discovery Cohort, it was shown that increasing, for example, LDL or serum thiocyanate levels resulted in an increase in the predicted number of raised lesions (Figure 3.2b-d). In contrast, increases in these risk factors, or changes in any other measured modifiable risk factor have no impact on the probability for an individual in the
high-risk subset to develop a raised lesion. This simple equation also allows for the rapid determination of the age at which the first raised lesion will form for a member of the high-risk subset of the Discovery Cohort. Solving this equation for \( N = 1 \), the model predicts that the first coronary raised lesion will be observed at an average age of 17.4, a result consistent with observations of the PDAY study.

### 3.2.3 Verification of Findings in Larger Population

Analysis of the Discovery Cohort revealed a subpopulation of approximately 13% that exhibited markedly accelerated coronary atherosclerosis. To validate the findings observed in this high-risk subset of the Discovery Cohort, the remainder of the PDAY study population was analyzed. From the PDAY study dataset of 2,651 subjects, the 739 subjects of the Discovery Cohort were removed. The remaining 1,912 subjects compose what will be referred to as the Verification Cohort. Due to the initial selection of the Discovery Cohort, the data available pertaining to measurements of risk factors within the Verification Cohort was somewhat less complete. However, there has been no demonstrated selection bias in the PDAY study, and the failure to obtain a particular risk value for a given individual could not be related in any way to the presence or absence of coronary atherosclerosis. Thus, it would be expected that subjects within the Verification Cohort would exhibit similar distributions of risk factors as those observed in the Discovery Cohort. Indeed, both the demographic profile (Figure 3.3a) and age distribution (Figure 3.3b) of the Verification Cohort are similar to those observed for the Discovery Cohort and the PDAY study population as a whole.

To begin this large-scale analysis, the distribution of coronary raised lesions for the subjects of the Verification Cohort was examined (Figure 3.3c). As seen for the
Figure 3.3: Identification and analysis of the Verification Cohort. (a) Demographic comparison of the entire PDAY study population (n = 2,651) and the identified Verification Cohort (n = 1,912). (b) Age distribution of the subjects in the Verification Cohort. (c) Distribution of the number of observed raised lesions in the RCA of subjects belonging to the Verification Cohort. (d) Identification of 250 subjects exhibiting an elevated number of coronary raised lesions (red circles). (e) Demographic analysis of the 250 subjects belonging to the high-risk subset of the Verification Cohort. (f) Comparison of the ages (left) and observed number of coronary raised lesions (right) between the 1,662 subjects belonging to the main population of the Verification Cohort (low-risk) and the identified high-risk subset of 250 subjects.
Discovery Cohort, the distribution of raised lesions within the Verification cohort was highly asymmetric, with approximately 80% of subjects exhibiting no raised lesions, but others with as many as nine raised lesions. The 1,912 subjects of the Verification Cohort were then rearranged, and the number of raised lesions was plotted as a function of age (Figure 3.3d). From this grouping, 250 subjects were identified as outliers exhibiting an elevated number of raised lesions for their age (Figure 3.3d, red circles). This high-risk subset represented approximately 13% of the Verification Cohort, a proportion strikingly similar to the value observed within the Discovery Cohort. Additional comparisons between the Verification and Discovery Cohorts could readily be made. Although the demographic distribution of the high-risk subset (Figure 3.3e) varied marginally from the remaining 1,662 subjects of the low-risk subset, the average ages of these two groups (Figure 3.3f) were nearly identical. However, the average of 2.27 raised lesions in the high-risk subset was significantly higher than the 0.10 raised lesions of the low-risk subset, an observation parallel to the previously examined Discovery Cohort.

Previously, the single population model generated for the Discovery Cohort had performed poorly in describing the observed number of raised lesions. Given the identified similarities between the Verification and Discovery Cohorts, it was determined that a two-group approach would be the most effective method to generate a GLM describing the 250-subject high-risk subset of the Verification Cohort. Using this two-group approach, the dependence of the number of raised lesions on the measured CVD risk factors and their interactions was determined. The resulting equation describing this relationship for the high-risk subset of the Verification Cohort is similar to that of the Discovery Cohort, and is as follows:
\[ \ln(N) = -1.10 + 0.073(\text{age}) \]  \hspace{1cm} (3.3)

As seen, this equation is independent of all modifiable risk factors, in addition to race and sex. A comparison of the equations to describe the high-risk subsets of the Discovery (Eq. 3.2) and Verification (Eq. 3.3) Cohorts reveal only minor differences. For example, the growth rate of lesions for subjects in the high-risk subset of the Verification Cohort was slightly lower than the corresponding rate of the Discovery Cohort (0.073 compared to 0.077). Additionally, the age of first lesion formation was predicted to be slightly lower for the high-risk subset of the Verification Cohort than for the high-risk subset of the Discovery Cohort (15.1 compared to 17.4). However, these equations suggest that for the high-risk subsets of both cohorts, the formation of coronary raised lesions is independent of the levels of modifiable risk factors.

3.2.4 Confirmation of Large-Scale Analysis and Cohort Comparisons

The 1,912 subjects of the Verification Cohort consisted of those individuals within the PDAY study population that lacked complete risk factor profiles (with those that possessed complete profiles belonging to the Discovery Cohort). It was determined that the equation describing the predicted number of raised lesions for the high-risk subset of this Verification Cohort exhibited no dependence on any of the measured risk factors aside from age. Thus, a potential concern is that the predicted lack of dependence on CVD risk factors may be a consequence of the incomplete measurements available for this cohort. To address this point, and to confirm the initial findings of the Discovery Cohort, one final cohort was generated from the PDAY study population.

Within the 1,912 subjects of the Verification Cohort, a subset of 751 subjects was identified for which measurements were available for all demographic and risk factor
Figure 3.4: Identification and analysis of the Confirmation Cohort. (a) Demographic comparison of the entire PDAY study population (n = 2,651) and the identified Confirmation Cohort (n = 751). (b) Age distribution of the subjects in the Confirmation Cohort. (c) Distribution of the number of observed raised lesions in the RCA of subjects belonging to the Confirmation Cohort. (d) Identification of 84 subjects exhibiting an elevated number of coronary raised lesions (red circles). (e) Demographic analysis of the 84 subjects belonging to the high-risk subset of the Confirmation Cohort. (f) Comparison of the ages (left) and observed number of coronary raised lesions (right) between the 667 subjects belonging to the main population of the Confirmation Cohort (low-risk) and the identified high-risk subset of 84 subjects.
categories with the exception of serum thiocyanate levels. Examination of this group, referred to as the Confirmation Cohort, revealed a high number of similarities to both of the previously identified cohorts. Comparisons of the demographic (Figure 3.4a), age (Figure 3.4b), and raised lesion (Figure 3.4c) distributions within the Confirmation Cohort reflected trends observed previously. Upon further analysis of this group, 84 subjects were identified as exhibiting accelerated atherosclerosis for their age (Figure 3.4d). These 84 subjects all belonged to the previously identified high-risk subset of the Verification Cohort, and exhibited a similar demographic distribution to other previous high-risk subsets (Figure 3.4e). Finally, this high-risk subset of the Confirmation Cohort demonstrated a similar age profile to the remainder of the cohort, but significantly increased atherosclerosis (Figure 3.4f), a trend observed across all cohorts.

Analysis of the Confirmation Cohort and generation of a GLM describing this group was performed in a manner similar to the Discovery and Verification Cohorts described previously. From the two-group model, the resulting equation describing the distribution of raised lesions within the high-risk subset of the Confirmation Cohort was only marginally more complex than the equations for the previous two cohorts, as shown:

$$\ln(N) = -0.864 + 0.050(age) + 0.0025(LDL)$$  \hspace{2cm} (3.4)

An immediate distinction of this equation from those describing either the Discovery or Verification Cohort is the inclusion of the term for levels of LDL. However, while this equation suggests some dependence on LDL, the predictive effects are minimal. The effect of a 10% increase in LDL concentration across the range of measured values generates an increase in the number of predicted raised lesions by less than 5%. Additionally, this equation continues to predict no significant dependence on any other
measured risk factor for the development of raised lesions in the identified high-risk subsets of the various established PDAY cohorts.

The equations established for the high-risk subsets of the Discovery (Eq. 3.2), Verification (Eq. 3.3), and Confirmation (Eq. 3.4) Cohorts allow for the comparison of each of these models. The predictions for the number of coronary raised lesions as a function of an individual’s age are largely consistent across all three models (Figure 3.5a). To account for the contribution of LDL to the model of the Confirmation Cohort, predictions for the number of raised lesions as a function of age were plotted at various LDL concentrations corresponding to the 5th (68.0 mg/dL), 50th (128.0 mg/dL), 75th (165.5 mg/dL), and 95th (230.6 mg/dL) percentiles as recorded for this cohort (Figure 3.5b). Importantly, all three equations predict a similar number of raised lesions for a subject of the high-risk subset at age 25 (Figure 3.5, dashed lines and Table 3.2). Moreover, when these models are extrapolated to age 45, they predict that the number of coronary raised lesions increases to between 4.9 and 8.9, values consistent with previous
analyses on the frequency of raised lesions in victims of CVD-related SCD (see Chapter 2) [158, 185].

Table 3.2: Number of raised lesions. Comparison of the observed and modeled distributions and correlation coefficients.

<table>
<thead>
<tr>
<th>Model</th>
<th>Age of Onset</th>
<th>Slope</th>
<th>N at 25 years</th>
<th>N at 45 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>17.4</td>
<td>0.077</td>
<td>1.79</td>
<td>8.38</td>
</tr>
<tr>
<td>Verification</td>
<td>15.1</td>
<td>0.073</td>
<td>2.06</td>
<td>8.89</td>
</tr>
<tr>
<td>Confirmation</td>
<td>9.5</td>
<td>0.05</td>
<td>2.17</td>
<td>5.89</td>
</tr>
</tbody>
</table>

3.3 Discussion

The current reanalysis of the data generated by the PDAY study sought to identify a subset of individuals exhibiting an elevated number of coronary raised lesions, a condition previously recognized as a reliable predictor for future incidence of SCD. Once identified, it was the objective of this reanalysis to determine a set of modifiable CVD risk factors that could classify subjects as belonging to this high-risk subset of the larger population. From the initial PDAY study dataset of 2,651 subjects, the present analysis first targeted a 739-subject group for which complete demographic and CVD risk factors profiles were available. Analysis of this Discovery Cohort revealed a set of 93 subjects (representing 12.6% of the cohort population) presenting an unusually high number of raised lesions within their RCA compared to members of their respective age groups. Through the use of general linear regression, it was demonstrated that the Discovery Cohort could not effectively be modeled as a single population, and thus that this high-risk subset of subjects was distinguishably unique. When modeled independently, it was observed that the contribution of measured modifiable risk factors varied between the
high-risk subset and the remainder of the Discovery Cohort. Specifically, these traditional risk factors were identified as important contributors to the formation and severity of coronary raised lesions in the general population as expected. In contrast, however, the model describing the identified high-risk subset exhibited no contributions from any modifiable risk factors toward lesion prevalence. This two-group model describing the Discovery Cohort was then validated using the larger dataset of the remaining 1,912 subjects of the PDAY study, comprising the Verification Cohort. A similar proportion of this cohort (approximately 13%) was identified as belonging to a high-risk subset of individuals. Subsequently, modeling of the two resulting groups demonstrated findings congruent with those of the Discovery Cohort; traditional risk factors are determinant for the extent of atherosclerosis in the general population, but were absent in the description of disease in the identified high-risk subjects. Finally, a selection of 751 subjects with the most complete dataset was isolated from within the Verification Cohort to confirm the findings of these first two large groups. Analysis of this Confirmation Cohort yielded results remarkably similar to the previous two cohorts, demonstrating a near complete independence from known modifiable risk factors (with the exception of serum LDL concentration) for description of the number of raised lesions observed in this cohort’s high-risk subjects. A summary of the cohorts identified, and the primary findings from each, is provided in Figure 3.6.

The results of this reanalysis support the growing body of literature derived from the PDAY study, indicating that for the majority of the population (approximately 87% of the population) traditional risk factors represent critical predictors for the development of coronary raised lesions. As demonstrated by the Discovery Cohort, the prediction of the
number of raised lesions for the 646 subjects not included in the high-risk subset required the involvement of nine of the ten measured CVD risk factors. As a result, modifications or reductions in any of these risk factors will result in a reduction in the overall extent of atherosclerosis for these individuals. However, the essential aspect of the present reanalysis is the identification of a high-risk subset of individuals representing ~13% of the general population, with markedly accelerated atherosclerosis (~20-fold faster) and starting at a strikingly young age (~16 years old) for which traditional risk factors seem to have limited to no effect on the development and progression of disease. In addition to the known and powerful contribution of aging, this group of high-risk young individuals seems to be affected by yet unrecognized risk factors contributing to accelerated atherosclerosis. Subsequently, for this high-risk group, traditional therapies targeted toward the reduction of modifiable risk factors may or may not be efficient in reducing raised lesion formation. As such, it is critical to identify this group prospectively, perhaps by utilizing novel genetic risk profiling. To date, numerous single nucleotide polymorphisms (SNPs) have been identified as conferring some level of increased risk to CVD [206]. Importantly, although the effect size of many SNPs is small, the combination of multiple of these SNPs has been shown to contribute to the genetic susceptibility of

**Figure 3.6:** Summary of cohort PDAY study population and cohort identification and analysis. The PDAY study population (white) consists of 2,651 subjects. The Discovery Cohort (green) of 739 was identified based on availability of demographic and risk factor measurements. From the Discovery Cohort, a high-risk subset (green, red outline) of 93 subjects (12.6%) was identified. The Verification Cohort (light blue) consisted of 1,912 subjects not included in the Discovery Cohort. A similar high-risk subset from the verification cohort (light blue, red outline) was identified and contained 250 subjects (13.1%). The Confirmation Cohort (dark blue) consisted of 751 subjects from within the Verification Cohort based on availability of risk factor measurements. The high-risk subset of this cohort (dark blue, red outline) included 84 subjects (11.2%).
Total PDAY Study Dataset
n = 2650

Discovery Cohort
n = 739

High-Risk
n = 93 (12.6%)

Verification Cohort
n = 1911

Confirmation Cohort
n = 751

High-Risk
n = 84 (11.2%)

Verification Cohort
n = 10 Subjects

Confirmation Cohort

Discovery Cohort

High Risk
coronary atherosclerosis [207]. Once identified, the extent of disease could then be confirmed through the use of non-invasive imaging modalities such as CT-angiography or MRI (see Chapter 1). In a majority of cases, SCD is believed to be secondary to the destabilization of a coronary raised lesion with subsequent thrombus formation, myocardial infarction, and death with asystole or ventricular fibrillation. Utilizing this strategy for the identification of individuals belonging to the high-risk subset of the population, new therapeutic strategies can be developed that can retard or block the development of new lesions. Ultimately, it is the hope that these treatments could then potentially reverse or regress the process of atherosclerosis progression, and prevent the onset of SCD or other thromboembolic coronary events.

