AWARD NUMBER: W81XWH-20-1-0615

TITLE: Clinical Study of Vascular Plaque Determination for Stroke Risk Assessment

PRINCIPAL INVESTIGATOR: D. Geoffrey Vince, PhD.

CONTRACTING ORGANIZATION: The Cleveland Clinic Foundation

REPORT DATE: October 2022

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
Oct 2022	Annual	15 Sep 2021 - 14 Sep 2022

4. TITLE AND SUBTITLE
Clinical Study of Vascular Plaque Determination for Stroke Risk Assessment

5a. CONTRACT NUMBER
W81XWH-20-1-0615

5b. GRANT NUMBER
PR191955

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
D. Geoffrey Vince
Russell J. Fedewa
Sheronica L. James

E-Mail: vinceg@ccf.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Cleveland Clinic Foundation, The
9500 Euclid Avenue
Cleveland, OH 44195-0001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The high rate of diabetes for the veteran population (~20%) tracks with high rates of obesity and these both have increased complications that include carotid stenosis and risk of stroke. Composition information is not available for carotid stenosis even though it is a known predictor of stroke risk. Our hypothesis is that there exist correlations between the information provided by the **C**ompositional **A**nalysis **S**ystem by **M**achine learning (CASM) algorithm and future outcomes for patients with carotid stenosis. This research effort involves enrolling up to 1500 subjects with significant carotid stenosis. The asymptomatic patients are followed for up to 3 years in order to determine if there are correlations between risk of stroke and the output from CASM. Also, the output of the CASM algorithm will be analyzed for key populations: e.g. diabetic vs non-diabetic and asymptomatic vs symptomatic. A total of **304 subjects have been enrolled** at the two sites with a total of **402 carotid arteries**. Also, **24** follow-up research ultrasound exams have been collected.

15. SUBJECT TERMS

diabetes, atherosclerosis, stroke, carotid stenosis, cerebrovascular accident, ultrasound, spectral analysis, tissue characterization, machine learning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	39	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. INTRODUCTION:

Higher rates of diabetes in the veteran population is a contributing factor in the higher rates of carotid stenosis and subsequent stroke as compared to the general population. Currently, the degree of stenosis is a key determinant in the recommended course of treatment for carotid plaque, but this measure is blind to the risk arising from composition which is not clinically available. To address this need for composition information, the Compositional Analysis System by Machine learning (CASM) algorithm was developed based on spectral analysis of backscattered ultrasound to provide a noninvasive measure of carotid plaque composition. This research effort is designed to determine correlations between output from the CASM algorithm and future outcomes regarding stroke for patients with carotid stenosis. To accomplish this, a prospective longitudinal clinical study is currently being run to enroll 1500 subjects at two sites. Each enrolled patient receives a research ultrasound which provides input data for the CASM algorithm and if they are asymptomatic with no interventions involving their carotid, then they are followed for up to 3 years. The measurements from the CASM algorithm will be obtained and compared between diabetic and non-diabetic, symptomatic and non-symptomatic, and over time for asymptomatic subjects. These findings will be used to find an ultrasound imaging system manufacturer who is interested in hosting the CASM algorithm and is willing to work with us on a 510k application based on the results. In parallel, the results will be used to design a future clinical trial to test the findings from this clinical study.

2. KEYWORDS:

Diabetes, atherosclerosis, stroke, carotid stenosis, cerebrovascular accident, ultrasound, spectral analysis, tissue characterization, machine learning

3. ACCOMPLISHMENTS:

- **What were the major goals of the project?**

  This research project, *Clinical Study of Vascular Plaque Determination for Stroke Risk Assessment*, contains six major tasks in order to achieve the three specific aims.

  The clinical study at the heart of this research effort has two sites: The Cleveland Clinic Foundation (CCF) and the Cleveland VA Medical Research and Education Foundation (VA).

AIM 1: Determine the Correlation between Ultrasonically Obtained Plaque Composition and Future Cerebrovascular Accident (CVA)

Major Task 1: Obtain major equipment and regulatory approval

- Update legal documents with Siemens (***goal: month 1***):
  Completed prior to the start of the contract.

- Obtain two modified research ultrasound machines from Siemens (***goal month 1***)
  - *Milestone*: Ultrasound system hardware delivery:

        - CCF: **Completed** prior to the start of the contract.
        - VA**:** **Completed** Month 8
    - Final Software Installation for Research Ultrasound Exam from Siemens:
        - CCF: **Completed** Month 7
        - VA: **Completed** Month 8

- *Milestone*: Local IRB Approval (***goal month 3*****)**
    - CCF: **Completed** Month 9**:** 11 Jun 2021
    - VA: **Completed** Month 9**:** 10 Jun 2021
- *Milestone*: HRPO Approval (***goal month 6***)
    - CCF: **Completed** Month 12**:** 19 Aug 2021
    - VA: **Completed** Month 11**:** 30 Jul 2021
- *Milestone*: Complete 510k Application (***goal month 48***):
    - TBD: Need to have an agreement in place with a medical device manufacturer to install the CASM algorithm prior to completing a 510k application to the FDA. (See prior Annual Report from 15 Oct 2021 for a discussion on this limitation).

Major Task 2: Fabricate plaque volumetric determination system

- Procurement of hardware and fabrication of device (***goal month 1***)
    - Procurement of hardware: **Completed** Month 1: 2 Oct 2021
    - Fabrication of Device: **Completed** Month 6: 17 Feb 2021
- *Milestone*: Creation of control software and integration with ultrasound system (***goal month 2***)
    - **Completed** Month 10

Major Task 3: Enroll patients for Aim1, 2, and 3

- Start Enrollment **(goal month 6):**
    - CCF: **Completed** 15 Sep 2021
    - VA: **Completed** 10 Sep 2021
- *Milestone*: Reach 1500 Enrollment for Study **(goal month 42)**
    - CCF: 179 of 900 **20% completed** (goal 50%)
    - VA: 125 of 600 **21% completed** (goal 50%)

Table 1: Targeted Enrollment from SOW

	Year 1				**Year 2**				**Year 3**				**Year 4**	
Target Enrollment (per quarter)	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**
Site 1: Cleveland Clinic	0	0	75	75	75	75	75	75	75	75	75	75	75	75
Site 2: LS VAMC	0	0	50	50	50	50	50	50	50	50	50	50	50	50
Target Enrollment (cumulative)	**0**	**0**	**125**	**250**	**375**	**500**	**625**	**750**	**875**	**1000**	**1125**	**1250**	**1375**	**1500**

Table 2: Actual enrollment

	Year 1				Year 2			
Actual Enrollment (per quarter)	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Site 1: Cleveland Clinic	0	0	0	0	65	51	33	30
Site 2: LS VAMC	0	0	0	1	15	44	37	28
Actual Enrollment (cumulative)	**0**	**0**	**0**	**1**	**81**	**176**	**246**	**304**

AIM 2: Does the Presence of Diabetes Affect Estimate of Carotid Plaque Composition?

Major Task 1: Determination of vascular geometry and composition

- Develop and test automated border detection algorithms
  (goal: start month 5 and complete month 30)
  - Subtask - Manual borders production **9 %** completed: started in **Month 12**
  - Subtask - Automated algorithm development: started **Month 18**
    (see below for further details)
- Create plaque compositional images using the CASM software
  (task is repeated in SOW under AIM 2 Task 2 and AIM 3 Task 1)
  (goal: start month 5 and complete month 48)
  0.1% completed started in month 12
  (Code has been tested and batch processing will be performed as needed.)
- *Milestone*: Geometry and composition of all vascular structures quantified
  0 % completed

Major Task 2: Determine the effect of diabetes on carotid plaque composition

- Identify diabetic patients and find matched controls
  (goal: start month 36 and complete month 48)
  - Start in month 36
- Create plaque compositional images using the CASM software
  (see above: same task as AIM 2 Task 1)
- *Milestone*: Statistical correlation between plaque composition and geometry between diabetic and non-diabetic patients produced. Plan to start in month 37 with interim analysis in month 30.

AIM 3: Does Estimate of Plaque Composition Correlate with Future Changes in Plaque Composition

Major Task 1: Determine the rate of plaque progression and correlate with risk factors and medication.

- Identify all patients with baseline and follow-up data
  (goal: start month 36 and complete month 48) plan to start in month 36

- Use segmentation software to identify vascular borders
  (goal: start month 30 and complete month 48) plan to start in month 36 (arising from enrollment delays)
- Create plaque compositional images using the CASM software
  (see above: same task as AIM 2 Task 1)
- *Milestone*: Statistical correlation between plaque composition, geometry, and risk factors.
  Plan to start in month 37

- **What was accomplished under these goals?**

  - **Clinical Study Enrollment**

Clinical study enrollment has reached 20% of the end goal. This is less than the target at the end of Year 2 of 50%. The initial numbers were calculated assuming one carotid artery per enrolled patient however roughly a third of the enrolled subjects had both of their carotid arteries imaged effectively raising the number of carotid arteries included in the study from 304 (20%) to 402 (27%), see Table 3. For most of the endpoint analyses each carotid artery can be treated independently since each has its own measure of stenosis, symptoms from carotid stenosis can generally be ascribed to a specific artery, and surgical intervention is largely determined based on measures for each artery. Correlations do exist between carotid arteries for each patient: i.e. the contralateral carotid artery of a patient with carotid stenosis is more likely to have carotid stenosis than a random carotid artery from the general population. Thus care will need to be taken during the analysis to account for this correlation.