3.4 Methods

3.4.1 Brief Description of the Dataset

The PDAY dataset utilized for this analysis consisted of information obtained from collected blood, arteries, and selected tissues from 2,651 individuals between the ages of 15 and 34. In addition to demographic data for the entire dataset, information regarding the number and classification of fatty streaks (a non-raised stage of atherosclerosis) and raised lesions in the right coronary artery (RCA), percentage vessel stenosis in the left anterior descending (LAD) artery, serum lipoprotein cholesterol concentrations, BMI, serum thiocyanate levels as an indicator of smoking status, and glycosylated hemoglobin levels as a measure of blood glucose and emergent diabetes mellitus were obtained [208, 209]. Additional detailed descriptions of the methods of collection and analysis have been published previously [194, 210, 211].
3.4.2 Statistical Analysis

All calculations were performed using MATLAB R2012b (Mathworks, Natick, MA) with the Statistical Toolbox. Box plots presented in Figures 3.1-3.4 were calculated using GraphPad Prism 6 (GraphPad Software, Inc. La Jolla, CA). Outliers were determined using the Tukey method of analysis in which the central mark in each box represents the median, while the edges of the box are the $q_1 = 25^{th}$ and $q_3 = 75^{th}$ percentiles. The whiskers extend to the most extreme data points not considered outliers (identified as those in the interval less than $q_1 - 1.5(q_3 - q_1)$ and greater than $q_3 + 1.5(q_3 - q_1)$). These outliers were then plotted as red circles.

3.4.3 Matrices of Correlation Coefficients

Matrices of correlation coefficients $R$ were calculated using the MATLAB function $[R, P] = \text{corrcoef}(x)$, which also generates and returns $P$, a matrix of $p$-values for testing the hypothesis of no correlation. Each $p$-value is the probability of getting a correlation as large as the observed value by random chance, when the true correlation is zero. Correlation coefficients found to be significant ($p < 0.05$) are highlighted in yellow in Tables B1-B5, while the $p$-values of these correlation coefficients are presented in Tables B6 and B7 in Appendix B.

3.4.4 Generalized Linear Regression Modeling

Generalized linear modeling was performed using the MATLAB function GeneralizedLinearModel.stepwise.m, creating a generalized linear regression model by stepwise regression. Since the number of raised lesions can be any non-negative integer, the Poisson distribution of the response was assumed, and the default ‘log’ link function for the Poisson distribution was used. Stepwise regression in the MATLAB Statistical
Toolbox systematically adds or removes terms to the generalized linear model based on their statistical significance in explaining the response variables. The stepwise procedure begins with an initial model, comparing the explanatory power of incrementally larger and smaller models based on the specified criterion. Here, the default ‘sse’ criterion was utilized (see MATLAB’s Statistical Toolbox documentation for additional details).

Details of the generated GLMs are located in Appendix B. The initial model generated for the Discovery Cohort assumed a ‘single population’ of all 739 members of this group. The equation generated from this model contained 27 terms, including intercept, 9 linear terms, 11 interaction terms, and 6 pure quadratic terms. However, even for this complicated model, the values of R-squared, adjusted R-squared, and adjusted generalized R-squared were low. For example, adjusted generalized R-squared reached only 0.39, indicating that a significant percentage of observed variance is left unexplained by this model. Figure B3 in Appendix B presents the values of the residuals for the model in both histogram and probability modes, indicating strong deviation of the residuals from a normal distribution.

Similar to the described model above, the GeneralizedLinearModel.stepwise.m function was utilized to generate two separate GLMs for the low-risk core group of 646 subjects and for the 93 high-risk outliers identified from the boxplot analysis on the observed number of raised lesions in the RCA of the Discovery Cohort population. When compared to the ‘single population’ model, this ‘two group’ model resulted in a reduction of the fraction of the population with large residuals, as well as the values of these residuals (Figure B3 in Appendix B). This procedure was repeated to generate and analyze GLMs for both the Verification and Confirmation Cohorts.
Chapter 4

Bioluminescence-based Apoptosis Detection

4.1 Perspective

The presence of controlled cell death has been identified in a diverse range of cell and tissue types as far back as the mid-nineteenth century [212, 213]. However, it was not until the introduction of the unifying term ‘apoptosis’ in the 1970’s that the importance of this ubiquitous process was rigorously examined (See Appendix C) [214]. Since that time, it has become apparent that apoptosis is not only crucial for normal homeostatic cell maintenance within whole tissues, but is also directly involved in a number of pathologies including cancer, neurodegenerative diseases (Alzheimer’s and Parkinson’s), cardiovascular disease, and autoimmune diseases [215-218]. Because both apoptosis and defects in apoptotic pathways are known to be essential components of these issues, a critical focus of research in this field has been centered on the development of techniques and platforms capable of selectively detecting apoptotic cells.

Although the numerous pathways of apoptosis initiation are varied and diverse, most coalesce at the activation of caspases (Figure 4.1). Once activated, these proteases are directly responsible for the biochemical and morphological changes that define the apoptotic process, including nuclear condensation, cellular blebbing, and – importantly – the exposure of the membrane phospholipid phosphatidylserine (PS) on the surface of the cell [219-223]. Herein, we describe the design, development, and implementation of a bioluminescent protein biosensor that specifically targets this morphological change,
Figure 4.1: Pathways of apoptosis (extrinsic – blue box, and intrinsic – orange box) and detection concept. Both pathways converge to the activation of caspases that are principally responsible for morphological and biochemical changes associated with apoptosis including phosphatidylserine (PS) exposure. Although PS is sequestered to the inner-face of the plasma membrane under healthy conditions, the ArFP construct exhibits the ability to bind PS at surface of apoptotic cells and generate a bioluminescent signal in the presence of substrate coelenterazine.
requires no post-expression modification, and provides an effective means to detect apoptotic cells both *in vitro* and *in vivo*.

Annexin V is an endogenous human protein (encoded by the ANXA5 gene) that binds specifically to PS with a high affinity (~$10^{-9}$ M) (Figure C1 in Appendix C) [224, 225]. Previous work describes Annexin V modified with either fluorescent [226-229] or radiolabel [230-232] reporters and used for the detection of apoptosis, but presents numerous challenges. Detection approaches utilizing fluorescence, specifically in the context of detection within a biological sample, inherently suffer from issues of high background signal and low sensitivity based on the prevalence of numerous fluorescent species found in biological samples [233]. Isotopic labeling carries restrictive regulations on handling, waste disposal, and health effects, while both of these approaches exhibit a high degree of variability in consistency of reagent preparation, presenting another major hurdle. Although the use of these detection methods is less than ideal, the *in vivo* detection of apoptosis still represents an important avenue of research, and has been employed to monitor a range of events such as cardiac allograft transplant stability and chemotherapeutic efficacy [234, 235]. To that end, the use of bioluminescence presents an opportunity to address the inherent challenges of apoptosis detection presented by these other methods.

*Renilla* luciferase (RLuc) is an enzyme natively expressed by the sea pansy *Renilla reniformis* that oxidatively decarboxylates its substrate coelenterazine to generate bioluminescence (Figure C1 in Appendix C) [236, 237]. Detection of this type of emission generally results in lower background and higher signal-to-noise ratios, providing a means of detection that is more sensitive than fluorescence, especially in the
context of biological samples [233, 238]. However, it has been shown that this wild-type RLuc is quickly rendered inactive in the presence of murine serum [239], strictly limiting previous work with a *Renilla* luciferase-labeled Annexin V probe to applications *in vitro* [240, 241]. Protein engineering of the wild type RLuc protein has yield a luciferase variant (RLuc8) that exhibits a 200-fold increase in serum stability, as well as a 4-fold increase in light output [239]. This affords a unique opportunity to couple its use with relevant *in vivo* disease models to detect apoptosis, and analyze the role which apoptosis plays within them.

We have utilized this RLuc8 mutant to generate a chimeric bioluminescence-based Annexin V apoptosis detection construct. We first analyze the structural and biochemical properties of this Annexin V-*Renilla* luciferase fusion protein (also referred to as ArFP) to verify that the construct retains the ability to both bind to PS while also remaining capable of bioluminescent emission. As shown in Figure 1, under apoptotic conditions when PS is no longer sequestered on the inner-face of the cell membrane, ArFP will bind to the surface of the apoptotic cells, allowing for detection via bioluminescence generation from addition of the substrate coelenterazine. Importantly, we demonstrate that, in addition to functioning as a specific sensor for apoptosis *in vitro*, this construct also allows for the bioluminescence imaging of apoptosis *in vivo*. With this functionality, we are able to examine the involvement of apoptosis within the context of entire tissues in processes such as surgery-induced vascular apoptosis and age-related macular degeneration (AMD), providing a new means to observe and utilize these and other models of disease-specific apoptosis.
4.2 Results

This research introduces a new and important expansion to the currently available toolkit of Annexin V-based detection systems for the specific identification of apoptosis. Specifically, herein we report the cloning and production of a chimeric fusion protein of Annexin V and RLuc8 (ArFP), and performed a comprehensive analysis of both the physical and biochemical characteristics of this fusion. We also demonstrated that the biosensor is capable of apoptosis detection in vitro. Moreover, this study represents the first description of a bioluminescent Annexin V apoptosis sensor successfully applied in vivo, providing new means to assess clinically relevant disease models involving apoptosis.

4.2.1 Design, Expression, and Characterization of ArFP

Genes for both Annexin V and RLuc8 were cloned out of their respective plasmids, and combined in a single fusion gene which was inserted into the pET-30 Xa/LIC vector to generate pET-30/ArFP (see Section 4.4 Methods, Figure C2 and C3 in Appendix C). Following transformation of this vector, ArFP was expressed in E. coli as soluble protein, and isolated using immobilized metal affinity chromatography. Before this novel bioluminescent Annexin V-based biosensor could be utilized in apoptosis detection, both the physical and biochemical properties were characterized. Analysis by SDS-PAGE indicated that isolated ArFP had a purity exceeding 95 % and exhibited an approximate size of 73 kDa, while purified samples of Annexin V and RLuc8 demonstrated similar purities with sizes of 36 kDa and 37 kDa respectively (Figure 4.2a). Identification of this isolated construct by antibodies against both Annexin V and RLuc8 via Western blot analysis (Figure 4.2b) supports the chimeric incorporation of both of these components.
into a single fusion, requiring no post-expression workup or modification. Similarities of the circular dichroism (CD) spectra for each of the three proteins in regions corresponding to secondary structure contributions also suggest that the fusion of Annexin V and RLuc8 has not compromised the individual secondary structure integrity of the proteins (Figure 4.2c). The far-UV CD spectrum for ArFP was further analyzed by the CDSSTR analysis package, generating secondary structure assignments for the fusion (Figure 4.2d). The presence of predominantly α-helical structure with minor β-sheet contribution in the fusion is in agreement with the known crystal structures of RLuc8 and Annexin V, the latter consisting nearly exclusively of α-helices (Figure C1 in Appendix C). In addition, the denaturation temperature (Tm) of the fusion protein was determined, and exhibited two distinct transition phases at approximately 48 °C and 73 °C (Figure 4.2e). Comparison of temperature-relevant sections of this curve to thermostability measurements made for Annexin V and RLuc8 indicate that each transition roughly corresponds with each component of the fusion (Figure 4.2f,g). Taken together, these results from SDS-PAGE, Western blot, and CD spectra analysis indicate that ArFP presents a combination of the structural characteristics of both Annexin V and RLuc8 with no need for post-expression modifications. Unlike previous Annexin V reporters relying on fluorescence or isotopic labeling, ArFP maintains a strict stoichiometric ratio of Annexin V and RLuc8 with no batch-to-batch variability. Additionally, the CD thermostability observed in the spectra of the fusion protein suggests that the two components of the fusion remain largely independent of each other, and should retain both PS-binding and bioluminescence activities at physiologically relevant temperatures.
Figure 4.2: (a) SDS-PAGE analysis of components and fusion protein (lane 1: Annexin V, lane 2: RLuc8, lane 3: ArFP, lane 4: Precision Plus Protein Ladder (Bio-Rad)). (b) Western blot analysis of components and fusion protein identified using anti-Annexin V antibody (lane 1: Precision Plus Protein Ladder, lane 2: Annexin V, lane 3: RLuc8, lane 4: ArFPP) or anti-Renilla luciferase antibody (lane 5: Precision Plus Protein Ladder, lane 6: Annexin V, lane 7: RLuc8, lane 8: ArFP). (c) Far-UV CD spectra of ArFP (purple), Annexin V (teal), and RLuc8 (blue). (d) Protein 2° structure assignment (CDSSTR analysis via Dichroweb) for ArFP. (e) Thermostability analysis (ellipticity at 222 nm) of ArFP. (f) Comparison of Annexin V and ArFP thermostability (upper) and first derivative analysis of curves (lower). (g) Comparison of RLuc8 and ArFP thermostability (upper) and first derivative analysis of curves (lower).
4.2.2 Activity Characterization of ArFP

Analysis of the protein activity of ArFP began with an investigation into its ability to bind PS. Isothermal titration calorimetry (ITC) was used to calculate the binding affinity of ArFP for PS exposed on the surface of lipid vesicles (Figure 4.3a). The dissociation constant (K_D) of 20.7 µM calculated for ArFP closely matched the previously reported value of 13 µM for native Annexin V, and suggests that ArFP will interact with exposed PS on cellular surfaces in an analogous manner (Table 4.1) [242]. Comparatively, the bioluminescence characteristics of ArFP were also found to be highly similar to those of native RLuc8 (Figure 4.3b-d). Minor variations were observed from RLuc8 to ArFP, including a red-shift in wavelength of peak emission, and a slight increase in emission half-life (Table 4.1). However, total luminescence output measured across a range of protein concentrations from 50 pM to 500 nM resulted in no statistical difference at any tested concentration (Figure 4.3d). Based on the results of this characterization of both the physical and biochemical properties of ArFP, we were confident that the fusion protein would be capable of binding to surface-exposed PS while also generating a detectable bioluminescent signal either \textit{in vitro} or \textit{in vivo}.