At the current rates we expect to achieve 40% of the enrollment goal by the end of Year 3 or if counting each artery independently that would 54%. We plan to continue enrolling for the first half of Year 4 with an expected total enrollment of 750 subjects and 1000 carotid arteries. See Section 5 for a discussion on the potential impact for this study.

Table 3: Carotid arteries enrolled in the study
Some subjects have both the right and left carotid arteries included in the study.

	Year 1				Year 2			
Arteries Enrolled (per quarter)	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Site 1: Cleveland Clinic	0	0	0	0	85	69	47	40
Site 2: LS VAMC	0	0	0	1	24	57	45	34
Total Arteries Enrolled (cumulative)	**0**	**0**	**0**	**1**	**110**	**236**	**328**	**402**

In addition to the initial research ultrasound exams that are performed at the time of enrollment, follow-up research ultrasound exams are progressing as depicted in Table 4.

Table 4: Follow-up research ultrasound exams (*number of vessels in parentheses*)

	Year 1				Year 2			
Actual Enrollment (per quarter)	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Site 1: Cleveland Clinic	0	0	0	0	0	0	3 (4)	7 (9)
Site 2: LS VAMC	0	0	0	0	0	0	4 (7)	17 (26)
Actual Enrollment (cumulative)	**0**	**0**	**0**	**0**	**0**	**0**	**7 (11)**	**24 (35)**

- **Image Segmentation**

Segmenting the carotid plaque is an immense effort. One that, if not automated well, will greatly impede clinical use of the CASM algorithm. Each carotid artery that is imaged produces at a minimum 120 frames to be segmented. At this point in the study, this segmentation is manual as we create the database for training an automated or semi-automated approach. For the 179 subjects enrolled at CCF, manual segmentation has been performed on 137. One of the 125 enrolled subjects from the VA site has been manually segmented as well. The data from the VA site was only recently made available due to legal negotiations between the VA and CCF that were not resolved until month 24. This did not delay the analysis since we spent the first half of year 2 training the team in how to perform image segmentation. This included multiple sessions of our lead image segmentation specialist (Jaqueline Loftis) working with the new members of the team independently and via group review sessions. The new hires are still improving but now can operate independently and know when to flag difficult frames in order to obtain feedback from the more experienced members of the research team. In addition there is an expert review of all segmented frames (researchers are not permitted to review their own segmentation work).

At this point in the project we have completed 138 of the 304 initial data collections for the 304 subjects (45%). This is 9% of the goal of 1500, however, as mentioned above we anticipate that we will enroll 750 and thus this is closer 18% of the final enrolled subjects. The plan to close this gap is discussed below in this section.

For automated image segmentation, effort has started but has not reached a point of implementation. In addition to our own efforts, we have approached IBM as a potential partner on solving the automation of image segmentation for carotid plaque. An engineering team that has experience working with automating plaque segmentation for intravascular ultrasound images of coronary arteries has expressed interest in working with us. In September we submitted a proposal to obtain IBM engineer time to be devoted to this project as part of the Cleveland Clinic-IBM Discovery Accelerator effort. This proposal is currently being reviewed by Cleveland Clinic and IBM personnel and if approved we anticipate starting to work together in Quarter 2 or Year 3 (Jan 2023).

- **3D Volume Reconstruction**

The 3D volume reconstruction performance of the modified Siemens probe was evaluated by freehand scanning cylindrical inclusions in a CIRS Model 044 tissue mimicking ultrasound phantom (Ultrasound Resolution Phantom model 044, Computerized Imaging Reference Systems, Inc., Norfolk, VA, USA). The CIRS phantom contains 4 groups of anechoic cylindrical inclusions of varying sizes and a series of nylon alignment targets. The inclusion used for this test was 18 mm in length and 12 mm in diameter as illustrated in Figure 1. The cylinder has a total volume of 2,036 mm^3.

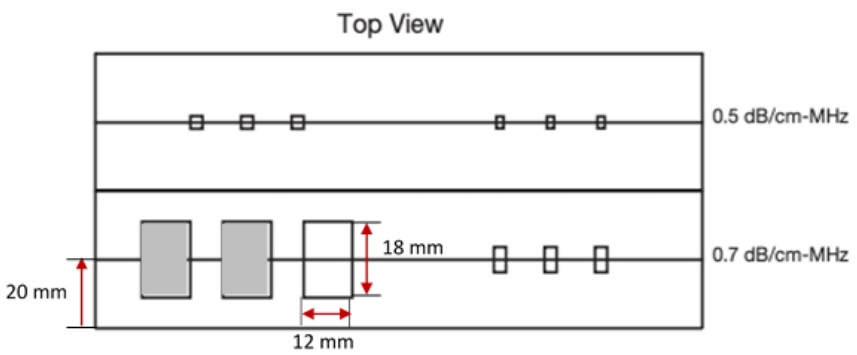

Figure 1: Schematic top-view diagram of the CIRS tissue-mimicking phantom Model 044. The inclusions used for testing are in gray.

Ultrasound data (demodulated IQ-data) were recorded using a Siemens Acuson Sequoia ultrasound system with research interface (Siemens Medical Solutions, Mountain View, CA, USA) equipped with a 10L4 linear array probe and the position/orientation tracking system described in the 2021 Annual Report. Data were collected at 12 frames per second for 10 seconds, for a total of 120 frames. Three scans were performed for volume estimation. For each scan, the probe first held stationary at the beginning of data collection, then moved in a medial direction from the edge of the scanning area while maintaining a uniform rate, thus mimicking the *in vivo* data collection.

The inclusion was manually segmented for each frame using customized Matlab software (The Mathworks Inc., Natick MA, USA) to obtain x and y coordinates. Initially, we had planned to use the inertial measurement unit of the position/orientation tracking system to obtain a measure of velocity for the probe. Unfortunately, this data is extremely noisy and only permits us to note the time when the probe begins motion and stops motion. An average velocity for each scan was calculated based on the start/stop time obtained from the position/orientation tracking system and visually recognizing the start and stop frame positions (a comparable distance can be determined based on the longitudinal measurements obtained from the clinical duplex ultrasound exam). An average velocity is then computed and applied to these frames. Average frame thickness for the three data acquisitions are 0.61 mm/frame, 0.51 mm/frame, and 0.48 mm/frame.

The 3D image was then constructed and the volume of the segmented region was computed using Matlab (function *covhull),* by computing the convex hull and volume of a set of 3D voxels. The volume error was calculated as the absolute value of the

difference between the segmented volume and the true volume, divided by the true volume. These estimates are listed below in Table 5. See section 5 below for further discussion.

Table 5: Estimated Volume and error for cylindrical inclusions from 3 scans.

Scan Number	Estimated Volume (mm^3)	Error Estimate (%)
1	2156	6%
2	2428	19%
3	2349	15%

- **Analysis of the CASM Algorithm**

Efforts continue to better understand and improve the CASM algorithm for providing a measure of plaque composition. These efforts are described in detail the abstract submitted to the 2022 IEEE Ultrasonics Symposium (IUS) and the manuscript accepted for publication in Ultrasonics which both appear in the appendix.

The IUS abstract demonstrates that the addition of spectral information improves on the performance of plaque composition estimation as compared with intensity only approaches. This determination is based on expert review of the plaque composition images compared to histology.

The manuscript accepted for the journal Ultrasonics describes the performance of the algorithm from a statistical nature for homogenous regions in the plaque. These results demonstrate that assuming a constant attenuation for compensation of the backscattered signal is acceptable and that using the nonlinearly generated second harmonic portion of the backscattered signal improves on the tissue composition estimation. See the appendix for more details.

- **What opportunities for training and professional development has the project provided?**

Nothing to Report

- **How were the results disseminated to communities of interest?**

Nothing to Report

- **What do you plan to do during the next reporting period to accomplish the goals?**

Enrollment: The difficulties in enrolling patients arises primarily from the fact that even though roughly 6000 and 5000 duplex ultrasound exams are performed each year at the CCF and VA sites respectively, the majority of these folks do not have sufficient plaque to qualify for this study (inclusion requires 40% or greater carotid stenosis). To counter this issue both sites are attempting to maximize recruitment of qualified candidates. Currently, the CCF site is negotiating with the CCF IRB to permit same day consent for patients who are scheduled at the last minute as well as for patients who have been contacted but did not receive the consent form prior to coming into CCF. The VA site is

permitted to perform same day consent and is expanding the patients that are contacted each day in order to improve enrollment, even though most of these patients are expected to have insufficient plaque. Unfortunately, the lack of patients with sufficient stenosis is beyond our control. At this point, even with our best efforts we anticipate under-enrollment to persist. See section 5 for further discussion.

Image Segmentation: The focus of Year 2 was in training the research team to perform the manual segmentation. Two new hires were added with little experience during the first half of Year 2. Neither of these individuals had any experience interpreting grayscale images. Thus there was a significant learning curve. With both of these team members trained, we anticipate a significant increase in manual segmentation and expect to be closing the gap between segmented and collected data sets.