\textbf{Table 4.1:} Summary of Protein Characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Calculated Molecular Weight (kDa)$^5$</th>
<th>Binding Affinity for PS (K_D, µM)</th>
<th>Maximum Bioluminescence Emission (nm)</th>
<th>Bioluminescence Half-Life (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annexin V</td>
<td>39.3</td>
<td>13$^6$</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RLuc8</td>
<td>39.5</td>
<td>--</td>
<td>487$^7$</td>
<td>3.11$^5$</td>
</tr>
<tr>
<td>ArFP</td>
<td>77.5</td>
<td>20.7</td>
<td>492</td>
<td>3.59</td>
</tr>
</tbody>
</table>

$^5$ Molecular weights determined from primary amino acid sequence  
$^6$ Binding affinity of native Annexin V reported in [242]  
$^7$ Emission characteristics of RLuc8 reported in [239]
4.2.3 Apoptosis Detection *In Vitro*

Having demonstrated that the physical and biochemical properties of both Annexin V and RLuc8 had been successfully incorporated into ArFP, we tested the ability of this...
construct to detect apoptosis *in vitro*. ArFP was added to Jurkat (human T lymphocyte) cells that had been treated with the known apoptosis inducer staurosporine [243, 244]. Bioluminescence intensities were then measured following the addition of the substrate coelenterazine, and as shown, exhibit increasing luminescence intensities as the staurosporine incubation time increases (Figure 4.4a). These observations were validated through the use of flow cytometric analysis, in which cells treated with staurosporine were subsequently stained using fluorescently labeled Annexin V (Annexin-FITC) and the vital dye propidium iodide (PI) (Figure 4.4b). As observed, an increase in staurosporine incubation results in an increase in apoptotic cells (Annexin-FITC⁺/PI⁻), while loss of membrane integrity (Annexin-FITC⁻/PI⁺ or Annexin-FITC⁺/PI⁺) wasn’t observed until the 6-hour time point. The activation of caspase-3 represents another reliable indicator of apoptosis induction as it serves as a convergence point for many different apoptotic pathways. This activation involves processing of caspase-3’s pro-enzyme form to its active proteolytic form, and is known to be responsible for the display of PS in apoptotic cells once activated [219, 245]. Western blot analysis of staurosporine treated and control Jurkat cell lysates clearly demonstrate a reduction in the levels of pro-

**Figure 4.4:** (a) Bioluminescence response of ArFP incubated with Jurkat cells after treatment with 1 µM staurosporine for various times. Bioluminescence intensity reported as the mean of three measurements ± one standard deviation. Error bars that are not visible are obstructed by the data point. (b) Flow cytometry analysis of Jurkat cells following treatment with 1 µM staurosporine for various times and dual stained with Annexin-FITC (FL1, 525/40 band pass filter) and the vital dye propidium iodide (PI) (FL11, 610/20 band pass filter) (Data acquisition: 2•10⁴ events per sample). (c) Western blot analysis of Jurkat cell lysates following treatment with 1 µM staurosporine for various times. (d) Bioluminescence response of ArFP incubated with Jurkat cells treated with 25 µM caspase inhibitor Z-VAD(OMe)-FMK followed by treatment with 1 µM staurosporine for various times. Bioluminescence intensity reported as the mean of three measurements ± one standard deviation. Error bars that are not visible are obstructed by the data point.
caspase-3 dependent on staurosporine incubation duration, suggesting the liberation of activated caspase-3 (Figure 4.4c). To investigate ArFP’s specificity in targeting PS exposed via this caspase activation during apoptosis, Jurkat cells were pre-incubated with the cell permeable, pan-caspase inhibitor Z-VAD(OMe)-FMK prior to staurosporine treatment. Following incubation with both the apoptosis inducer and the caspase inhibitor, ArFP was added to the treated cells, but the resulting bioluminescence measurements showed no significant intensity increase with respect to incubation time (Figure 4.4d). These data demonstrate that the induction of apoptosis and the activation of caspases within the cell are responsible for an increase in the number of bound ArFP constructs, directly resulting in an increase in bioluminescence signal intensity. Additionally, bioluminescence-based apoptosis detection of staurosporine-treated cells was performed using a second cell line, namely, Caco-2 human epithelial colorectal adenocarcinoma cells (Figure C4 in Appendix C). Finally, as a confirmation of staurosporine impact on the health of these cells, the cellular metabolic activity was examined with respect to staurosporine incubation duration for both Caco-2 (Figure C5 in Appendix C) and Jurkat (Figure C6 in Appendix C) cells. These experiments confirmed that ArFP is capable of \textit{in vitro} apoptosis detection by the generation of bioluminescence signal in a rapid and highly specific manner.

4.2.4 Apoptosis Detection \textit{Ex Vivo}

Apoptosis has an important role in numerous diseases ranging from cancer, to cardiovascular disease, to degenerative diseases such as AMD. Thus, the ability to detect and image apoptosis not just in cell culture, but also in tissues composed of heterogeneous populations of cells is highly desired. In our work, we began by
investigating vascular apoptosis by employing a rat model of arterial injury, as well as corneal and retinal apoptosis models in mice. Specifically, in our first animal study, we performed balloon angioplasty on *Sprague-Dawley* rats to create conditions of ischemia/reperfusion in the iliac arteries (Figure 4.5a). Upon exposure of our ArFP reagent to the vessels, a significantly increased (*p* < 0.01) bioluminescence signal was obtained from the angioplasty-treated artery (right iliac, Figure 4.5b, teal) when compared to either the ischemia-only (left iliac, Figure 4.5b, purple) or control (carotid, Figure 4.5b, blue) arteries. This suggests an elevated level of apoptosis from balloon angioplasty and reperfusion over ischemia/reperfusion as expected. The presence of apoptosis in treated vessels was confirmed through bioluminescence imaging with IVIS (Figure 4.5c), and histological analysis (Figure 4.5d). These results indicated that the fusion protein successfully binds to apoptotic cells in whole tissues as well. The low level of background observed in the carotid artery (Figure 4.5b, blue) is likely due to the known autoluminescence of coelenterazine in biological samples [246], and can be accounted for by setting specific thresholds of bioluminescence emission.

### 4.2.5 Apoptosis Detection *In Vivo* and Disease Model

In addition to examining vascular apoptosis, we also investigated the feasibility of employing our apoptosis detection technology in both corneal and retinal tissues of mice *in vivo*. We employed a mouse model of ethanol-induced corneal apoptosis in which we

**Figure 4.5:** (a) Diagram of operational procedure, illustrating position of treated and collected tissues (teal: right iliac/treated condition, purple: left iliac/reperfusion control, blue: carotid/healthy control). (b) Luminescence comparisons of treated (teal) and untreated (purple) iliac arteries as well as the carotid (blue) artery as a healthy control, performed in triplicate. (c) IVIS image of the carotid (1), left iliac (2), and right iliac (3) arteries after treatment. (d) Histological analysis of treated vascular tissue indicating the emergence of apoptotic cells following treatment.
compared EtOH-treated and PBS control mice following topical application of our ArFP reagent and its coelenterazine substrate (Figure 4.6a) [247, 248]. Controls for this experiment included either topical application of ArFP or coelenterazine alone to the cornea of the animal (Figure 4.6a, Control 1 and Control 2, respectively). Both controls showed that either ArFP or coelenterazine alone were insufficient to generate a detectable bioluminescence signal. In contrast, the application of ArFP and substrate in an eye lacking apoptosis induction (Figure 4.6a, Untreated) resulted in a small, yet detectable signal of very low intensity, likely due to some non-specific binding of the protein to the cornea. However, when applied together to an apoptosis-induced cornea, a significantly increased (p < 0.01) bioluminescence signal was generated (Figure 4.6a, Treated and Figure 4.6b). We confirmed that the observed increase in bioluminescence corresponds to a dramatic increase in apoptotic cells by using the TUNEL analysis of corneal whole mount and cross-sections from these same mice (fluorescein-positive/green) (Figure 4.6c). These results indicate that ArFP is capable of specifically detecting apoptosis in the corneal tissues of mice and represents the first use of bioluminescence-based in vivo Annexin V apoptosis detection.

The demonstrated success of ArFP in the bioluminescence-based detection of apoptosis in corneal tissues led us to investigate further its ability to detect disease-
relevant cell apoptosis and death in deeper tissues, such as the retina. It has been reported that intravenous sodium iodate (NaIO$_3$) injection causes retinal degeneration, resulting in a model of age-related macular degeneration (AMD) [249-252]. Specifically, mechanisms of apoptosis have been implicated as the underlying cause of this degeneration, and could thus be identified using the ArFP construct [251, 252]. As shown, a strongly positive bioluminescence signal was observed from ArFP applied to the eyes of NaIO$_3$ treated mice, while only a limited amount of background is observed for the control mouse receiving only PBS (Figure 4.7a). These results were verified through histological analysis of retinal cross sections, revealing a large number of

![Figure 4.7](image_url)

**Figure 4.7:** (a) IVIS analysis of bioluminescence comparing control (Saline injection) and retinal degeneration model (NaIO$_3$ injection) mice. (b) Histological analysis of retinal sections from control (A1, PBS injection) and retinal degeneration model (A2 and A3, NaIO$_3$ injection) mice. Left: H&E stain of retinal cross sections. Right: TUNEL analysis of retinal cross sections indicating presence of apoptotic (green) and non-apoptotic (blue) nuclei.
apoptotic cells in the NaIO$_3$-injected mice (Figure 4.7b, A2 and A3) that are completely absent in the saline control mouse (Figure 4.7b, A1).

4.3 Discussion

In this work, we report a new bioluminescence-based Annexin V biosensor for the detection of apoptosis. Because of its involvement in both natural cell clearance within multicellular organisms as well as in numerous disease states, the ability to detect and monitor apoptosis is critically important. In addition to the identification of excessive or abnormal cell death, apoptosis monitoring can also be applied to the development and screening of numerous therapeutic agents such as those developed to induce selective cell death within tumors.

Since its identification over thirty years ago, Annexin V has been coupled with diverse reporters for the identification of apoptotic cells via recognition and binding of phosphatidylserine. To date, the vast majority of these reporters have relied on either isotopic markers or conjugation to fluorophores. As previously discussed, the chimeric nature of the ArFP construct eliminates the challenges associated with batch-to-batch variability presented by earlier platforms reported in the literature, while also taking advantage of the inherently higher sensitivity and lower background of bioluminescence-based detection. The analysis of the structural, physical, and biochemical properties of our bioluminescent biosensing system demonstrates that the ArFP fully retains the properties of both Annexin V and RLuc8 with no need for post-translational modification.

Importantly, we demonstrated the feasibility of using ArFP for the first time as a tool for the design and development of a novel class of bioluminescence-based Annexin
V biosensors for in vivo detection of apoptosis. Using three different animal models, namely, rat vascular, mouse corneal, and mouse retinal apoptosis, we validated our in vitro detection of apoptosis, and showed successful bioluminescence imaging of disease-relevant of apoptosis. This further demonstrates that our ArFP biosensing detection method can be applied to the monitoring and study of different diseases. We envision that our ArFP biosensing system can be employed in a diverse range of clinically relevant diseases encompassing AMD, cancer, and even atherosclerosis and cardiovascular disease. Additionally, unique fusion proteins between Annexin V and other bioluminescence proteins such as aequorin, obelin, Gaussia luciferase, etc., can be created, enhancing the toolbox of sensing proteins with tailored characteristics, such as varied emission wavelengths, kinetics, and bioluminescence half-lives. The ability to prepare a series of fusion proteins endowed with specific characteristics allows for the design of a new class of biosensors for a variety of targeted applications. We foresee that these biosensors will find numerous applications in both in vitro as well as in vivo identification, therapeutic screening, and risk assessment of diseases involving apoptosis.

4.4 Methods

4.4.1 Reagents and Kits

LB broth (Miller), LB Agar (Miller), dimethylsulfoxide (DMSO), KOD hot start master mix, pET-30 Xa/LIC Vector Kit, and ApopTag® Peroxidase In Situ Oligo Ligation Kit were purchased from EMD Millipore (Billerica, MA). Kanamycin sulfate, ampicillin (sodium salt), and potassium chloride were obtained from Amresco (Solon, OH). From VWR International (West Chester, PA), sodium chloride, ammonium sulfate, methanol, and 4-(2-hydroxymethyl)-1-piperazine-1-ethanesulfonic acid (HEPES) were purchased.
Calcium chloride, magnesium chloride, sodium iodate, L-arabinose, imidazole, heat inactivated fetal bovine serum (FBS), staurosporine, β-mercaptoethanol, Tween® 20, paraformaldehyde, nucleotide primers, the Roche In situ Cell Death Detection (TUNEL) kit, as well as mono- and di-basic phosphate salts of both potassium and sodium were all purchased from Sigma-Aldrich (St. Louis, MO). Triton X-100 was purchased from Alfa Aesar (Haverhill, MA). Sodium citrate was purchased from Fisher Scientific (Pittsburgh, PA), while normal saline solution was purchased from G Biosciences (St. Louis, MO). Problock™ Gold Bacterial 2D Protease Inhibitor Cocktail and isopropyl β-D-1-thiogalactopyranoside (IPTG) were purchased from Gold Biotechnology (St. Louis, MO). 3-(N-morpholino)propanesulfonic acid (MOPS) and the Pierce® BCA Protein Assay Kit were obtained from Thermo/Life Technologies (Rockford, IL). Lipids POPC (16:0-18:1 PC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) and POPS (16:0-18:1 PS, 1-palmitoyl-2-oleoyl-sn-glycero-3-[phospho-L-serine]) were purchased as chloroform solutions from Avanti Polar Lipids, Inc. (Alabaster, AL). Both high glucose Dulbeccos’ Modified Eagle Medium (DMEM) and Roswell Park Memorial Institute (RPMI) 1640 medium were purchased from Gibco/Life Technologies (Grand Island, NY), while trypsin-versene (EDTA) was purchased from Lonza (Basel, Switzerland), antibiotic-antimycotic solution was obtained from Cellgro/Corning (Manassas, VA), and Accutase® cell detachment solution was purchased from Innovative Cell Technologies (San Diego, CA). The CellTiter 96® AQueous One Solution Cell Proliferation Assay kit was obtained from Promega (Madison, WI). Primary antibodies were purchased from Abcam (Cambridge, MA) and Cell Signaling Technology, Inc. (Danvers, MA), while native coelenterazine was purchased from Nanolight Technology (Pinetop, AZ). The pan
caspase inhibitor Z-VAD(OMe)-FMK and the Annexin V-FITC Early Apoptosis Detection Kit were also purchased from Cell Signaling Technology, Inc. Competent NEB5α and T7 Express Escherichia coli strains were obtained from New England Biosciences (Ipswich, MA). Ni-NTA agarose and the QIaprep Spin Miniprep Kit were purchased from Qiagen (Hilden, Germany). Laemmli sample buffer and Mini-Protean® TGX™ (4-20%) gels were obtained from Bio-Rad (Hercules, CA). Odyssey® Blocking Buffer (PBS) and secondary antibodies were purchased from LI-COR Biosciences (Lincoln, NE). Isoflurane, USP was purchased from Piramal Healthcare (Bethlehem, PA), and VECTASHIELD Antifade Mounting Medium with DAPI was purchased from Vector Laboratories (Burlingame, CA).

4.4.2 Fusion Gene Design and Transformation

Plasmid pPROEX HTb containing Annexin V was provided by Prof. Seamus Martin, Trinity College, Dublin, Ireland. The Annexin V gene was isolated using forward (GGTATTGAGGGTCGCATGGCACAGGTTCTC) and reverse (AGAACCACCAG-AACCACCCTCATCTTTCTCCACA) primers with appropriate regions for inclusion into the pET-30 Xa/LIC vector (bolded) as well as for use in overlap-extension PCR (OE-PCR, underlined). The RLuc8 gene was isolated from pBAD/Myc-HIS A::RLuc8 using similarly designed forward (GGTGGTTCTGGTGGTTCTATGGCTTCCAAGGTG) and reverse (AGAGGAGAGTTAGAGCCTGCTCGTTCTTCAG) primers. Following fusion gene generation via OE-PCR, cloning was performed as per the pET-30 Xa/LIC Vector Kit’s instructions, and the product was transformed into competent NEB5α and T7 Express Escherichia coli. Nucleotide sequencing analysis performed on plasmids isolated from colonies of these bacteria confirmed the presence of a fusion gene
consisting of Annexin V and RLuc8, henceforth referred to as pET-30/ArFP (Figure C2 in Appendix C).

4.4.3 Protein Expression

Five milliliters Growth Media (LB Broth (Miller) with 35 µg/mL kanamycin) was inoculated with a T7 Express bacterial colony and grown overnight at 37 °C/250 rpm. After this growth period, 300 mL Growth Media was inoculated with 1 mL of this overnight culture and grown at 37 °C/250 rpm to an OD600 of 0.6. Protein expression was induced via addition of IPTG to a final concentration of 1 mM, and expression continued for two hours at 37 °C/250 rpm. Cells were harvested via centrifugation at 5,000 xg for 25 minutes at 4 °C, and the pellet was resuspended in 10 mL Lysis Buffer (50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH = 8.0) and 100 µL Problock™ Gold Bacterial Protease Inhibitor Cocktail was added. Purification of the ArFP construct was achieved utilizing an encoded N-terminal 6x Histidine tag as described by Qiagen. Expression and purification of native Annexin V and RLuc8 was achieved using this protocol with minor changes: both were expressed in Amp Growth Media (LB Broth (Miller) with 0.1 mg/mL ampicillin), and expression of RLuc8 was induced via addition of 0.2 % arabinose.

4.4.4 Verification and Characterization of ArFP

4.4.4.1 Western Blot Analysis of Protein Expression

A total of 1 µg purified protein was mixed with 2x Laemmli Sample Buffer with 5% β-mercaptoethanol, resolved by SDS-PAGE, and transferred to nitrocellulose membranes for blotting. Either anti-Annexin V mouse monoclonal (ab54775) or anti-Renilla luciferase rabbit polyclonal (ab187338) antibody were used as indicated, and membranes
were imaged using the LI-COR Odyssey® Classic (Model 9120) Imaging System (LI-COR Biosciences, Lincoln, NE).

4.4.4.2 Bioluminescence Measurement and Characterization

Purified ArFP and RLuc8 were dialyzed into ArFP Binding Buffer (10 mM HEPES, 100 mM NaCl, 5 mM KCl, 1.8 mM CaCl$_2$, 1 mM MgCl$_2$, pH = 7.4), and an aliquot of 100 µL was added in triplicate to a 96 well plate at concentrations ranging from 50 pM to 500 nM. Bioluminescence intensities were recorded using a PolarSTAR Optima (BMG LABTECH GmbH, Ortenberg, Germany) following injection of 2.5 µg/mL coelenterazine. Decay kinetics were measured by integrating the bioluminescence emission signal from the 500 nM sample in 0.24 second intervals for 50 seconds (coelenterazine injection occurred at 10 seconds). Bioluminescence spectra were obtained for both proteins using a Luminoskan Ascent Microplate Luminometer, recorded from 400 nm to 700 nm.

4.4.4.3 Circular Dichroism (CD) Spectroscopy Analysis

Samples of ArFP, Annexin V, and RLuc8 were dialyzed into CD Buffer (10 mM potassium phosphate, 100 mM ammonium sulfate, pH = 7.4) at a concentration of 0.1 mg/mL, and the far-UV CD spectra were recorded using a Jasco J-815 Spectropolarimeter (Jasco, Inc., Easton, MD) (Scan mode: continuous; Scan speed: 50 nm/min; Data pitch: 0.5 nm; Bandwidth: 1 nm; Data Integration Time (D.I.T.): 2 s; Accumulations: 3) [253]. Analysis of the ArFP CD data was performed using Dichroweb (http://dichroweb.cryst.bbk.ac.uk, Department of Crystallography, Institute of Structural and Molecular Biology, Birkbeck College, University of London, United Kingdom)[254], and secondary structure assignments were generated using the CDSSTR analysis package
Thermostability analysis was measured as the ellipticity at 222 nm as a function of temperature from 25 °C to 90 °C (Scan speed: 1 °C/min; Bandwidth: 2 nm; D.I.T: 2 s) [257].

4.4.4.4 Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) was performed using a Nano ITC (TA Instruments, New Castle, DE) and a modified protocol for the characterization of Annexin V [242]. ArFP was dialyzed into ITC Buffer (20 mM MOPS, 100 mM KCl, 0.75 mM CaCl₂, pH = 7.5) to a final concentration of 0.029 mM, while vesicles composed of 60:40 POPC:POPS were prepared in a sample of the final dialysate to a concentration of 30 mM (protocols for generation and sizing of these lipid vesicles found in Supplemental Information, Supplemental Figure S4). Titration of vesicles into ArFP solution was performed via 24 x 9.14 µL injections, and the integrated heat signals were analyzed with the NanoAnalyze Software (v 2.3.6, TA Instruments, New Castle, DE) using an independent binding model.

4.4.5 Cell Culture

Human T lymphocyte interleukin-2 (Jurkat, Clone E6) cells (ATCC® TIB-152™) were grown in RPMI 1640 supplemented with 10 % FBS, 100 mg/L penicillin, 100 mg/L streptomycin, and 2 mM glutamine.

4.4.5.1 Apoptosis Induction

Cells were grown in Corning® T25 Tissue Culture Flasks to an approximate density of 5·10⁶ cells/mL. Apoptosis was induced via addition of 1 µM of the known inducer staurosporine [258]. For apoptosis inhibition studies, the synthetic peptide Z-VAD(OMe)-FMK was added to the cells at a final concentration of 25 µM one hour prior
to the induction of apoptosis via staurosporine. For bioluminescence-based detection of apoptosis, cells were collected following incubation, centrifuged, and resuspended in Binding Buffer containing 1 nM ArFP for 5 minutes. Cells were then washed three times with Binding Buffer, and an equal number of cells were transferred to a 96 well plate in triplicate for bioluminescence measurements.

4.4.5.2 Flow Cytometric Analysis of Apoptosis

After induction of apoptosis, cells from each treatment condition were washed once in PBS and resuspended in Annexin Binding Buffer and stained with FITC-conjugated Annexin V and propidium iodide (PI) as per the manufacturer’s instructions. After staining, cells were then washed with Annexin Binding Buffer, resuspended in a fixing solution of PBS containing 4% paraformaldehyde (PFA), and stored at 4 °C in the dark. Flow cytometry of samples was performed using a CytoFLEX S (Beckman Coulter Life Sciences, Indianapolis, IN). Annexin-FITC fluorescence (FL1) was collected through a 525/40 band pass filter, while PI fluorescence (FL11) was collected through a 610/20 band pass filter. Data acquisition (2•10^4 events per sample) was performed using the CytExpert 1.2 software (Beckman Coulter Life Sciences, Indianapolis, IN).

4.4.5.3 Western Blot Analysis of Caspase Activation

Cells treated with staurosporine (or control) were collected (approximately 5•10^7 total cells per condition) and washed with PBS. Cells were lysed via addition of 1 mL ice-cold RIPA buffer containing 10 µL Halt™ Protease and Phosphatase Inhibitor (100X), and were stored at 4 °C for 30 minutes. Lysates were clarified via centrifugation at 12,000 xg for 25 minutes, and 1 µg of total protein (cell lysate) from each condition was analyzed
via western blot as previously described using anti-pro-caspase 3 rabbit polyclonal (CST #9662) and anti-β-actin mouse monoclonal (CST #3700) antibodies.

4.4.6 Animal Models of Apoptosis

All animal studies performed within this work were carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and were based on protocols approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Miami.

4.4.6.1 Ex vivo Rat Iliac Artery Apoptosis Induction

Three male Sprague Dawley rats (12 weeks old, 280-320 g, Harlan, Indianapolis, IN) were used to induce vascular apoptosis in the right iliac artery via balloon injury as previously described [259]. Briefly, the abdominal and iliac arteries were clamped at proximal and distal sites, respectively, and a 2 Fr. Fogarty Arterial Embolectomy Catheter with a 4 mm balloon (Edwards Lifesciences, Irvine, CA) was directed to the right iliac to induce injury. Blood flow was restored, and 10 % BSA was used to reduce non-specific binding. Following washing with Binding buffer, 100 µL 2 nM ArFP solution was added directly to the vessel surface, and incubated for 3 minutes. The vessels were again washed with Binding Buffer, and the right iliac (treated), left iliac (untreated), and carotid (control) arteries were excised and stored in Binding Buffer. Bioluminescence measurements were performed via addition of 50 µL of 0.1 mg/mL coelenterazine and recorded using either a Turner BioSystems 20/20n Luminometer (Promega, Madison, WI), or a Caliper/Xenogen IVIS® SPECTRUM in vivo imaging system (IVIS, Caliper, Hopkinton, MA). Immunohistochemical staining of balloon-injured (or control) arteries harvested 30 minutes after treatment were either visualized
with hematoxylin and eosin (H&E) stain, or used for apoptosis detection with the ApopTag Peroxidase In Situ Oligo Ligation Kit (EMD Millipore, Billerica, MA) according to the manufacturer’s guidelines. Apoptotic cells were identified by a dark brown nucleus surrounded by a red cytoplasm.

4.4.6.2 In Vivo Mouse Corneal Apoptosis Detection

Corneal injury was induced in the right eye of eight female BALB/c mice (8 weeks old, Charles River, Boston, MA) via exposure to 20 % ethanol for 40 seconds following anesthetization via ketamine/xylene (1.5 mg/0.3 mg) injection [247, 260]. The left eyes of these mice were treated with PBS as internal controls, after which both eyes were rinsed with normal saline. Five additional mice were used as untreated controls, receiving PBS on both eyes. 4 hours post-treatment, 10 µL 2 nM ArFP solution was added to the examined eye (treated or control) for 1 minute. Eyes were washed with Binding Buffer, and bioluminescence was recorded via addition of 5 µL of 0.1 mg/mL coelenterazine and imaged using IVIS. Whole mount and cross-section corneal samples were fixed using 2 % paraformaldehyde, and apoptosis was visualized using the terminal deoxynucleotidyltransferase (TdT) dUTP-biotin nick-end labeling (TUNEL) kit (In Situ Cell Death Detection Kit, Roche, Basel, Switzerland) according to the manufacturer’s protocol. Images of mounted samples were obtained with a Zeiss confocal microscope exciting at 488 nm for fluorescein (green) and 405 nm for DAPI (blue).

4.4.6.3 In Vivo Mouse Retina Apoptosis Detection

Following anesthetization with ketamine/xylene (1.5 mg/0.3 mg), four female BALB/c mice (8 weeks old, Charles River, Boston, MA) received tail-vein injection of 30 mg/kg sodium iodate (NaIO₃) while a control mouse received a volume matched injection of
PBS. Three days following injection, 10 µL 2 nM ArFP solution was added topically to the right eye of each mouse for 1 minute, and washed with Binding Buffer. Bioluminescence imaging performed via addition of 5 µL of 0.1 mg/mL coelenterazine followed immediately by image capture via IVIS. Retinal samples from an additional 3 mice (two NaIO₃-treated, one PBS-control) were prepared in a cryomold and sectioned for histological analysis. Sections were then either stained using H&E or treated with the TUNEL assay kit for apoptosis detection as previously described.