For automated segmentation, we predicted that we would need 300 cases to seriously begin creating algorithms. We have just reached that mark in month 24 (instead of month 18) and anticipate having manual borders completed for these 300 cases in quarter 2 of Year 3. This also coincides with the expected beginning of the collaboration with IBM experts in ultrasound image segmentation and with our interim review (see below). If the collaboration with IBM does not happen, we will continue building up our skillset for creating convolution neural network AI segmentation tools within our group and expect to implement them at that time.

3D Volume Estimation: As described in section 5, the noise of the inertial measurement unit (IMU) has been larger than expected and thus the position/orientation system is under-performing. We are still able to obtain the time difference from the IMU and these times are associated with the collected frames. The distance traveled is obtained by locating the start and stop frames in the longitudinal images from the clinical duplex ultrasound. Extracting this data and improving the 3D volume estimation will be part of the coming year's efforts.

Month 30 Interim Review: In order to better understand the quality of the data that has been collected and the impact of the under-enrollment, an interim review will be performed at the end of Quarter 2 in Year 3. This review will test our established image segmentation, CASM algorithm, and 3D volume estimation. In addition, we will be able to review the computation of the number needed to enroll. The target of 1500 was derived from a sampling of 30 subjects from a prior study. This prior study did not measure the entire plaque and only a single frame from the region of greatest stenosis was used to represent the plaque for each of these 30 subjects. The size of the hemorrhagic and/or necrotic core (H/NC) region was then used to determine the power for this clinical study. However, we are not only collecting data at the point of greatest stenosis but also for the entire plaque. The use of AI to combine various measures from the overall plaque burden, to a direct measure of the degree stenosis, along with output from the CASM algorithm will be investigated to determine the association of our measurements with clinically relevant status of the subjects. Thus we have a much richer data set which may indicate a smaller size needed for this study. Only data collected at the initial visit will be included in this interim analysis.

4. IMPACT

- **What was the impact on the development of the principal discipline(s) of the project?**

  Nothing to Report

- **What was the impact on other disciplines?**

  Nothing to Report

- **What was the impact on technology transfer?**

  Nothing to Report

- **What was the impact on society beyond science and technology?**

  Nothing to Report

5. CHANGES/PROBLEMS:

- **Changes in approach and reasons for change**

  Nothing to Report

- **Actual or anticipated problems or delays and actions or plans to resolve them**

  - **Enrollment**

  We are under enrolling for the clinical study at both sites at roughly half of the target enrollment. Despite sufficient number of duplex ultrasound exams of the carotid artery, the vast majority of patients do not have clinically significant carotid stenosis. In addition to this primary issue, both sites have had issues with staffing amongst the ultrasound technicians which has also negatively impacted enrollment. (Data collection requires an ultrasound technician working with a member of the research team.) A mitigating factor is that roughly a third of the enrolled subjects qualify for both carotid arteries thus resulting in an expected effective under enrollment of one third when treating each qualifying carotid towards the enrollment goal.

  The plan moving forward is to perform an interim analysis in Quarter 2 of Year 3 to re-evaluate the enrollment needed to power this study while continuing to look for approaches that will increase enrollment.

- **Automating Image Segmentation**

In Year 2, we were presented with the opportunity to work with IBM experts on image segmentation of ultrasound images involving atherosclerotic plaque. This opportunity is based on an agreement between IBM and CCF to accelerate medical research using the skills and capabilities of IBM with the clinical expertise and data available at CCF. Specifically, the two companies created the Cleveland Clinic-IBM Discovery Accelerator program which accepts non-monetary research proposals. Effectively, we have submitted a joint proposal to request effort from IBM engineers to assist with artificial intelligence aspects of this research. Primarily, this is the automation effort for image segmentation. An initial informal review at IBM was supportive. For now, we are awaiting the official response which is expected in Month 26 (late October 2022). If this proposal is not supported, we will revert to the current plan in progress: a convolution neural network approach for automating based on data from the first 300 cases.

- **3D Volume Estimation**

The inertial measurement unit (vn-100 rugged from VectorNav, Dallas, TX) at the heart of the position/orientation tracking system has been found to be too noisy. We are able to obtain the start and stop times from it but do not have confidence in the velocity estimations. Thus, we are implementing a backup plan that relies on the following:

- Ultrasound images collected: determine the start and stop frames when imaging the entire plaque.
- IMU data: determine the time the probe was moving.
- Clinical Duplex Ultrasound: determine the locations of the start and stop frames and thus deduce the distance traveled.
- Assume an average velocity over this distance.
- Apply this average velocity to produce a 3D volume estimation.
- Synchronization of the IMU data and the RF ultrasound data collection can be used as a check that the start movement and stop movement frames are aligned with the start and stop times from the IMU.

The initial implementation of this approach is described above in Section 3. We will continue to refine this approach in order to minimize the volume estimation. Errors in the volume estimation will directly impact all volume measurements, but will not affect measurements that rely on a 2D view of the stenosis. For example, direct measure of the maximum stenosis will be unaffected as will the CASM results at the point of maximum stenosis.

- **Changes that had a significant impact on expenditures**

  The limited enrollment is directly associated with the cost for ultrasound technician time (i.e. sonographer).

  Travel limitations and disruptions due to COVID has shifter attendance to professional conferences into Year 3.

  The hiring of the research coordinator was delayed into quarter 2 (rather than the expected hiring in quarter 1) as a result of a shortage of trained clinical research coordinators in the region.

  The expected replacement for Tanujit Dey (Key personnel, statistician) has been further delayed and is expected in quarter 2 of year 3, which is also when we anticipate performing an interim analysis of the clinical data.

- **Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents**

  - **Significant changes in use or care of human subjects**

    Nothing to Report

  - **Significant changes in use or care of vertebrate animals.**

    Nothing to Report

  - **Significant changes in use of biohazards and/or select agents**

    Nothing to Report

6. PRODUCTS

- **Publications, conference papers, and presentations**

  - **Journal publications**

    James S, Fedewa RJ, Lyden S, Vince DG, *Spectral Analysis of Ultrasound Backscatter for Non-invasive Measurement of Plaque Composition*, Ultrasonics, accepted for publication Sept 2022.
    Federal support from this grant is acknowledged.
    Accepted manuscript is included in the Appendix.

  - **Books or other non-periodical, one-time publications**

    Nothing to Report

- **Other publications, conference papers, and presentations**

  James S, Fedewa R, Lyden S, Vince DG, "Integrated Backscatter Versus Spectral Parameters for in Vivo Estimation of Human Carotid Plaque Composition", accepted for Poster presentation at the 2022 IEEE International Ultrasonics Symposium, Oct 10-13, 2022.
  Submitted abstract is included in the Appendix.

- **Website(s) or other Internet site(s)**

  Nothing to Report

- **Technologies or techniques**

  Nothing to Report

- **Inventions, patent applications, and/or licenses**

  Nothing to Report

- **Other Products**

  Nothing to Report

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

- **What individuals have worked on the project?**

D. Geoffrey Vince – no change
Michael A. Rosenbaum – no change
Sheronica L. James – no change
Jaqueline Loftis – no change

Name:	Russell J. Fedewa
Project Role:	Co- Investigator at CCF site
Researcher Identifier (e.g. ORCID ID):	0000-0002-0690-9472
Nearest person month worked:	11
Contribution to Project:	Managing regulatory submissions with key sites and collaborators. Development of clinical study protocol and support documents with assistance from co-investigators. Oversight and review of lab personnel for image segmentation. Support for signal and image processing and programming with support for image segmentation and lead for automatic segmentation effort.
Funding Support:	NA

Name:	Maya Mays
Project Role:	Research Technician at CCF site
Researcher Identifier (e.g. ORCID ID):	NA
Nearest person month worked:	12
Contribution to Project:	Screening, enrolling, and data acquisition are the primary tasks with secondary task of image segmentation.
Funding Support:	NA

Name:	Allison Smollen
Project Role:	Clinical Research Coordinator at CCF site
Researcher Identifier (e.g. ORCID ID):	NA
Nearest person month worked:	8
Contribution to Project:	Initial screening of patients, initial contact and follow-up, capturing data from medical records, handling communications with IRB, record keeping for the research study operation.
Funding Support:	NA

Name:	Jerad Williams
Project Role:	Research Nurse / Study Coordinator at VA site
Researcher Identifier (e.g. ORCID ID):	NA
Nearest person month worked:	12
Contribution to Project:	Mr. Williams has been responsible for the screening, recruitment, and consenting of subjects. He has also assisted with study data collection.
Funding Support:	NA

Name:	Manda Double
Project Role:	Research Regulatory and Compliance specialist at VA site
Researcher Identifier (e.g. ORCID ID):	NA
Nearest person month worked:	3
Contribution to Project:	Ms. Double has managed the IRB submission and documentation for the Northeast Ohio VA Health System IRB, tracking of human subjects' certification and Conflicts of Interest documentation, and assisted with regulatory audits.
Funding Support:	NA

- **Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**

  Name/Role: Russell J. Fedewa / Co-Investigator
  Description of Change: Percent effort has been reduced from 100% to 80% with 20% effort applied to the new research effort listed below. A new technician has been hired to work on image segmentation in response.
  Title: **Broad Bandwidth Transducers for High Resolution Information Rich IVUS**
  Funding Agency: Department of Health and Human Services, National Institutes of Health, National Heart, Lung, and Blood Institute: *1R61HL156154-01A1*
  Start and End Date: 06/10/2022 – 04/30/2024
  Level (%) of Effort in the Project: 20
  Potential Overlap: None
  Role: Co-investigator

Name/Role: Michael A. Rosenbaum / Co-Investigator
Description of Change: Participation in the following research effort has **ended**
Title: **Endothelial Healing is Inhibited by PI3 Kinase-induced activation of TRPC6**
Funding Agency: Department of Veterans Affairs Biomedical Research and Development Service, IK2BX003628
Start and End Date: 01/10/2017 – 12/31/2021
Level (%) of Effort in the Project: 75
Potential Overlap: None
Role: PI

Name/Role: Michael A. Rosenbaum / Co-Investigator
Description of Change: Participation in the following research effort has **begun**
Title: **Endothelial Healing is Inhibited by Activation of TRPC6 Channels**
Funding Agency: Department of Veterans Affairs Biomedical Research and Development Service, I01BX005823
Start and End Date: 01/01/2022 – 12/31/2025
Level (%) of Effort in the Project: 37.5
Potential Overlap: None
Role: PI

Name/Role: Michael A. Rosenbaum / Co-Investigator
Description of Change: Participation in the following research effort has **begun**
Title: **Targeted Abrogation of the FXII-uPAR-pAkt2 Axis in Neutrophils for Treatment of Chronic Wounds**
Funding Agency: National Institutes of Health/National Heart, Lung, and Blood Institute, 1R01 HL137695
Start and End Date: 01/25/2019 – 12/31/2023
Level (%) of Effort in the Project: 10
Potential Overlap: None
Role: Co-Investigator

Name/Role: Michael A. Rosenbaum / Co-Investigator
Description of Change: Participation in the following research effort has **begun**
Title: **Effects of Lipids on Vascular Graft Healing**
Funding Agency: National Institutes of Health/National Heart, Lung, and Blood Institute, 2R01 HL064357
Start and End Date: 12/01/2017 – 11/30/2023
Level (%) of Effort in the Project: 10
Potential Overlap: None
Role: Co-Investigator

- **What other organizations were involved as partners?**

Organization Name: Siemens Medical Solutions USA, Inc.
Location of Organization: 51 Valley Stream Parkway, Malvern PA 19355, USA

Partner's contribution to the project: In-kind support
Equipment loan: 2 Siemens Sequoia Ultrasound Systems with associated probes, software, service, and engineering support.

Organization Name: Cleveland VA Medical Research and Education Foundation
Location of Organization: 10701 East Blvd, Cleveland, OH 44106-1702, USA
Partner's contribution to the project: Collaboration and Facilities
Second site for clinical study and scientific collaboration with site personnel.

8. SPECIAL REPORTING REQUIREMENTS

Nothing to Report

9. APPENDICES

- Abstract: James S, Fedewa R, Lyden S, Vince DG, *Integrated Backscatter Versus Spectral Parameters for in Vivo Estimation of Human Carotid Plaque Composition*, accepted for Poster presentation at the 2022 IEEE International Ultrasonics Symposium, Oct 10-13, 2022.
- Manuscript accepted for publication:
James S, Fedewa RJ, Lyden S, Vince DG, *Spectral Analysis of Ultrasound Backscatter for Non-invasive Measurement of Plaque Composition*, Ultrasonics, accepted for publication Sept 2022.

Integrated Backscatter versus Spectral Parameters for In Vivo Estimation of Human Carotid Plaque Composition

Sheronica James[1], Russell Fedewa[1], Sean Lyden[2], D. Geoffrey Vince[1], [1]Cleveland Clinic Department of Biomedical Engineering, Cleveland, United States, [2]Cleveland Clinic Department of Vascular Surgery, Cleveland, United States

Background, Motivation and Objective

Carotid plaque composition is a missing piece of information in the treatment of carotid stenosis. Non-invasive approaches for estimating composition include ultrasound grayscale based analysis, ARFI, MRA, and spectral analysis. This study evaluates a spectral analysis based approach versus an intensity only approach (comparable to an idealized grayscale) using both fundamental and harmonic bandwidths. The intensity approach utilizes the integrated backscatter (IB), while the spectral analysis approach uses the slope, intercept and mid-band fit from an estimate of the backscatter transfer function. Plaque is classified into three types: hemorrhagic and/or necrotic core (HNC), fibrous (F), and calcium (Ca).

Statement of Contribution/Methods

Backscattered ultrasound RF data from carotid plaque were acquired in vivo in transverse orientation with respect to the carotid artery prior to carotid endarterectomy for 134 subjects. Serial histology of the excised plaque were prepared following surgery and spatially independent ultrasound frames were matched to the histology. Regions of interest (ROI) were selected in the RF data corresponding to 1.2 mm by 1.2 mm homogenous regions within the histology. A data set of 120 F, 125 HNC, and 130 Ca ROI's were randomly selected from the total ROI's to train and test two random forest classification models relying on either 1) IB or 2) linear fit parameters as inputs.

Color maps for 30 randomly selected matched frames were produced from the two random forest classifiers and compared to the matched histology based on a blinded expert review.

Results/Discussion

For classifying HNC, accuracy and specificity were slightly better for the linear fit based approach than IB (accuracy, 0.63 ± 0.05 vs 0.57 ± 0.06 and specificity, 0.77 ± 0.1 vs 67 ± 0.1). Sensitivity was the same for both (0.36 ± 0.07). Figure 1 depicts one data set comparing both the grayscale, histology, and color maps. The blinded review found the linear fit parameter based model provided a better representation of the plaque than the IB based model for 26 of the 30 frames. These results support the understanding that spectral information can improve on the performance of intensity only based approaches for ultrasound based tissue characterization of carotid plaque.

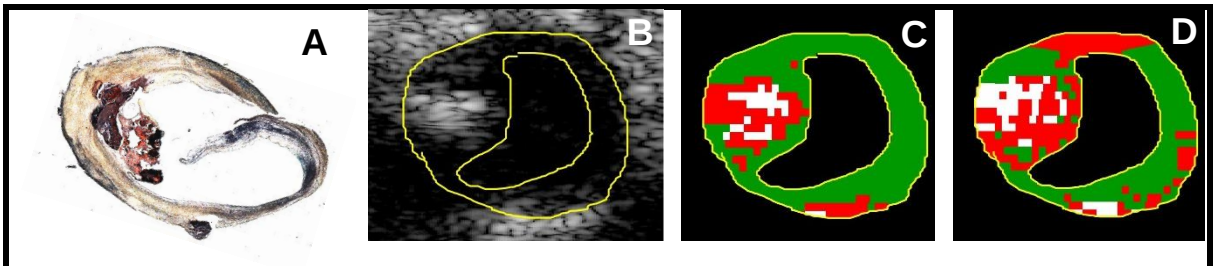

Fig.1. Comparison of (A) matched histology slide with (B) grayscale ultrasound image of *in vivo* carotid plaque and classification output images relying on different input parameter sets: (C) Linear fit parameters and (D) integrated backscatter only. The histology slide of the carotid plaque depicts a calcified (white) necrotic core (red) region on the left with a calcified nodule on the bottom and the remainder being fibrous (green).

Spectral Analysis of Ultrasound Backscatter for Non-invasive Measurement of Plaque Composition

Sheronica L. James[a*] (james4@ccf.org); Russell J. Fedewa[a] (fedewar@ccf.org); Sean Lyden[b] (lydenss@ccf.org); D. Geoffrey Vince[a] (vinceg@ccf.org)

[a] Department of Biomedical Engineering, Cleveland Clinic, 9500 Euclid Ave, Cleveland, OH 44195, USA
[b] Department of Vascular Surgery, Cleveland Clinic, 9500 Euclid Ave, Cleveland, OH 44195, USA
*Corresponding author

This work was supported by the Office of the Assistant Secretary of Defense and Health Affairs, through the Peer Reviewed Medical Research Program under Award No. W81XWH-16-1-0608. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense or U.S. Army.
Ultrasound imaging system and probe loaned by Siemens Medical Solutions USA, Inc.
Research supported by AHA SDG Award 15SDG25700258

ABSTRACT

Carotid atherosclerotic plaque composition may be an important indication of patient risk for future cerebrovascular events. Ultrasound spectral analysis has the potential to provide a robust measure of plaque composition *in vivo* if the backscatter transfer function can be sufficiently isolated from the effects of attenuation from overlying tissue, receive and transmit transfer functions from the ultrasound system and transducer, and diffraction. This study examines the usefulness of the nonlinearly generated second harmonic portion of the backscatter signal and the effects of a variety of attenuation compensation techniques for noninvasively characterizing human carotid plaque using spectral analysis and machine learning. Post-beamformed ultrasound backscatter radiofrequency (RF) data were acquired from 6 normal subjects and 119 carotid endarterectomy patients prior to surgery. Plaque obtained following surgery was histologically processed, and regions of interest (ROI) corresponding to homogenous tissue types (fibrous, hemorrhagic and/or necrotic core and calcium) were selected from RF data. Both the harmonic and fundamental power spectra for each ROI was obtained and normalized by data from a uniform phantom (0.5 dB/cm-MHz slope of attenuation). Additional attenuation compensation approaches were compared to simply using the reference phantom: 1) optimum power spectral shift estimation, (2) one-step adventitial, or (3) two-step adventitial. Spectral parameters extracted from both the fundamental and harmonic estimates of the backscatter transfer function of 363 ROI's from 152 plaque specimens were used to train and test random forest and support vector machine classification models. The best results came from using spectral parameters derived from both the fundamental and second harmonic bands with a predictive accuracy of 65-68%, kappa statistic of 0.49-0.54, and accuracies of 84% for fibrous, 68-74% for hemorrhagic and/or necrotic core, and 78-81% for calcium ROI's. The result indicated that the nonlinearly generated second harmonic portion of backscatter is useful for carotid plaque tissue characterization and that a reference phantom approach with a 0.5 dB/cm-MHz slope of attenuation works as well as more complicated approaches.