4.4.6.4 Statistical Analyses of Animal Studies

For purposes of in vivo apoptosis assessment, the number of animals was limited to avoid unnecessary use of living subjects. Although power analysis was not performed prior to data acquisition, the number of animals selected (n = 3 for vascular apoptosis and n = 13 for corneal apoptosis) was sufficient to generate a statistically relevant difference in values between populations. Data is given as mean and standard deviations, while data points for each individual are shown for all statistical analyses of the animal studies. Obtained data were assumed normally distributed, and statistical significance of the collected data was determined with the Student’s t-Test (unpaired, two-tailed) using the analysis software Prism 6.0 (GraphPad Software, Inc., La Jolla, CA). Healthy individuals were used for all experiments with no requisite randomization. Animals were neither selected for nor against additional specific criteria, and researchers were not blinded to these selections.
Chapter 5

Conclusions and Future Work

5.1 Perspective

With growing at-risk populations in lower- and middle-income countries around the globe, the importance and impact of cardiovascular disease is unlikely to be attenuated in the coming years. In the face of this growing epidemic, healthcare technologies and interventional measures will continue to improve, and must become increasingly common throughout the world. The first implantable pacemaker was deployed in 1958, and failed approximately 3 hours later [261]. In spite of this, the recipient survived, requiring the replacement of 22 different pulse generators over the course of his life. Today, there are well over 1 million patients with implanted pacemakers as a means to avert arrhythmias of the heart that could otherwise potentially lead to fatal cardiac events [262, 263]. These devices have seen dramatic improvements, with modern battery lives extended to as many as 15 years. Additionally, improvements in imaging and diagnostic tools will continue to advance the understanding of the underlying mechanisms of CVD and its subsequent manifestations. A myriad of imaging techniques have evolved over the past few decades, resulting in both invasive and non-invasive imaging modalities capable of providing critical information regarding the structure and function of the heart and the vasculature (see Chapter 1). Recently, the greatest advances in imaging technologies for the assessment of CVD have incorporated hybrid-imaging techniques, compounding the strengths of multiple imaging modalities. Most
notably, PET-MRI and PET-CT have both been shown to be extremely beneficial in the non-invasive assessment of disease. In addition, imaging also provides the means to bridge important gaps between translational research and clinical application. The detection of specific disease biomarkers in vivo using various imaging modalities allows for early disease diagnosis, and provides a means to explore novel therapies and their efficacy. Moreover, imaging can be utilized to monitor patients receiving newly developed therapeutics. In this way, these imaging modalities can assist in decisions regarding either a reduction or an increase of therapeutic dosage, and can potentially identify off-target effects of the treatment on the heart or vasculature. Most importantly, advances in CVD imaging present the capability of shifting CVD research further toward understanding the disease on a patient-specific level. As these and other technologies examining the underlying mechanisms of CVD advance, the ability to successfully treat and save individual patients will concurrently improve.

5.2 Modeling the Future

One manifestation of technology being applied to the understanding CVD is that of computational population modeling. The use of both discrete event simulation (DES) modeling, as well as generalized linear regression modeling (GLM), allowed for the examination of subpopulations at elevated risk of CVD and subsequent sudden cardiac death (SCD). Using a post-mortem analysis of the extent of atherosclerosis in an older population, it was determined that a high number of raised lesions at a young age served as a reliable biomarker of future SCD. These results were then applied to the study of a young population to examine the
contribution of a set of traditional CVD risk factors to the prevalence of atherosclerosis (see Chapter 2 and 3). However, due to the nature of the developed models, they are not restricted to the analysis of the pathological data and risk factor evaluation presented within. Factors contributing to CVD that are either currently recognized or that may be identified in the future can be incorporated into these flexible models. Therefore, as more information from large population studies becomes available, these models can be refined, providing ever more accurate descriptions of disease in both younger and older populations. Specifically, the contribution of various comorbid conditions such as diabetes, obesity, and depression can be modeled as independent chains, or as modifying events within the existing DES model framework. Additionally, as suggested by the re-analysis of the PDAY Study dataset, information pertaining to the genetic contributions to atherosclerosis in younger individuals will likely be of great importance. Once collected, this data can be incorporated into the models that have been generated for the high-risk subset of individuals identified in the PDAY population, potentially providing a means to rapidly identify those subjects at greatest risk of future clinical events leading to SCD.

5.3 Illuminating Potential Directions

The work presented within demonstrated the successful application of a chimeric fusion of Annexin V and RLuc8 toward the detection of apoptotic cells both in vitro as well as in vivo. Critically, apoptosis plays a central role in the progression of simple atherosclerotic lesions to complicated plaques. Specifically, as macrophages become engorged with lipids and lipid products in the intima, many become
apoptotic, and begin accumulating into fatty streaks within the intima. Ultimately, due to inefficient clearance of these apoptotic cells, secondary necrosis of the macrophages leads to a growing necrotic core within the atherosclerotic plaque. This growing pool of cellular debris and pro-inflammatory cytokines propagates apoptotic signaling, leading to a significantly increased number of apoptotic cells within the atherosclerotic plaque and in the immediately surrounding tissue. Having demonstrated that ArFP was capable of detecting apoptotic cells in a vascular environment (see Chapter 4), it is reasonable to expect that this construct would be able to identify atherosclerotic lesions based on the presence of increased apoptosis.

To explore the potential application of ArFP toward atherosclerosis detection, the apolipoprotein E-deficient (ApoE−/−) strain of mouse was utilized. It has been shown that normal, or wildtype (wt), mice will not develop atherosclerotic lesions unless fed a high-cholesterol, Western-type diet for extended periods of time [264]. In comparison to humans, wt mice exhibit significantly lower total serum concentrations (85 mg/dL in mice compared to the “normal” cholesterol concentration of ≤200 mg/dL for humans). Additionally, a large portion of this cholesterol in mice is carried by high-density lipoprotein (HDL). To generate a mouse model of atherosclerosis, the ApoE protein was inactivated in C57BL/6 mice. This glycoprotein, synthesized mainly in the liver, is a ligand for receptors responsible for clearing very low-density lipoprotein (VLDL) from the blood. By knocking out the gene for this ligand, ApoE−/− mice that are fed a standard chow diet exhibit total cholesterol levels of >500 mg/dL, a value that can be nearly quadrupled through the use of a Western diet [264]. Importantly, by 32 weeks of age on a
standard chow diet, these ApoE−/− mice develop fibrous atherosclerotic plaques in their major arteries, including the aorta.

To test the ability of ArFP to detect atherosclerosis, the aorta of an ApoE−/− mouse at 32 weeks of age was removed. After washing, the aorta was incubated with the ArFP construct. The sample was then washed again to remove any unbound fusion, and bioluminescence was recorded following addition of the substrate coelenterazine. As a control, the aorta of an age-matched wt C57BL/6 mouse was also collected and treated using the same protocol. As shown, the recorded bioluminescence signal from the aorta collected from the ApoE−/− mouse was significantly increased compared to the signal from the tissues of the C57BL/6 mouse (Figure 5.1). Although there is a high amount of background signal from the non-atherosclerotic tissue, these results are similar to findings observed with radio-labeled Annexin V (using ⁹⁹Technetium), and could be further improved through optimization of the binding and imaging protocol [265].

Another potential avenue to expand the application of the developed Annexin V fusion protein (AFP) platform is to replace the bioluminescent reporter. Specifically, by incorporating other photoproteins or bioluminescent enzymes, the characteristics of the resulting AFP can be modified. Preliminary work on three potential modifications has already been performed, and investigates the possibility of generating fusions of Annexin V with aequorin, *Gaussia* luciferase (GLuc), and a truncated variant of *Vargula* luciferase (tVLuc) (Figure 5.2). Although RLuc8, aequorin, and GLuc all generate bioluminescence through oxidative decarboxylation of the same substrate (coelenterazine), each of these additional potential fusion
proteins exhibits unique beneficial characteristics. Using site-directed mutagenesis, a large number of aequorin mutants have been generated that exhibit a range of bioluminescent characteristics, including shifted wavelengths of maximum emission and unique emission kinetic profiles [266]. These mutants have also been combined with a number of modified coelenterazine substrates to further expand the combinations of spectral and kinetic profiles available. By generating a fusion protein of Annexin V with aequorin, these tailored bioluminescence emission

**Figure 5.1:** Comparison of bioluminescence recorded from ArFP incubated with aorta samples from either wtC57BL/6 mouse (left) or ApoE⁻/⁻ mouse (right). To ensure valid comparison of data, the signals were normalized to the mass of the collected tissue.

**Figure 5.2:** Expression of additional Annexin V fusion proteins (AFPs). (a) SDS-PAGE analysis of three AFPs: Annexin V-Renilla luciferase (ArFP), Annexin V-aequorin (AqFP), and Annexin V-Gaussia luciferase (AgFP). Calculated molecular weights for each component and fusion protein provided. (b) Western blot analysis of ArFP and components using anti-Annexin V monoclonal mouse antibody (left) or anti-RLuc rabbit polyclonal antibody (right). (c) Western blot analysis of AqFP and components using anti-Annexin V monoclonal mouse antibody (left) or anti-aequorin rabbit polyclonal antibody (right). (d) Western blot analysis of AgFP and components using anti-Annexin V monoclonal mouse antibody (left) or anti-GLuc rabbit polyclonal antibody (right).
A

Annexin V: ~34.7 kDa
Rluc: ~37.8 kDa
Aequorin: ~25.5 kDa
Gluc: ~24.8 kDa

Annexin V/Renilla luciferase (ArFP): ~63.7 kDa
Annexin V/Aequorin (AqFP): ~58.2 kDa
Annexin V/Gaussia luciferase (AgFP): ~56.9 kDa

B
C
D

Ladder  Annexin V  Rluc  ArFP  Ladder  Annexin V  Rluc  ArFP  Ladder  Annexin V  Aequorin  AqFP  Ladder  Annexin V  Aequorin  AqFP  Ladder  Annexin V  Gluc  AgFP  Ladder  Annexin V  Gluc  AgFP

Anti-Annexin Ms mAb  Anti-Rluc Rb pAb  Anti-Aequorin Rb pAb  Anti-Annexin Ms mAb  Anti-Rluc Rb pAb  Anti-Aequorin Rb pAb  Anti-Annexin Ms mAb  Anti-Rluc Rb pAb  Anti-Aequorin Rb pAb
characteristics can be incorporated into the apoptosis-detecting construct. GLuc offers additional benefits for inclusion into an AFP. Although the emission characteristics are similar to RLuc8, the GLuc protein is much smaller. This reduction is an important characteristic when considering applications in vivo in which the fusion protein may need to pass between tissues (such as into the intima of an atherosclerotic vessel from the lumen). Finally, unlike aequorin, RLuc, or GLuc, the inclusion of tVLuc would confer unique kinetic bioluminescence characteristics to the fusion protein. Specifically, with an emission half-life of over two hours, tVLuc exhibits glow-type bioluminescence [267]. This unique characteristic could thus provide a means to perform time-resolved studies of apoptosis both in vitro as well as in vivo, providing important insight into various conditions including atherosclerosis. Preliminary work has been performed on the expression and purification of these additional fusion proteins. With the exception of the fusion of Annexin V and tVLuc, each of these fusions demonstrates a high degree of purity following expression and isolation (Figure 5.2a), as well as reactivity with antibodies specific for each reporter (Figure 5.2b-d). As the isolation and characterization of these constructs continues, novel applications for apoptosis detection, specifically related to atherosclerosis, will be explored.
Appendix A

Discrete Event Simulation Model of Sudden Cardiac Death
Supporting Information

A.1 Structure of the DES Model

The discrete event simulation (DES) model of sudden cardiac death (SCD) was developed utilizing the SimEvents simulation engine, a component of the MATLAB 2011b family of software tools (Mathworks, Natick, MA). Independent chains of events for the formation and progression of plaques with varying morphology were developed. These chains represent thin-cap fibroatheromas (Figure A1), medium-cap fibroatheromas (Figure A2), and thick-cap atheromas (Figure A3). An additional block of events in the DES model was utilized to determine the likelihood of SCD following arrhythmias caused by the formation and progression of atheromas (see Chapter 2, Figure A4).

A.2 Parameters of the Model

Formation and growth of atheromas were simulated utilizing the Weibull distribution (see Chapter 2). The parameters of this distribution were varied between each type of atheroma (thin-cap, medium-cap, or thick-cap). Table A1 provides exact values for all of these simulation parameters in addition to values determined for growth rates, threshold sizes, and average growth times.
Figure A1: Structure of the SCD-Predict model for thin-cap atheroma formation and growth.
Figure A2: Structure of the SCD-Predict model for medium-cap atheroma formation and growth.
Figure A3: Structure of the SCD-Predict model for thick-cap atheroma formation and growth.
Figure A4: Block of the SCD-Predict model simulating the likelihood of SCD for each arrhythmia.
Table A1: Parameters of the DES Model.

<table>
<thead>
<tr>
<th></th>
<th>Thin-Cap Atheroma</th>
<th>Clot from Thin-Cap Atheroma</th>
<th>Medium-Cap Atheroma</th>
<th>Clot from Medium Cap Atheroma</th>
<th>Thick-Cap Atheroma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-Generation Time Distribution</td>
<td>Weibull: Scale: 15 Shape: 1.5</td>
<td>Forced by thin-cap atheroma rupture</td>
<td>Weibull: Scale: 22 Shape: 3</td>
<td>Forced by medium-cap atheroma reaching threshold size</td>
<td>Weibull: Scale: 9 Shape: 3</td>
</tr>
<tr>
<td>Initial Size Distribution</td>
<td>Weibull: Scale: 0.36 Shape: 4.5</td>
<td>Weibull: Scale: 0.6 Shape: 10</td>
<td>Weibull: Scale: 0.36 Shape: 4.5</td>
<td>Weibull: Scale: 0.5 Shape: 10</td>
<td>Weibull: Scale: 0.36 Shape: 4.5</td>
</tr>
<tr>
<td>1/Growth Rate</td>
<td>12 years</td>
<td>1.5 years</td>
<td>25 years</td>
<td>5 years</td>
<td>19 years</td>
</tr>
<tr>
<td>Threshold Size</td>
<td>0.6</td>
<td>0.75</td>
<td>0.5</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>Average Growth Time</td>
<td>12 years</td>
<td>8 months</td>
<td>14.5 years</td>
<td>5.5 years</td>
<td>32 years</td>
</tr>
<tr>
<td>Probability to Survive First Arrhythmia</td>
<td>0.22</td>
<td></td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
</tr>
</tbody>
</table>
Appendix B

Identification of a Subpopulation in Youth with Accelerated Atherosclerosis Supporting Information

B.1 Univariate Analysis of CVD Risk Factors

Using the identified Discovery Cohort of 739 subjects from the PDAY study population of 2,651 individuals (see Chapter 3), an analysis of the correlation of measured CVD risk factors with observed numbers of coronary raised lesions was performed. The subjects of the Discovery Cohort were grouped according to the observed number of raised lesions in their RCA, and the mean values for the measured risk factors were determined for each group. These values were then divided by the average of the corresponding risk factor as observed in the entire Discovery Cohort, resulting in a value that would indicate if a specific risk factor was elevated (values greater than 1.0) or decreased (values less than 1.0) for individuals with a certain number of raised lesions compared to the cohort population as a whole (Figure B1a). Interestingly, the mean risk factor values for the group with the highest number of raised lesions never exceeds a 40% increase from the population average, and does not represent the group with the greatest elevation of risk factors. The greatest absolute increase in any risk factor for a given group is that of serum thiocyanate for the group of subjects with six coronary raised lesions. However, this increase seems inconsistently high when compared to the changes in serum thiocyanate for other similar groups (those with five or seven raised lesions demonstrate a significantly smaller increase in this risk factor, Figure B1a).