Key Words: Atherosclerosis, plaque characterization, spectral analysis, attenuation, harmonics, stroke

1. INTRODUCTION

Stroke is the second leading cause of death and the third leading cause of disability, globally [1]. Ischemic stroke, which accounts for roughly 70% – 85% of strokes worldwide [1], [2], is typically by caused by rupture of carotid atherosclerotic (stenotic) plaque. While degree of stenosis based on a Doppler measure of blood velocity has been the gold standard in guiding clinical decision-making [3], research shows that plaque composition is likely to be a more powerful indication of patient risk for future cerebrovascular events (CVE) than the current standard measure of carotid stenosis [4], [5].

Grayscale (B-mode) analysis, which relies on analyzing the envelope-detected, time-domain signals scattered from different tissue structures, has been used to provide a measure of plaque composition that distinguishes symptomatic from asymptomatic plaque [6], [7] and to predict the degree of instability of asymptomatic plaque [8]. However, grayscale analysis struggles at distinguishing vulnerable plaque components, such as hemorrhage and necrotic core [9] from other tissue types [10], [11] , thus limiting its usefulness for clinical care.

Ultrasound spectral analysis of backscattered radiofrequency (RF) data can be used to extract information about tissue microstructure that is lost during the ultrasound image formation process and can complement grayscale-based characterization approaches to provide a more robust measure of carotid plaque composition *in vivo*. Many have leveraged spectral analysis of ultrasound backscatter for tissue characterization [12]–[22], but few have utilized spectral analysis of the backscattered RF signal for estimating carotid plaque composition *in vivo*, and to our knowledge, none have utilized the nonlinearly generated second harmonic bandwidth. Published efforts for *in vivo* plaque characterization have demonstrated that integrated backscatter (IB) of the normalized power spectrum can differentiate fibrous, fatty and calcified lesions of the [12]–[14], [23] arterial wall, but rarely considered hemorrhage or necrosis. However, since IB is a summation across all frequencies, it does not provide information on the variation of scattering with respect to frequency or nonlinearly generated harmonics. Center frequency shift [24] and attenuation coefficient [25] have been used to identify plaques with histological vulnerable phenotypes, but have not focused on differentiating plaque components.

Vulnerable carotid plaques share some histological features of coronary plaques, such as a lipid-rich necrotic core, fibrous tissue, and calcified tissue [26], [27] that have been identified by VH-IVUS with integrated backscatter (IB) and multiple other spectral parameters including: maximum and minimum power with corresponding frequency, slope, y-intercept, and mid-band fit [21], [27]. The invasive nature of IVUS makes it impractical for routine use in carotid arteries, but the relatively shallow depth of the carotid artery facilitates the use of clinical linear-array transducers to collect RF data non-invasively. Use of a clinical linear array enables the application of the nonlinearly generated second harmonic for spectral analysis which is not possible for current IVUS catheters due to limited bandwidth.

Transcutaneous insonification of the carotid presents several difficulties for tissue characterization based on spectral analysis. Attenuation and distortion arising from ultrasonic propagation through the different tissue layers between the transducer and the target plaque tissue can complicate spectral analysis. System-dependent effects, such as focusing and time-gain compensation, as well as patient motion and respiratory artifacts are additional sources of distortion. These factors impact the measures of spectral parameters. Normalizing to a well-characterized reference phantom has been demonstrated to compensate for some of these effects (i.e. transducer and system transfer functions, diffraction, and attenuation) [28], but this compensation degrades as the reference phantom acoustic properties (e.g. attenuation and velocity) differ from the *in vivo* tissue. The heterogeneous nature of the path to the carotid artery means that the acoustic properties of a uniform

ultrasound phantom will always have some differences with the tissue between the carotid and ultrasound transducer. Thus, additional compensation approaches may be needed beyond the reference phantom.

This study examines the effects of a variety of attenuation compensation techniques and the use of the nonlinearly generated second harmonic bandwidth for noninvasively characterizing human carotid plaque using spectral analysis and machine learning. Specific focus is placed on identifying hemorrhagic and/or necrotic core (H/NC) tissue, since this type of plaque tissue is most often associated with vulnerable or rupture-prone atherosclerotic lesions. This effort is a part of the longer-term effort to provide a clinically useful tool to improve stroke risk stratification for patients with carotid stenosis.

2. DATA ACQUISITION

2.1. Clinical Study

As part of a pilot clinical study, ultrasound backscatter RF data was collected prior to subjects receiving a carotid endarterectomy (CEA). Following the CEA, the excised plaque was collected to use as a 'gold' standard for the tissue classification algorithm. Inclusion was limited to patients with atherosclerotic plaque within the distal common carotid artery, the carotid bulb, or the proximal internal carotid artery. Patients with a prior intervention with the ipsilateral carotid artery were excluded as were patients who could not provide consent in English.

The *in vivo* ultrasound employed a Siemens S3000 HELX ultrasound system with Axius Direct Ultrasound Research Interface (URI) software enabling the acquisition of pulse inversion beamformed RF data at 40 MHz sampling. Using a 9L4 linear array transducer, two sets of data were acquired while imaging the carotid artery in a transverse orientation. The first set of data was acquired while holding the transducer still (i.e. static sites) at sites separated by 1 cm from the proximal to distal ends of the carotid plaque region. Next, RF data loops of 54 frames were collected while moving the transducer from the proximal end of the plaque to the distal end. This loop data served as a backup to the static sites and provided context for matching *in vivo* frames to histology. Reference RF data were acquired from a 0.5 dB/cm-MHz tissue-mimicking phantom (Ultrasound Resolution Phantom model 044, Computerized Imaging Reference Systems, Inc., Norfolk, VA, USA). Identical system settings were used for both the *in vivo* data collection and the phantom data collection. The depth was set to 4 cm with a transmit focus of 2cm and 456 lines per frame. The analysis in this paper is obtained from the first 119 subjects

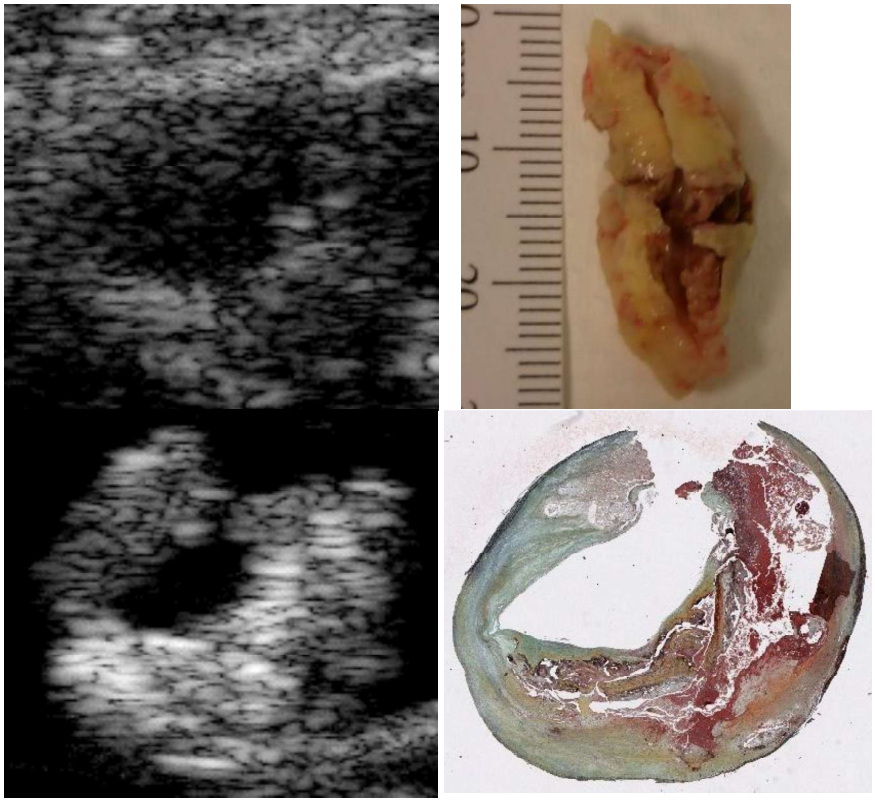

Figure 1: Depiction of the matching process starting in the upper left *in vivo* image of the internal carotid artery for UMP0050 Site 01. The upper right image depicts the plaque shortly after being removed (mm scale shown). The scale was photographed with the excised plaque. The lower left *ex vivo* image corresponds to the dashed line in the upper right image. Bottom right is the histology match with a Movat pentachrome stain. Histology shows hemorrhagic necrotic core from roughly 12-8 o'clock.

enrolled following this arrangement (the first 6 subjects enrolled in the study are not included since the acquisition for these subjects did not employ pulse inversion). In addition to the CEA subjects, six normal subjects had ultrasound backscatter from both of their carotid arteries collected. This normal data set was used for attenuation compensation.