Next, differences in individual risk factors were compared between the high-risk subset of 93 subjects of the Discovery Cohort and those individuals within this cohort
presenting zero raised lesions (Figure B1b). As demonstrated, the zero-lesion group consistently demonstrates mean risk factor values lower than the average of the entire Discovery Cohort. However, even for high-risk subset, no risk factor exceeded a 20% increase from the corresponding average of the entire cohort population, demonstrating an inability of individual risk factors to distinguish subjects belonging to the identified high-risk subset of the Discovery Cohort.

**Figure B1**: Comparison of the values of the known CVD risk factors for the groups of subjects with various numbers of raised lesions. (a) Comparison of individual known risk factors of CVD for subjects of the Discovery Cohort grouped by number of raised lesions in their RCA. Values given as average risk factor value for the group divided by the average value of the Discovery Cohort population. Values greater than 1.0 indicate risk factors elevated for a given group, while values less than 1.0 indicate risk factors decreased for the given group. (b) Comparison of individual known risk factors of CVD for subjects of the Discovery Cohort with zero raised lesions and those belonging to the identified 93-subject high-risk subset (see Chapter 3).
Finally, to determine if any statistical dependence between the measured levels of individual risk factors and the observed number of coronary raised lesions existed, the correlation coefficients between these values were calculated. The members of the Discovery Cohort were first separated according to their age, generating a ‘younger’ group of subjects from ages 15-24, and an ‘older’ group from ages 25-34. However, the calculated correlation coefficients for each of these age groups as well as the entire Discovery Cohort population demonstrate weak relationships between the individual risk factors and the observed number of raised lesions. From this analysis, the highest calculated correlation coefficient was observed between subject age and number of raised lesions in the entire Discovery Cohort, but resulted in a value of only 0.26 (Figure B2a). Following these results, the Discovery Cohort was reorganized into the previously identified 93-subject high-risk subset and the remaining 646 subjects of the cohort population. Correlation coefficients between individual risk factors and the observed number of raised lesions again demonstrate marginal relationships aside from age for the high-risk subset of the Discovery Cohort (Figure B2b). However, it should be noted that a number of calculated correlation coefficients were not found to be significant (p < 0.05). Correlation coefficients not meeting this criterion are indicated with an asterisk (*). Complete matrices of correlation coefficients between individual risk factors and the observed number of raised lesions are given in Table B1-B5, while the calculated p-values for the correlation coefficients determined for the entire Discovery Cohort and the high-risk subset of 93 subjects can be found in Table B6 and B7, respectively.
B.2 Generalized Linear Regression Modeling

B.2.1 Description of the Models

The Wilkinson notation (standard for GLM) was used to describe the factors present in the formula for each model. Thus, the term “1” in the model equation represents the intercept term, and $x_i \times x_j$ includes $x_i$, $x_j$, and $x_i$ multiplied by $x_j$. The interaction term
$x_i x_j$ includes only $x_i$ multiplied by $x_j$, and $x_i^2$ indicates a pure quadratic term. For the following GLM analyses, the terms are identified as follows:

- $x_1$: age/mean age
- $x_2$: race/mean race
- $x_3$: sex/mean sex
- $x_4$: cholesterol/mean cholesterol
- $x_5$: HDL/mean HDL
- $x_6$: LDL/mean LDL
- $x_7$: thiocyanate/mean thiocyanate
- $x_8$: (cholesterol/HDL)/mean (cholesterol/HDL)
- $x_9$: glycosylated hemoglobin/mean glycosylated hemoglobin
- $x_{10}$: BMI/mean BMI

Note here that the variables of the models are normalized by their mean values in order to allow for comparison of similar terms in models for different cohorts and subsets. Therefore, the terms of the models described here must be divided by the mean value of their respective variable to generate the coefficients in the equations describing these models (see Eqs. 3.1-3.4 in Chapter 3). For these models, the response variable “$y$” represents the number of raised lesions in the RCA of the subject. Additionally, to account for non-quantifiable information (specifically race and sex), the following values were assigned:

- Race: white = 1     black = 2
- Sex: female = 1     male = 2
B.2.2 Single Population GLM for the Discovery Cohort (739 Subjects)

The equation generated by the GLM for this group is as follows:

\[ \ln(y) \sim \text{[Linear formula with 27 terms in 10 predictors]} \]  

(B1)

Distribution: Poisson

Estimated Coefficients and model statistics given in Table B8.

B.2.3 Two-Group Approach GLM for the Low-Risk Subset of the Discovery Cohort (646 Subjects)

The equation generated by the GLM for the low-risk subset of the Discovery Cohort is as follows:

\[ \ln(y) \sim \text{[Linear formula with 25 terms in 9 predictors]} \]  

(B2)

Distribution: Poisson

Estimated Coefficients and model statistics given in Table B9.

B.2.4 Two-Group Approach GLM for the High-Risk Subset of the Discovery Cohort (93 Subjects)

The equation generated by the GLM for the high-risk subset of the Discovery Cohort is as follows:

\[ \ln(y) \sim 1 + x_1 + x_2 \times x_3 \]  

(B3)

Distribution: Poisson

Estimated Coefficients and model statistics given in Table B10.

Using the determined coefficients, the number of raised lesions in the RCA of subjects belonging to the high-risk subset of the Discovery Cohort can be determined using the following equation:

\[ \ln(N) = 0.564 + 0.077(\text{age}) + 0.734(\text{race})(\text{sex}) - 1.327(\text{race}) - 1.025(\text{sex}) \]  

(B4)
B.2.5 Two-Group Approach GLM for the High-Risk subset of the Verification Cohort (250 Subjects)

The equation generated by the GLM for the high-risk subset of the Verification Cohort is as follows:

\[ \ln(y) \sim 1 + x_1 \]  
(B5)

Distribution: Poisson

Estimated Coefficients and model statistics given in Table B11.

B.2.6 Two-Group Approach GLM for the High-Risk Subset of the Confirmation Cohort (250 Subjects)

The equation generated by the GLM for the high-risk subset of the Verification Cohort is as follows:

\[ \ln(y) \sim 1 + x_1 + x_6 \]  
(B6)

Distribution: Poisson

Estimated Coefficients and model statistics given in Table B12.

B.3 Tables of Correlation Coefficients and p-Values

Note that for Tables B1-B7, the term “Chol/HDL represents the ratio of total serum cholesterol to HDL.

**Table B1**: Correlation Coefficients. Discovery Cohort, all 739 Subjects. Significant Correlations (p<0.05) Highlighted in Yellow.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Race</th>
<th>Sex</th>
<th>Cholesterol</th>
<th>HDL</th>
<th>LDL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1</td>
<td>-0.07</td>
<td>-0.07</td>
<td>0.19</td>
<td>0.09</td>
<td>0.17</td>
</tr>
<tr>
<td>Race</td>
<td>-0.07</td>
<td>1</td>
<td>0.03</td>
<td>-0.09</td>
<td>0.13</td>
<td>-0.15</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.07</td>
<td>0.03</td>
<td>1</td>
<td>0.01</td>
<td>-0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.19</td>
<td>-0.09</td>
<td>0.01</td>
<td>1</td>
<td>0.4</td>
<td>0.93</td>
</tr>
<tr>
<td>HDL</td>
<td>0.09</td>
<td>0.13</td>
<td>-0.06</td>
<td>0.4</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>LDL</td>
<td>0.17</td>
<td>-0.15</td>
<td>0.04</td>
<td>0.93</td>
<td>0.04</td>
<td>1</td>
</tr>
</tbody>
</table>
Table B2: Correlation Coefficients. Discovery Cohort, 340 Subjects Aged 15-24
Significant Correlations (p<0.05) Highlighted in Yellow.

<table>
<thead>
<tr>
<th></th>
<th>Thiocyanate</th>
<th>Chol/HDL</th>
<th>Glycosylated Hemoglobin</th>
<th>BMI</th>
<th>Number of Raised Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.27</td>
<td>0.11</td>
<td>0</td>
<td>0.09</td>
<td>0.26</td>
</tr>
<tr>
<td>Race</td>
<td>-0.09</td>
<td>-0.18</td>
<td>0.13</td>
<td>0.03</td>
<td>-0.05</td>
</tr>
<tr>
<td>Sex</td>
<td>0.02</td>
<td>0.08</td>
<td>0.04</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.15</td>
<td>0.29</td>
<td>0.02</td>
<td>0.09</td>
<td>0.2</td>
</tr>
<tr>
<td>HDL</td>
<td>0.11</td>
<td>-0.64</td>
<td>-0.07</td>
<td>-0.06</td>
<td>-0.03</td>
</tr>
<tr>
<td>LDL</td>
<td>0.12</td>
<td>0.57</td>
<td>0.05</td>
<td>0.12</td>
<td>0.23</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>1</td>
<td>0.03</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>0.03</td>
<td>1</td>
<td>0.08</td>
<td>0.12</td>
<td>0.16</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>-0.06</td>
<td>0.08</td>
<td>1</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>BMI</td>
<td>-0.02</td>
<td>0.12</td>
<td>0.07</td>
<td>1</td>
<td>0.13</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>0.14</td>
<td>0.16</td>
<td>0.08</td>
<td>0.13</td>
<td>1</td>
</tr>
</tbody>
</table>
**Table B3**: Correlation Coefficients. Discovery Cohort, 399 Subjects Aged 25-34
Significant Correlations (p<0.05) Highlighted in Yellow.

<table>
<thead>
<tr>
<th></th>
<th>Thiocyanate</th>
<th>Chol/HDL</th>
<th>Glycosylated Hemoglobin</th>
<th>BMI</th>
<th>Number of Raised Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.19</td>
<td>-0.02</td>
<td>-0.04</td>
<td>0.11</td>
<td>0.13</td>
</tr>
<tr>
<td>Race</td>
<td>-0.07</td>
<td>-0.15</td>
<td>0.22</td>
<td>0.14</td>
<td>-0.01</td>
</tr>
<tr>
<td>Sex</td>
<td>0.06</td>
<td>0.01</td>
<td>0.13</td>
<td>0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.12</td>
<td>0.21</td>
<td>0.04</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>HDL</td>
<td>0.15</td>
<td>-0.64</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.13</td>
</tr>
<tr>
<td>LDL</td>
<td>0.06</td>
<td>0.52</td>
<td>0.06</td>
<td>0.08</td>
<td>0.15</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>1</td>
<td>-0.06</td>
<td>-0.04</td>
<td>0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>-0.06</td>
<td>1</td>
<td>0.09</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>-0.04</td>
<td>0.09</td>
<td>1</td>
<td>0.13</td>
<td>0.05</td>
</tr>
<tr>
<td>BMI</td>
<td>0.01</td>
<td>0.06</td>
<td>0.13</td>
<td>1</td>
<td>0.17</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>0.15</td>
<td>0.05</td>
<td>0.05</td>
<td>0.17</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Race</th>
<th>Sex</th>
<th>Cholesterol</th>
<th>HDL</th>
<th>LDL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1</td>
<td>-0.06</td>
<td>-0.01</td>
<td>0.07</td>
<td>-0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Race</td>
<td>-0.06</td>
<td>1</td>
<td>0</td>
<td>-0.14</td>
<td>0.14</td>
<td>-0.2</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.01</td>
<td>0</td>
<td>1</td>
<td>0.03</td>
<td>-0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.07</td>
<td>-0.14</td>
<td>0.03</td>
<td>1</td>
<td>0.36</td>
<td>0.93</td>
</tr>
<tr>
<td>HDL</td>
<td>-0.01</td>
<td>0.14</td>
<td>-0.11</td>
<td>0.36</td>
<td>1</td>
<td>-0.01</td>
</tr>
<tr>
<td>LDL</td>
<td>0.07</td>
<td>-0.2</td>
<td>0.07</td>
<td>0.93</td>
<td>-0.01</td>
<td>1</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>0.18</td>
<td>-0.09</td>
<td>0.02</td>
<td>0.12</td>
<td>0.06</td>
<td>0.11</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>0.13</td>
<td>-0.19</td>
<td>0.13</td>
<td>0.32</td>
<td>-0.65</td>
<td>0.6</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>-0.09</td>
<td>0.08</td>
<td>0.01</td>
<td>0</td>
<td>-0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>BMI</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.09</td>
<td>0.09</td>
<td>-0.12</td>
<td>0.14</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>0.16</td>
<td>-0.06</td>
<td>0.1</td>
<td>0.18</td>
<td>-0.1</td>
<td>0.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Thiocyanate</th>
<th>Chol/HDL</th>
<th>Glycosylated Hemoglobin</th>
<th>BMI</th>
<th>Number of Raised Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.18</td>
<td>0.13</td>
<td>-0.09</td>
<td>0.03</td>
<td>0.16</td>
</tr>
<tr>
<td>Race</td>
<td>-0.09</td>
<td>-0.19</td>
<td>0.08</td>
<td>-0.05</td>
<td>-0.06</td>
</tr>
<tr>
<td>Sex</td>
<td>0.02</td>
<td>0.13</td>
<td>0.01</td>
<td>0.09</td>
<td>0.1</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.12</td>
<td>0.32</td>
<td>0</td>
<td>0.09</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>HDL</td>
<td>LDL</td>
<td>Thiocyanate</td>
<td>Chol/HDL</td>
<td>Glycosylated Hemoglobin</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>HDL</td>
<td>0.06</td>
<td>-0.65</td>
<td>-0.09</td>
<td>-0.12</td>
<td>-0.1</td>
</tr>
<tr>
<td>LDL</td>
<td>0.11</td>
<td>0.6</td>
<td>0.04</td>
<td>0.14</td>
<td>0.23</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>0.06</td>
<td>0.1</td>
<td>0.07</td>
<td>0.15</td>
<td>0.18</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>0.06</td>
<td>0.1</td>
<td>0.07</td>
<td>0.15</td>
<td>0.18</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>-0.08</td>
<td>0.07</td>
<td>0.04</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>-0.06</td>
<td>0.15</td>
<td>0.04</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>0.1</td>
<td>0.18</td>
<td>0.08</td>
<td>0.11</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table B4:** Correlation Coefficients. Discovery Cohort, 646 Subjects of the Low-Risk Subset. Significant Correlations (p<0.05) Highlighted in Yellow.
<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Race</th>
<th>Sex</th>
<th>Cholesterol</th>
<th>HDL</th>
<th>LDL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1</td>
<td>-0.016</td>
<td>-0.187</td>
<td>0.121</td>
<td>-0.131</td>
<td>0.175</td>
</tr>
<tr>
<td>Race</td>
<td>-0.016</td>
<td>1</td>
<td>0.023</td>
<td>-0.082</td>
<td>0.083</td>
<td>-0.116</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.188</td>
<td>0.023</td>
<td>1</td>
<td>0.132</td>
<td>0.117</td>
<td>0.101</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.121</td>
<td>-0.082</td>
<td>0.132</td>
<td>1</td>
<td>0.349</td>
<td>0.946</td>
</tr>
<tr>
<td>HDL</td>
<td>-0.131</td>
<td>0.083</td>
<td>0.117</td>
<td>0.349</td>
<td>1</td>
<td>0.027</td>
</tr>
<tr>
<td>LDL</td>
<td>0.175</td>
<td>-0.116</td>
<td>0.101</td>
<td>0.946</td>
<td>0.027</td>
<td>1</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>0.217</td>
<td>-0.034</td>
<td>0.17</td>
<td>0.035</td>
<td>0.087</td>
<td>0.007</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>0.091</td>
<td>-0.164</td>
<td>-0.027</td>
<td>0.267</td>
<td>-0.649</td>
<td>0.509</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>0.009</td>
<td>-0.088</td>
<td>-0.037</td>
<td>0.034</td>
<td>-0.118</td>
<td>0.077</td>
</tr>
<tr>
<td>BMI</td>
<td>-0.003</td>
<td>-0.079</td>
<td>0.113</td>
<td>-0.069</td>
<td>-0.122</td>
<td>-0.032</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>0.493</td>
<td>-0.006</td>
<td>-0.073</td>
<td>0.174</td>
<td>-0.123</td>
<td>0.228</td>
</tr>
</tbody>
</table>