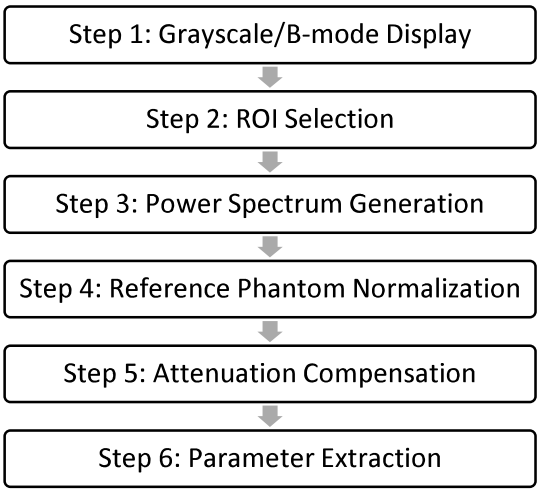

Figure 2: Diagram showing spectral analysis steps.

2.2 Histology Matching to Ultrasound Data

On the same day as the CEA, the excised plaque tissue is placed in saline and ultrasound images are collected in order to locate regions of the plaque corresponding to the *in vivo* static sites. The tissue was then placed in 10% buffered formalin for at least 24 hours followed by two rounds of de-calcification (Cal-Rite, Thermo Fisher Scientific, Runcorn, UK) for at least 24 hours for each round. Prior to paraffin embedding the target sites for matching were marked on the tissue with dye. Serial histology was performed with slide pairs taken at 1 mm steps before and after the marked region and 0.5 mm steps through the target region. One slide of each pair received hematoxylin and eosin stain while the other slide received Movat pentachrome stain. The best matching slide to each *in vivo* grayscale site was found by one member of the research team and reviewed by a different member, and discrepancies were discussed and resolved. An example set of *in vivo* and *ex vivo* images for a matched site is shown in Figure 1. A pathology expert than read the matched slide and the nearest neighbor slides defining regions of homogenous tissue across all three slides. Three tissue types were defined: calcified regions (Ca), fibrous and/or fibro-fatty regions (F), and hemorrhagic and/or necrotic core regions (H/NC). Next, regions of interest (ROI's) were drawn within the grayscale *in vivo* images corresponding to these homogenous regions in histology (see Figure 4). Customized Matlab software (The Mathworks Inc., Natick MA, USA) was used to extract the RF data corresponding to these ROI's. Each ROI was 64 points by 15 lines which corresponds to 1.2 mm by 1.2 mm.

3. SIGNAL PROCESSING

3.1. Spectral Analysis Procedure

Determining tissue composition based on spectral analysis of backscattered ultrasound relies on isolating the scattering properties of the target tissue. This is performed via a sequence of operations (summarized in Fig.2) applied to the acquired *in vivo* RF data, shown in Fig. 3, and described as the follows:

Step 1: Grayscale B-mode images were formed from ultrasound backscatter RF data acquired *in vivo* prior to subjects receiving a carotid endarterectomy (CEA). Each frame is composed of 456 scanlines by 2076 points digitized at 40 MHz. Each scanline corresponds to 2 RF data lines arising from the use of pulse inversion for isolating the nonlinearly generated second harmonic portion of the backscatter. Thus, for each frame, two sets of RF data and corresponding grayscale images are formed: fundamental portion of the backscatter formed by computing the pulse inversion difference and nonlinearly generated second harmonic portion of the backscatter form by computing the pulse inversion summation followed by a high pass filter.

Step 2: Regions of interest (ROI's) were drawn within the grayscale *in vivo* images corresponding to homogenous regions in histology (see Figure 2). Customized Matlab software (The Mathworks Inc., Natick MA, USA) was used to extract the RF data corresponding to these ROI's. Each ROI was 64 RF data points

by 15 scanlines which corresponds to approximately a 1.2mm by 1.2 mm area. Thus for each ROI, two sets of RF data are extracted: fundamental and harmonic.

Step 3: An average power spectrum for each ROI was obtained by computing the power spectrum of each scanline line within the ROI using a Yule-Walker Autoregressive (AR) Method of order 24, converting to dB scale, and then averaging over all 15 lines. Two average power spectra were computed for each ROI: one for the fundamental and one for the harmonic portion of the backscatter.

Step 4: For *in vivo* applications in soft tissue, Yao *et. al.* have demonstrated that normalizing to a well-characterized reference phantom effectively removes the effects of the transducer, system, diffraction, and some attenuation [28]. Assuming that scattering in soft tissue is weak enough to ignore multiple scattering (Born approximation) and that the tissue within the ROI is homogeneous, the value of the power spectrum of the backscattered RF signal at frequency f and depth z within a target ROI, $S(f,z)$ can be expressed as

$$S(f,z) = G(f)\,D(f,z)\,BSC(f,z)e^{-4\,\alpha(f)\,z}, \tag{1}$$

where $G(f)$ represents the combined transducer and ultrasound system effects from transmitting and receiving an RF signal, $D(f,z)$ accounts for diffraction, and $BSC(f,z)$ is the backscatter coefficient within the ROI, $e^{-4\,\beta\,fz}$ is the cumulative attenuation along the propagation path between the transducer surface and the ROI, and β is the slope of attenuation of the medium expressed in dB/cm-MHz. Then, the ratio of the backscattered signal power spectrum from a plaque tissue ROI to that from a reference phantom ROI at the same depth can be described as:

$$\frac{S_{TIS}(f,z) = G(f)_{TIS}\,D(f,z)_{TIS}\,BSC_{TIS}(f,z)e^{-4\beta_{TIS}fz}}{S_{REF}(f,z) = G(f)_{REF}D(f,z)_{REF}\,BSC_{REF}(f,z)e^{-4\beta_{REF}fz}} \tag{2}$$

where the subscripts *TIS* and *REF* denote the plaque tissue and reference phantom media, respectively. Assuming that the speed of sound within the tissue sample, reference phantom and system beamformer is approximately the same, simplifying equation Eq. (2) to

$$\frac{S_{TIS}(f,z)}{S_{REF}(f,z)} = \frac{B_{TIS}(f,z)}{B_{REF}(f,z)}\; e^{-4(\beta_{TIS}-\beta_{REF})\,fz} \tag{3}$$

The backscatter coefficient of the reference phantom need not be similar to the target tissue, but must be uniform within the phantom to obtain an estimate of the backscatter transfer function (*eBTF*) relative to the reference phantom for a given plaque tissue ROI. This estimate is then expressed as

$$eBTF(f,z) = \frac{S_{TIS}(f,z)}{S_{REF}(f,z)}\; e^{4(\beta_{TIS}-\beta_{REF})\,fz}\quad . \tag{4}$$

The reference power spectrum was determined by forming the fundamental and harmonic RF data sets as described for the in vivo data. Then extracting corresponding RF signals from the same depth as the *in vivo* ROI across all 456 scanlines for 10 frames. The same AR approach described above was used to estimate the power spectrum of each line and converted to dB. The average is then subtracted from the corresponding ROI power spectrum found in the prior step for both the fundamental and harmonic bands. This provides an estimate of the backscatter transfer function, $eBTF(f,z)$ relative to the phantom.

Step 5: As shown in equation 4, the uniform reference phantom may be insufficient to account for the attenuation of the propagation path between the transducer and the ROI with this difference represented by $\beta_{TIS} - \beta_{REF}$. Thus, three attenuation compensation methods in addition to simply relying on the reference phantom normalization are investigated:: (1) optimum power spectral shift estimator (OPSSE); (2) 1-step normalization using adventitia and (3) 2-step normalization using adventitia. See Section 4, for more details.

Step 6: Spectral parameters were extracted from both the fundamental and harmonic *eBTF* of each ROI. These parameters, along with the histology determined tissue type, form the basis for training and testing the machine learning based classifiers. The operational bandwidths for the parameters were defined by the 20dB down locations obtained from the phantom data: 2.5 MHz to 6.9 MHz for the fundamental and 4.9 MHz to 10.1 MHz for the harmonic. Three parameters are obtained from a linear fit to the *eBTF* for each band: slope, intercept, and mid-band fit. Integrated backscatter (IB) was computed as the area under the *eBTF*. The final four parameters are obtained by locating the maximum (P_{max}) and minimum (P_{min}) values within each band and the corresponding frequencies (F_{max} and F_{min}). The computation of these parameters for the fundamental spectrum are illustrated in Figure 3. Each set of spectral features were then chosen alone or in combination as input for classification.

3.2 Non-linearly generated second harmonic

The non-linearly generated second harmonic signal produced by a clinical imaging system can be separated into two key portions: the nonlinear propagation of the signal from the transducer to the target tissue and the backscattered signal which has a greatly reduced amplitude and behaves linearly. When producing ultrasound images, the harmonic portion of the backscattered signal is treated comparable to a linearly propagating signal after it is separated from the fundamental portion of the backscatter. To determine if we can do the same for spectral analysis we need to examine the production of second harmonic during the propagation out to the target tissue. Is it safe to assume that the finite amplitude distortion will be reasonably robust across subjects?

Assuming that the attenuation is reasonably approximated as linear with frequency; the attenuation for the second harmonic α_2 is roughly twice the attenuation of the fundamental α_1. Then when transmitting a finite amplitude plane wave at the fundamental frequency, f, with starting position $z = 0$ and peak sound pressure p_0, the second harmonic portion of the pressure amplitude, p_2, at a distance z can be written as [29],

$$p_2 = \left(1 + \frac{B}{2A}\right) \frac{\pi p_0^2 f z}{\rho_0 c_0^3} e^{-\left(\alpha_1 + \frac{\alpha_2}{2}\right) z} \tag{5}$$

where $(1+B/2A)$ is the coefficient of nonlinearity, ρ_0 is the density, and c_0 is the velocity. In treating the second harmonic in the same manner as the fundamental potion of backscatter we are assuming that differences in the coefficient of nonlinearity, attenuation, density, and velocity for the tissues located between the skin and target plaque can be ignored for a given depth and transmit frequency.

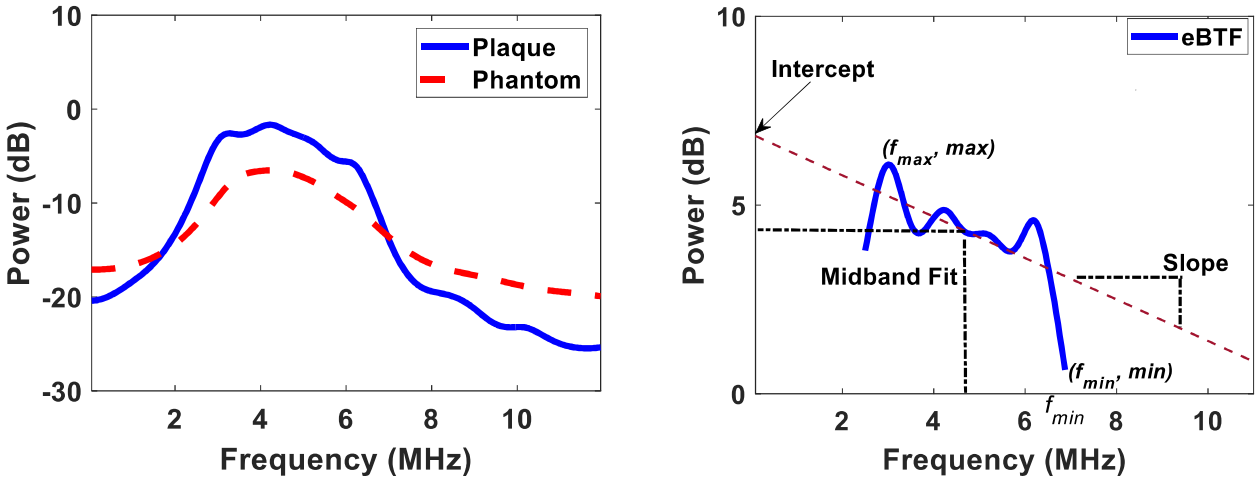

Figure 3: Diagram illustrating computation of spectral parameters. Left - Sample average region-of-interest (ROI) power spectrum from plaque tissue and average reference phantom power spectrum. Right - Estimated backscatter transform function (*eBTF*) within the usable 20dB down bandwidth of 2.5 -6.9 MHz. Each short-time segment of radiofrequency data (RF) that passes through the ROI is extracted from the raw data. An autoregressive model is used to compute the power spectrum of each of the A-lines and then the resulting power spectra are averaged across the scanlines in the ROI. A power spectrum from a tissue phantom, from the same depth as the ROI, is used to normalize the average ROI power spectrum to get the *eBTF*. Spectral parameters are computed from the *eBTF* over a limited bandwidth of 2.5 -6.9 MHz as determined from the reference phantom. A regression line is fit to the *eBTF*, and the mid-band fit, slope and y-intercept of the regression line are determined. Four extreme points are obtained from the *eBTF*: maximum, minimum, and corresponding frequencies. IB, a mean intensity parameter, was also calculated, but is not shown here.

4. Attenuation Compensation

From Equation 4 we see that the difference between the reference phantom attenuation and the attenuation of overlying tissue layers may still have an impact on the backscattered signals. Below we describe three approaches for compensating for this difference in attenuation. A total of four attenuation compensation approaches are investigated, reference phantom alone and the three approaches described below.

4.1 Optimum power spectral shift estimation (OPSSE)

The optimum power spectral shift estimator (OPSSE) is a correlation-based centroid shift estimator used to approximate local attenuation coefficients in soft tissue [30]. It is a variation of the hybrid method [31] developed by Kim and Varghese that estimates the slope of attenuation using a combination of the reference phantom method and a spectral cross-correlation algorithm that best minimizes the variance in the estimation. While the approach was investigated to obtain an intensity insensitive measure of the attenuation properties of the target tissue [30], we apply it here as a method to compensate for attenuation arising from the tissue path to and from the target tissue that may not be accounted for by phantom normalization.

To calculate the slope of attenuation of the plaque tissue medium, RF data from a large window of interest (≈10mm x 5.5mm) was divided into 5 axially consecutive blocks with 75% overlap. An average power spectrum was calculated for each b4.1lock and normalized to a reference phantom. Following normalization, a Gaussian function $G(f)$ at the transmit center frequency f_c with a bandwidth σ comparable to the transmit pulse,

$$G(f) = e^{\left\{ -\frac{(f-f_C)^2}{2\sigma^2} \right\}}, \quad (6)$$

was applied to the normalized spectrum, or $eBTF(f,z)$, to yield the Gaussian filtered normalized spectrum $GS(f,z)$, where

$$GS(f,z) = e^{\left\{ -\frac{(f-f_C)^2}{2\sigma^2} \right\}} \frac{S_{TIS}(f,z)}{S_{REF}(f,z)} \; e^{4(\beta_{TIS}-\beta_{REF})fz} \quad . (7)$$

Following normalization and Gaussian-filtering of the $eBTF(f,z)$, the OPSSE approach correlates the Gaussian filtered $eBTF(f,z)$ with the weight, B_{opt}, defined as follows:

$$B_{opt} = -\frac{d}{df}\left(\frac{1}{G(f)}\right), \quad (8)$$

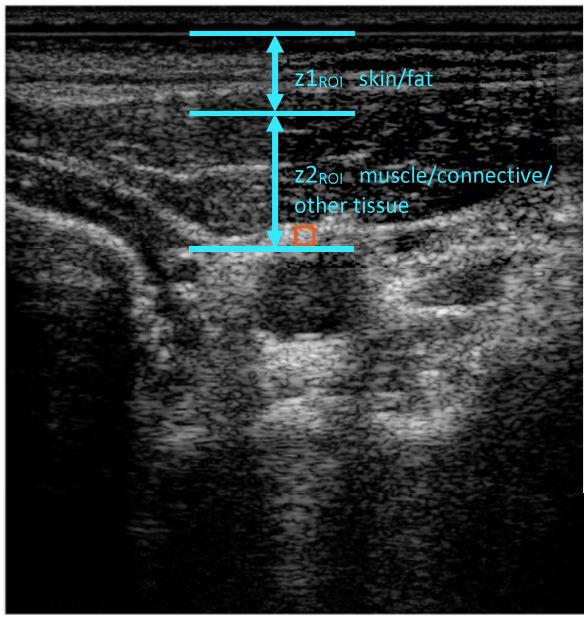

Figure 4: Sample transverse grayscale image of a carotid artery with an adventitial ROI in orange. The total depth of the ROI from the transducer surface is divided into two layers: $z1_{ROI}$, skin/fat tissue and $z2_{ROI}$, muscle, connective and other tissue.

and searches for the zero to estimate the shift in center frequency. Shown in work by others [31], the shifted center frequency $f_c(z)$ at depth z, is related to the slope of the attenuation coefficient of the target region, β_{TIS}, by

$$\beta_{TIS}(db/cm \cdot MHz) = \frac{1}{4\,\sigma^2}\frac{df_c(z)}{dz} + \beta_{REF} \quad (9)$$

A linear least-squares fit of consecutive center frequency shift estimates along depth was used to calculate the slope of the attenuation coefficient of the tissue layers overlying the carotid artery and thus provide an additional attenuation compensation for propagation beyond the reference phantom.

4.2 One-Step and Two-Step Adventitial Based Approaches

Adventitia has been used as a reference for tissue characterization [32]. Our adventitia-based attenuation compensation approaches use backscatter from the adventitia of normal subjects as a reference to compensate for attenuation differences between the reference phantom and tissue layers overlying the ROI's. The adventitia of 6 normal subjects was analyzed and 209 ROI's were obtained. An average power spectrum was estimated for each ROI and normalized by the reference phantom as described above to obtain the corresponding *eBTF*. A linear least-squares regression model was obtained for each frequency bin within the bandwidth by fitting a line to the graph of $eBTF(f,z)$ versus depth for all 209 ROI's.

The one-step approach defined a single depth, z_{ROI}, as the distance from the surface of the transducer to the starting edge of the ROI. The 2-step approach breaks this path into two pieces based on the user defined boundary between the skin/fat layer ($z1_{ROI}$) and the underlying muscle layer ($z2_{ROI}$) as shown in Figure 4.

An additional attenuation compensation was applied to each plaque ROI by inserting the ROI depth into the regression equations and subtracting the output from the $eBTF(f,z)$:

$$One\ Step\ Correction = -[\, m_{adv}(f) \cdot z_{ROI} + b_{adv}(f)\,] \quad (10)$$

$$Two\ Step\ Correction = -[\, m1_{adv}(f) \cdot z1_{ROI}(f) + m2_{adv}(f) \cdot z2_{ROI} + b_{adv}(f)\,] \quad (11)$$

where $m_{adv}(f)$ and $b_{adv}(f)$ are the slope and intercept, respectively, obtained from the linear fit of the adventitial $eBTF(f,z)$ and z_{ROI} is the depth of the start of the tissue sample ROI from the transducer.