**Table B5: Correlation Coefficients. Discovery Cohort, 93 Subjects of the High-Risk Subset. Significant Correlations (p<0.05) Highlighted in Yellow.**

<table>
<thead>
<tr>
<th></th>
<th>Thiocyanate</th>
<th>Chol/HDL</th>
<th>Glycosylated Hemoglobin</th>
<th>BMI</th>
<th>Number of Raised Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.217</td>
<td>0.091</td>
<td>0.009</td>
<td>-0.003</td>
<td>0.493</td>
</tr>
<tr>
<td>Race</td>
<td>-0.034</td>
<td>-0.164</td>
<td>-0.088</td>
<td>-0.079</td>
<td>-0.006</td>
</tr>
<tr>
<td>Sex</td>
<td>0.17</td>
<td>-0.027</td>
<td>-0.037</td>
<td>0.113</td>
<td>-0.073</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.035</td>
<td>0.267</td>
<td>0.034</td>
<td>-0.069</td>
<td>0.174</td>
</tr>
<tr>
<td>HDL</td>
<td>0.087</td>
<td>-0.649</td>
<td>-0.118</td>
<td>-0.122</td>
<td>-0.123</td>
</tr>
<tr>
<td>LDL</td>
<td>0.007</td>
<td>0.509</td>
<td>0.077</td>
<td>-0.031</td>
<td>0.228</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>1</td>
<td>-0.108</td>
<td>-0.174</td>
<td>-0.035</td>
<td>0.193</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>-0.108</td>
<td>1</td>
<td>0.127</td>
<td>0.053</td>
<td>0.167</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>-0.174</td>
<td>0.127</td>
<td>1</td>
<td>0.032</td>
<td>0.045</td>
</tr>
<tr>
<td>BMI</td>
<td>-0.035</td>
<td>0.053</td>
<td>0.032</td>
<td>1</td>
<td>0.051</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>0.193</td>
<td>0.167</td>
<td>0.045</td>
<td>0.051</td>
<td>1</td>
</tr>
</tbody>
</table>
Table B6: p-Values. Discovery Cohort, all 739 Subjects.

<table>
<thead>
<tr>
<th>Age</th>
<th>Race</th>
<th>Sex</th>
<th>Cholesterol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.00E+00</td>
<td>4.46E-02</td>
<td>7.53E-02</td>
</tr>
<tr>
<td>Race</td>
<td>4.46E-02</td>
<td>1.00E+00</td>
<td>4.95E-01</td>
</tr>
<tr>
<td>Sex</td>
<td>7.53E-02</td>
<td>4.95E-01</td>
<td>1.00E+00</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>2.19E-07</td>
<td>1.35E-02</td>
<td>7.51E-01</td>
</tr>
<tr>
<td>HDL</td>
<td>1.04E-02</td>
<td>2.35E-04</td>
<td>1.30E-01</td>
</tr>
<tr>
<td>LDL</td>
<td>3.92E-06</td>
<td>2.75E-05</td>
<td>3.39E-01</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>7.25E-14</td>
<td>1.36E-02</td>
<td>6.42E-01</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>3.88E-03</td>
<td>1.32E-06</td>
<td>3.36E-02</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>9.05E-01</td>
<td>5.39E-04</td>
<td>2.44E-01</td>
</tr>
<tr>
<td>BMI</td>
<td>1.52E-02</td>
<td>4.42E-01</td>
<td>5.36E-02</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>2.96E-13</td>
<td>1.62E-01</td>
<td>4.37E-02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HDL</th>
<th>LDL</th>
<th>Thiocyanate</th>
<th>Chol/HDL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.04E-02</td>
<td>3.92E-06</td>
<td>7.25E-14</td>
</tr>
<tr>
<td>Race</td>
<td>2.35E-04</td>
<td>2.75E-05</td>
<td>1.36E-02</td>
</tr>
<tr>
<td>Sex</td>
<td>1.30E-01</td>
<td>3.39E-01</td>
<td>6.42E-01</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>1.69E-30</td>
<td>0.00E+00</td>
<td>5.61E-05</td>
</tr>
<tr>
<td>HDL</td>
<td>1.00E+00</td>
<td>2.81E-01</td>
<td>2.71E-03</td>
</tr>
<tr>
<td>LDL</td>
<td>2.81E-01</td>
<td>1.00E+00</td>
<td>1.45E-03</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>2.71E-03</td>
<td>1.45E-03</td>
<td>1.00E+00</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>3.81E-85</td>
<td>5.56E-66</td>
<td>4.07E-01</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>5.90E-02</td>
<td>1.69E-01</td>
<td>1.13E-01</td>
</tr>
<tr>
<td>BMI</td>
<td>1.19E-01</td>
<td>8.64E-04</td>
<td>6.48E-01</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>3.90E-01</td>
<td>1.42E-10</td>
<td>8.56E-05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Glycosylated Hemoglobin</th>
<th>BMI</th>
<th>Number of Raised Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>9.05E-01</td>
<td>1.52E-02</td>
</tr>
<tr>
<td>Race</td>
<td>5.39E-04</td>
<td>4.42E-01</td>
</tr>
<tr>
<td>Sex</td>
<td>2.44E-01</td>
<td>5.36E-02</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>5.72E-01</td>
<td>1.36E-02</td>
</tr>
<tr>
<td>HDL</td>
<td>5.90E-02</td>
<td>1.19E-01</td>
</tr>
<tr>
<td>LDL</td>
<td>1.69E-01</td>
<td>8.64E-04</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>Race</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Age</td>
<td>1.00E+00</td>
<td>8.78E-01</td>
</tr>
<tr>
<td>Race</td>
<td>8.78E-01</td>
<td>1.00E+00</td>
</tr>
<tr>
<td>Sex</td>
<td>7.17E-02</td>
<td>8.29E-01</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>2.47E-01</td>
<td>4.36E-01</td>
</tr>
<tr>
<td>HDL</td>
<td>2.12E-01</td>
<td>4.29E-01</td>
</tr>
<tr>
<td>LDL</td>
<td>9.42E-02</td>
<td>2.69E-01</td>
</tr>
<tr>
<td>Thiocyanate</td>
<td>3.69E-02</td>
<td>7.47E-01</td>
</tr>
<tr>
<td>Chol/HDL</td>
<td>3.87E-01</td>
<td>1.15E-01</td>
</tr>
<tr>
<td>Glycosylated Hemoglobin</td>
<td>9.32E-01</td>
<td>4.00E-01</td>
</tr>
<tr>
<td>BMI</td>
<td>9.77E-01</td>
<td>4.53E-01</td>
</tr>
<tr>
<td>Number of Raised Lesions</td>
<td>5.28E-07</td>
<td>9.55E-01</td>
</tr>
</tbody>
</table>

Table B7: p-Values. Discovery Cohort, 93 Subjects of the High-Risk Subset.
### Table B8: Estimated Coefficients and Model Statistics, Single Population Model of the Discovery Cohort (739 Subjects).

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-14.646</td>
<td>4.1215</td>
<td>-3.554</td>
<td>3.800E-04</td>
</tr>
<tr>
<td>x1</td>
<td>10.612</td>
<td>2.2474</td>
<td>4.722</td>
<td>2.338E-06</td>
</tr>
<tr>
<td>x2</td>
<td>-0.422</td>
<td>1.6763</td>
<td>-0.252</td>
<td>8.013E-01</td>
</tr>
<tr>
<td>x3</td>
<td>-0.083</td>
<td>2.2929</td>
<td>-0.036</td>
<td>9.710E-01</td>
</tr>
<tr>
<td>x4</td>
<td>0.000</td>
<td>0.0000</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>x5</td>
<td>7.587</td>
<td>3.1003</td>
<td>2.447</td>
<td>1.440E-02</td>
</tr>
<tr>
<td>x6</td>
<td>0.631</td>
<td>2.5774</td>
<td>0.245</td>
<td>8.067E-01</td>
</tr>
<tr>
<td>x7</td>
<td>0.199</td>
<td>0.6884</td>
<td>0.290</td>
<td>7.720E-01</td>
</tr>
<tr>
<td>x8</td>
<td>5.084</td>
<td>2.6429</td>
<td>1.924</td>
<td>5.438E-02</td>
</tr>
<tr>
<td>x9</td>
<td>-1.573</td>
<td>1.3381</td>
<td>-1.175</td>
<td>2.399E-01</td>
</tr>
<tr>
<td>x10</td>
<td>0.751</td>
<td>0.2643</td>
<td>2.839</td>
<td>4.520E-03</td>
</tr>
<tr>
<td>x1:x5</td>
<td>-5.310</td>
<td>1.3160</td>
<td>-4.035</td>
<td>5.452E-05</td>
</tr>
<tr>
<td>x1:x8</td>
<td>-2.009</td>
<td>1.0062</td>
<td>-1.996</td>
<td>4.591E-02</td>
</tr>
<tr>
<td>x2:x3</td>
<td>2.501</td>
<td>0.9831</td>
<td>2.545</td>
<td>1.094E-02</td>
</tr>
<tr>
<td>x2:x5</td>
<td>-1.961</td>
<td>0.9589</td>
<td>-2.045</td>
<td>4.089E-02</td>
</tr>
<tr>
<td>x2:x6</td>
<td>1.551</td>
<td>0.7328</td>
<td>2.117</td>
<td>3.429E-02</td>
</tr>
<tr>
<td>x3:x8</td>
<td>-1.788</td>
<td>0.9175</td>
<td>-1.949</td>
<td>5.134E-02</td>
</tr>
<tr>
<td>x3:x5</td>
<td>-2.688</td>
<td>1.3652</td>
<td>-1.969</td>
<td>4.897E-02</td>
</tr>
</tbody>
</table>
Table B9: Estimated Coefficients and Model Statistics, Two-Group Model of the Discovery Cohort, Low-Risk Subset (646 Subjects).

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-4.213</td>
<td>9.0567</td>
<td>-0.465</td>
<td>6.418E-01</td>
</tr>
<tr>
<td>x1</td>
<td>14.046</td>
<td>2.7450</td>
<td>5.117</td>
<td>3.106E-07</td>
</tr>
<tr>
<td>x2</td>
<td>-15.278</td>
<td>10.4240</td>
<td>-1.466</td>
<td>1.427E-01</td>
</tr>
<tr>
<td>x3</td>
<td>-7.166</td>
<td>6.5732</td>
<td>-1.090</td>
<td>2.756E-01</td>
</tr>
<tr>
<td>x5</td>
<td>-0.131</td>
<td>3.3859</td>
<td>-0.039</td>
<td>9.691E-01</td>
</tr>
<tr>
<td>x6</td>
<td>-4.874</td>
<td>3.7811</td>
<td>-1.289</td>
<td>1.974E-01</td>
</tr>
<tr>
<td>x7</td>
<td>13.333</td>
<td>3.3439</td>
<td>3.987</td>
<td>6.685E-05</td>
</tr>
<tr>
<td>x8</td>
<td>-5.255</td>
<td>3.4083</td>
<td>-1.542</td>
<td>1.232E-01</td>
</tr>
<tr>
<td>x9</td>
<td>-4.884</td>
<td>4.0972</td>
<td>-1.192</td>
<td>2.332E-01</td>
</tr>
<tr>
<td>x10</td>
<td>-0.366</td>
<td>1.4293</td>
<td>-0.256</td>
<td>7.978E-01</td>
</tr>
<tr>
<td>x1:x7</td>
<td>-3.925</td>
<td>1.8604</td>
<td>-2.110</td>
<td>3.486E-02</td>
</tr>
<tr>
<td>x2:x3</td>
<td>13.815</td>
<td>9.1264</td>
<td>1.514</td>
<td>1.301E-01</td>
</tr>
<tr>
<td>x2:x6</td>
<td>3.407</td>
<td>1.2792</td>
<td>2.663</td>
<td>7.737E-03</td>
</tr>
<tr>
<td>x2:x8</td>
<td>-3.382</td>
<td>1.2302</td>
<td>-2.750</td>
<td>5.967E-03</td>
</tr>
</tbody>
</table>
Table B10: Estimated Coefficients and Model Statistics, Two-Group Model of the Discovery Cohort, High-Risk Subset (93 Subjects).

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.564</td>
<td>1.0669</td>
<td>0.529</td>
<td>5.971E-01</td>
</tr>
<tr>
<td>x1</td>
<td>1.910</td>
<td>0.4008</td>
<td>4.766</td>
<td>1.877E-06</td>
</tr>
<tr>
<td>x2</td>
<td>-2.025</td>
<td>1.0349</td>
<td>-1.956</td>
<td>5.042E-02</td>
</tr>
<tr>
<td>x3</td>
<td>-1.783</td>
<td>0.9271</td>
<td>-1.923</td>
<td>5.445E-02</td>
</tr>
<tr>
<td>x2:x3</td>
<td>1.949</td>
<td>0.9708</td>
<td>2.008</td>
<td>4.468E-02</td>
</tr>
</tbody>
</table>

Number of Observations 93
Error Degrees of Freedom 88
Dispersion 1
χ²-squared Statistic vs. constant model 27.8 p-value 1.37E-05

R-squared
Ordinary 0.3300
Table B11: Estimated Coefficients and Model Statistics, Two-Group Model of the Verification Cohort, High-Risk Subset (250 Subjects).

<table>
<thead>
<tr>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-1.100</td>
<td>0.2654</td>
<td>-4.145</td>
</tr>
<tr>
<td>x1</td>
<td>1.870</td>
<td>0.2485</td>
<td>7.528</td>
</tr>
</tbody>
</table>

Number of Observations 250
Error Degrees of Freedom 248
Dispersion 1
χ²-squared Statistic vs. constant model 59 p-value 1.58E-14

R-squared
Ordinary 0.2418
Adjusted 0.2388
LLR 0.0661
Deviance 0.7713
Adjusted Generalized 0.2163

Table B12: Estimated Coefficients and Model Statistics, Two-Group Model of the Confirmation Cohort, High-Risk Subset (84 Subjects).

<table>
<thead>
<tr>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.864</td>
<td>0.4615</td>
<td>-1.872</td>
</tr>
<tr>
<td>x1</td>
<td>1.266</td>
<td>0.4340</td>
<td>2.916</td>
</tr>
<tr>
<td>x6</td>
<td>0.39402</td>
<td>0.20709</td>
<td>1.9027</td>
</tr>
</tbody>
</table>

Number of Observations 84
Error Degrees of Freedom 81
Dispersion 1
χ²-squared Statistic vs. constant model 15.8 p-value 3.71E-04

R-squared
Ordinary 0.2053
Adjusted 0.1856
LLR 0.0534
Deviance: 0.8024
Adjusted Generalized: 0.1767

Figure B3: Analysis of the residuals for the GLM analyses of the Discovery Cohort. Residuals generated from the single population (a,b), low-risk subset (c,d), and high-risk subset (e,f). Residuals for these models presented in either histogram (a,c,e) or probability (b,d,f) modes.
Appendix C

Bioluminescence-based Apoptosis Detection Supporting Information

C.1 Apoptosis and Its Detection

C.1.1 Defining Apoptosis

The ability to regulate the division and controlled removal of cells is a hallmark of multicellular organisms. Regulated cell removal is required for a range of diverse physiological roles, such as the remodeling or deletion of structures during development, maintenance of tissue homeostasis, and elimination of potentially dangerous cells such as self-reactive lymphocytes [215, 216]. Today, multiple mechanisms of Programmed Cell Death (PCD) to regulate cell removal have been identified, each a highly structured and tightly regulated process [218, 268, 269]. One of the best characterized and most understood of these PCD processes is that of apoptosis. Two core routes of apoptosis progression have been well established; the cell receptor-mediated extrinsic pathway, and the mitochondria-controlled intrinsic pathway, which results from non-receptor stimuli such as DNA damage or a withdrawal of growth factors (see Figure 4.1 in Chapter 4).

Although these two pathways are initiated by differing factors, a central step in both routes to apoptosis is the activation of caspases (cysteine-dependent aspartyl-specific proteases), a family of proteins that are ubiquitously expressed throughout cell types as inactive zymogens (proenzymes). Once activated via cleavage at specific locations, however, caspases are responsible for cleaving plasma membrane, cytoskeletal, and nuclear proteins; activating cytoplasmic endonucleases that degrade nuclear material; and
activating additional zymogens, creating a caspase cascade that ultimately results in the biochemical and morphological changes associated with apoptosis. Kerr, Currie, and Wyllie first described these characteristic features of apoptosis in the early 1970s, using electron microscopy to define apoptosis as a form of cell death involving nuclear condensation and cytoplasmic shrinkage, accompanied by preservation of membrane integrity [214].

C.1.2 Targeting Apoptosis

In addition to the prognostic features identified by Kerr et al., cells undergoing apoptosis have been shown to rapidly display phosphatidylserine (PS) on the external surface of their cellular membranes. Under healthy cellular conditions, PS is sequestered to the inner-leaf of the plasma membrane. This membrane asymmetry is achieved by proteins known as flippases (aminophospholipid translocases) that actively transport PS and phosphatidylethanolamine to the inner-leaf of the plasma membrane in a unidirectional and ATP-dependent fashion (Figure 4.1 in Chapter 4) [221]. Identified members of this family contain caspase-3 cleavage sites, rendering them inactive upon induction of apoptosis [219]. However, this inactivation is insufficient to account for the observed rapid exposure of PS on the surface of apoptotic cells, due to the relatively slow passive diffusion rates of phospholipids across the membrane [270]. Scramblases are proteins that, unlike flippases, move phospholipids through membranes in a bidirectional and ATP-independent fashion, and are activated by caspase cleavage [222, 223]. Together, these processes work to ensure that, while PS is generally isolated to the inner-leaf of the plasmalemma, apoptosis induction quickly results in displayed PS on the outer surface of the cell.
Drastic changes in composition of the external plasma membrane face have for many years been proposed as an ‘eat-me’ signal, or means by which apoptotic cells can be cleared by neighboring cells or professional phagocytes (such as macrophages or immature dendritic cells) [219, 220, 271]. These changes include variances in surface charge, altered levels of glycoprotein epitope exposure, exposure of the calcium ion binding protein calrecticulin, and, importantly, the previously mentioned exposure of phosphatidylserine [220, 272, 273]. Calculations of phospholipid content of the plasma membrane indicate that PS comprises approximately 10% of the total phospholipid content of the bilayer, while across multiple cell types, less than 5% of this PS is found at the cell surface [221, 270, 274]. The rapid increase in the levels of externalized phosphatidylserine during apoptosis (over a 250-fold increase over healthy levels in some cases) has been reported to participate either directly in phagocyte binding, or through bridge molecules that link PS to phagocyte receptors [219, 220, 272]. Research toward identifying apoptotic cells has taken advantage of this specific, yet universal, morphological change by utilizing a high-affinity binding partner of PS. The protein Annexin V (annexin A5, encoded by human gene ANXA5) is a member of an extensive family of calcium ion- and phospholipid-binding proteins, with members expressed ubiquitously in cells throughout the plant and animal kingdoms [275-277]. In addition to their calcium- and phospholipid-binding properties, annexins are defined by the presence of a conserved structural repeat of approximately 70 amino acid residues that form a highly α-helical domain through which calcium and lipid binding is achieved. Annexin V is an approximately 36 kDa protein, and contains four such domains (domains I-IV) roughly arranged in a disc shape, creating a convex and concave face (Figure C.1a). On
the convex surface of the disc, cationic calcium is bound by each of the annexin cores through interactions with the peptide backbone as well as negatively charged amino acid residues. Subsequently, through bound calcium, Annexin V interacts with, and binds to, the negatively charged head group of phosphatidylserine, tethering the protein to the membrane surface. Early work using exogenously applied, fluorescence (fluorescein isothiocyanate)-labeled Annexin V demonstrated that this specific binding was sufficient to target and identify apoptotic cells \textit{in vitro} [226, 229, 278]. Since that time, protocols for the overexpression of recombinant Annexin V [224, 225] have led to an abundance of progress toward the detection of apoptotic cells via the binding of PS by Annexin V with reporters ranging from radio-labeling (using $^{99m}$Tc, $^{94m}$TC, $^{123}$I, $^{111}$In, $^{18}$F, and others)

\textbf{Figure C1:} (a) Structure of Annexin V (PDB accession number 1ANX) with four repeated annexin core domains colored individually. The convex face of the protein binds PS at the surface of apoptotic cells. (b) Structure of \textit{Renilla} luciferase with bound substrate coelenterazin (PDB accession number 2PSJ). (c) Oxidative decarboxylation of RLuc8 substrate coelenterazin.
fluorescence (FITC, EGFP, Cy5.5, among others) and most recently through the use of bioluminescence.

Previous work toward the detection of apoptosis through the use of bioluminescence has heavily relied upon intracellular expression of bioluminescent reporters, and has examined apoptotic features such as the alteration of cellular ATP levels and caspase activity either through luciferin substrate liberation or split luciferase activity. However, by exploiting the ability of Annexin V to specifically target PS on the surface of apoptotic cells, a new class of bioluminescence-based apoptosis detection sensors can be developed. Renilla luciferase (RLuc) is an approximately 36 kDa protein which is natively expressed by the sea pansy Renilla reniformis. Unlike many of the commonly used beetle luciferases (such as Firefly luciferase), RLuc functions independently of ATP to generate bioluminescence via oxidative decarboxylation of its substrate coelenterazine to coelenteramide (Figure C.1b,c). This property allowed for the generation of an RLuc/Annexin probe which was successful in detection of apoptotic cells in vitro. However, previous work has shown that this wildtype RLuc is quickly rendered inactive in the presence of murine serum, severely limiting its use to applications in vivo. In response to this finding, Loening et al. developed a mutant luciferase (RLuc8) through mutagenesis based on consensus sequence homology resulting in a 200-fold increase in serum stability, as well as a 4-fold increase in light output (Figure C.1b). Utilizing this RLuc8 mutant, an enhanced bioluminescence-based Annexin V apoptosis detection construct was generated for use both in vitro as well as in vivo.
C.2 Fusion Gene Design and Transformation

Genes for both Annexin V and RLuc8 were cloned out of plasmids pPROEX-HTb and pBAD/Myc-HIS A::RLuc8, respectively. Resulting PCR products were visualized on a 1% agarose gel, verifying gene sizes of 969 nucleotides and 993 nucleotides for Annexin V and RLuc8, respectively (Figure C2a). These genes were combined using overlap-extension PCR (Figure C3), and the product was visualized on a 1% agarose gel, verifying a gene size of 1944 nucleotides for the ArFP gene (Figure C2b). This fusion gene was then treated according to the pET-30 Xa/LIC vector kit and inserted into this vector to generate pET-30/ArFP (Figure C4). Following transformation, isolated plasmid was treated with restriction enzymes EcoRI and NdeI to verify the presence of the ArFP gene insert, while the resulting banding pattern was verified using SerialCloner v2.6.1.

C.3 Dynamic Light Scattering (DLS) of Lipid Vesicles

In order to determine the binding affinity of ArFP for PS, lipid vesicles were prepared at a molar ratio of 60:40 POPC:POPS. These vesicles were subsequently downsized via repeated extrusion through a 0.2 µm polycarbonate filter a minimum of 21 times. The size of these vesicles was subsequently determined via dynamic light scattering (DLS) measurements (Figure C5).

C.4 Cell Culture of Caco-2 Cells

Caco-2 heterogeneous human epithelial colorectal adenocarcinoma cells (ATCC® HTB-37™) were obtained from ATCC. These cells were maintained in DMEM, supplemented with 10% fetal bovine serum (FBS), 100 mg/L penicillin, 100 mg/L streptomycin, and 2 mM glutamine.
C.4.1 Apoptosis Induction

Cells to be used for apoptosis induction studies were grown in multiple Corning® T25 Tissue Culture flasks to confluence. To these flasks, 1 µM the known apoptosis inducer staurosporine was added for times ranging from 30 minutes to 6 hours, while untreated cells served as controls. Following incubation, the media was carefully aspirated, and the cells washed with 5 mL PBS. Washed cells were then removed from the growth surface via addition of 5 mL Accutase® cell detachment solution. Detached cells were collected, centrifuged at 150 xg for 5 minutes, resuspended in a solution of Binding Buffer.
Figure C3: Schematic of OE-PCR. 1) Annealing of primers with appropriate LIC (orange) and linker (green) regions to template. 2) Extension of primers. 3) Gene product containing LIC or linker extension. 4) Annealing of primers to product from (3) while template strands are used for additional cycles. 5) Extension of primers. 6) Generation of complete gene with LIC and linker sequences. 7) Annealing of linker sequences and primers. 8) Extension of primers. 9) Generation of complete fusion gene. 10) Amplification.
containing a final concentration of 1 nM ArFP, and incubated for 5 minutes to allow for binding. The cells were then washed three times with 5 mL Binding Buffer. A cell count was performed for each induction condition using a Moxi Z™ Mini Automated Cell Counter (ORFLO Technologies, Ketchum, ID), and equal numbers of cells were transferred in triplicate to a 96-well microtiter plate. Bioluminescence measurements of ArFP were then obtained using a POLARstar® Optima (BMG LABTECH GmbH, Ortenberg, Germany) via injection of 50 µL coelenterazine at a concentration of 2.5 µg/mL. These bioluminescence measurements generated a time-dependent increase in bioluminescence that corresponded to increased incubation with staurosporine (Figure C6a).

**Figure C4**: pET-30/ArFP vector expression system plasmid map.
C.4.2 Cell Metabolic Activity (MTS) Assay

The time-dependent cytotoxicity of staurosporine was studied by MTS (3-(4,5-dimethylthiazol-2-yI)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay in 96-well microtiter plates. Caco-2 cells were transferred, maintained, and grown to confluence in the 96-well microtiter plate. In triplicate, staurosporine was added to a final concentration of 1 µM at various times ranging from 30 minutes to 6 hours, with untreated cells serving as negative controls. Following apoptosis induction, cell proliferation was monitored using the CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega, Madison, WI). From these measurements, a time-dependent decrease in absorbance at 490 nm is observed, corroborating the results observed from bioluminescence measurements of apoptosis-induced cells (Figure C6b).
Figure C6: (a) Bioluminescence response of ArFP incubated with Caco-2 cells after treatment with 1 μM staurosporine for various times. (b) Absorbance at 490 nm of wells containing cells incubated with MTS reagent for 4 hours following treatment with 1 μM staurosporine for various times.
References


173


<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
<th>Details</th>
</tr>
</thead>
</table>