5. MACHINE LEARNING CLASSIFICATION

5.1 Machine learning algorithms

Two machine learning (ML) classification algorithms were explored: random forest classification and multi-class support vector machine (SVM) classification. The Matlab functions *treebagger* [33] and *fitcecoc* [34] were used to produce the random forest and SVM models respectively The random forest algorithm, introduced in its entirety by Brieman in 2001 [35], uses an ensemble of decision trees on random subset of data samples, then combines the predictions from each of them to select the best output by means of voting.. A support vector machine is a binary learning model [36]. Thus a multiclass approach was implemented for the SVM model based on error-correcting output codes (ECOC) combined with SVM binary learners [37]. An error-correcting output code (ECOC) model with a one vs one coding design was used to reduce the 3-class problem to a set of binary classification problems as described by Dietterich and Bakiri [38].

Previous work in our lab has shown overlap in spectral parameter values of carotid plaque tissue types [39]. This presents an issue where SVM classification problems are not linearly separable, meaning they do not have a simple hyperplane as a useful separation criterion. In these cases, it is useful to apply transformations to the data, which map the data points from the original space into a higher dimensional feature space [40]. The goal is that the classes will become linearly separable after transformation to the higher dimensional space, enabling us to fit a decision boundary, or hyperplane, that separates the classes and make predictions [40]. Three commonly used kernel functions were used in this study: Linear, Polynomial and Gaussian Radial Basis Function (RBF).

5.2 Dataset

ML was used to classify plaque tissue ROIs as F, H/NC or Ca. Our initial dataset contained a total of 737 ROI's (F = 125, H/NC = 477 and Ca = 135) from 92 of the 119 subjects. Data was not obtained from some subjects due to an inability to match the serial histology to the ultrasound data. To form the training and testing subsets, a random subsample of 125 H/NC ROI's was selected to balance the dataset. A balanced set minimizes bias that can occur during training. Stratified random sampling was used to partition this smaller dataset into a training set and test set with the same relative class frequencies in each. The test set contains about one-third of the data (F = 37, H/NC = 37 and C = 41 from 59 subjects), while the remaining two-thirds was used for training. While outliers were included in the test set to obtain an unbiased measure of the classifier performance, for simplification, they were removed from the training set by keeping only data within two standard deviations of the mean of IB and slope for each tissue type. The final training set consisted of 80 F, 79 H/NC and 89 C ROI's from 80 subjects where 22 ROI's were excluded for being outliers.

5.3 Hyperparameter Optimization

The first step of the classification process was to optimize model hyperparameters to select the best performing predictive model. Because a separate validation set was not feasible with our sample size, model

hyperparameters were tuned using cross-validation: out-of-bag (OOB) error for the random forest classifier, and k-fold cross validation for the SVM classifier. The OOB error is commonly considered as a proxy of the performance estimation obtained using an independent test dataset or through resampling-techniques such as cross-validation. All random forest models were grown to 100 trees, and the OOB error was used to optimize the maximum number of decision splits per tree, minimum leaf size and number of predictors to select at random for each split. A list of all the optimized hyperparameters for both ML models and their ranges of values are presented in Table 1.

TABLE 1 Optimized hyperparameters for each ML algorithm and their range of values

Machine Learning Algorithm	Cross-validation Method	Optimized Hyperparameters	Range of Values
Random Forest	Out of Bag Error	Min. Leaf Size	Integer values in the range [1, 20]
		Max No. decision splits	3^n where n = {1,2,3,4,5}
		No. of predictors	Integer values in the range [1, 16]
Support Vector Machine	K-fold Cross Validation[a]	Box constraint	10^n where n = {-3,-2,-1,0 1 2 3}
		Kernel Scale[b]	10^n where n = {-3,-2,-1,0 1 2 3}
		Kernel type	Linear, Polynomial and Gaussian RBF
		Polynomial degree[c]	Integer values in the range [2, 4]

[a] K=5 for this study
[b] Used only with Gaussian RBF kernels
[c] Used only with polynomial kernels

K-fold cross-validation involves randomly dividing a set of observations into k groups, or folds, of approximately equal size. The first fold is treated as a validation, or hold-out, set, and the model is trained on the remaining k − 1 folds. This is repeated with each of the folds being the hold-out. For linear and polynomial kernel functions, only box constraint is optimized, while both box constraint and kernel scale are optimized simultaneously for Gaussian RBF kernels. Box constraint controls the trade-off between maximizing the margin of separation and minimizing the training error. This allows us to create a 'soft-margin' SVM, which allows for some points to be misclassified at the expense of penalizing the cost function proportionally to the degree of misclassification. In addition to the three kernel types, polynomials of different degree were also explored.

Cross-validation error was evaluated for all combinations of the 4 attenuation compensation methods (Reference Phantom, OPSSE, 1-step and 2-step adventitia) and 3 spectral parameter types as described in step 6 of section 3.1 (fundamental parameters only, harmonic features only, or a combination of all fundamental and harmonic parameters), for a total of 12 models. The combination of parameters with the lowest cross-validation error for each ML algorithm was selected for training and testing the final classifiers. Classification accuracy, sensitivity, specificity, and Kappa statistic for these models are reported in this paper.

5.4 Statistical Measures

Following hyperparameter optimization, the best performing classification model for each ML algorithm was using the full training set. The predicted outcomes of the optimized models were compared to the known plaque types from histology to obtain the predictive accuracy, sensitivity and specificity: An overall measure of agreement between the predictions and histology interpretation was determined by the Kappa statistic κ. The measure falls between 0 and 1, where zero indicates agreement due purely to chance and one is perfect

agreement [41]. The most commonly used scale rates a kappa of 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial and 0.81 to 0.99 as almost perfect agreement with histology [42]. These statistical measures were used to determine the overall best classification model.

6. RESULTS

6.1. Hyperparameter optimization results

The grid-search method was used to tune the hyperparameters of each model. With this technique, we simply built cross-validated classification models for each possible combination of all the hyperparameter values provided and selected the combination of parameters which produced the lowest cross-validation error. Fig. 5 shows the cumulative OOB error with increasing number of trees in the random forest for each of the 12 RF models. RF models were grown to 100 trees and produced a maximum OOB error ≤ 0.3 with less than 5% difference among all 12. Table 2 lists the OOB and cross-validation errors for each of the 12 models following parameter optimization. For SVM models, the linear kernel algorithm performed as well as or better than the other kernels for all 12 combinations of parameter inputs. For both ML classification methods, models with a reference phantom or OPSSE attenuation compensation and combined spectral parameters had the lowest OOB (0.25) or cross-validation error (0.29). There was no apparent trend in hyperparameter values within the random forest group; values for minimum leaf size, maximum number of splits and number of predictors was 9, 27 and 2 for the reference phantom combined model compared to 3, 81 and 5 for the combined OPSSE. However, both optimal SVM models had a value of 0.01 for box constraint.

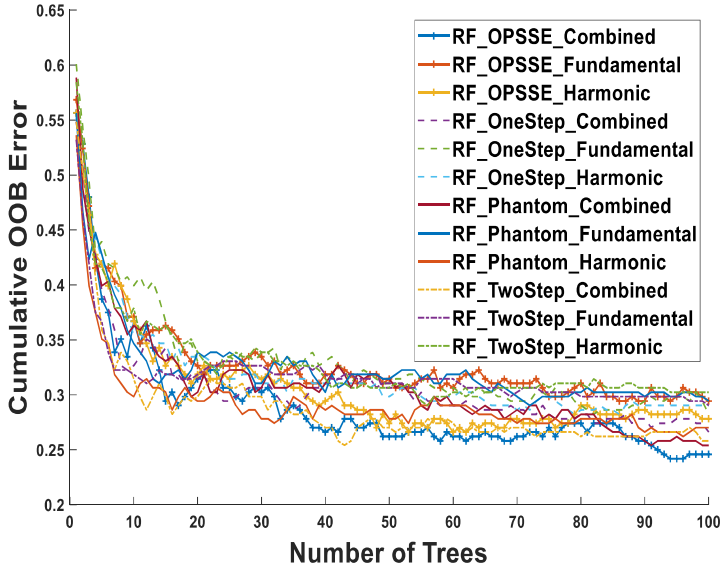

Figure 5: Cumulative out of bag error vs number of trees for each optimized random forest classification model. Each forest was grown out to 100 trees.

TABLE 2 Cross-validation error for machine learning algorithms with optimized hyperparameters. Reported values are Out-of-bag error for Random Forest models and 5-fold cross-validation error for Multiclass Support Vector Machine (SVM) models.

CLASSIFICATION MODEL		CROSS-VALIDATION ERROR	
Attenuation Compensation	Spectrum	Random Forest	SVM
Reference Phantom	Fundamental	0.29	0.31
	Harmonic	0.27	0.32
	Combined	**0.25**	**0.29**
One Step Adventitia	Fundamental	0.28	0.31
	Harmonic	0.29	0.33
	Combined	0.27	0.30
Two Step Adventitia	Fundamental	0.29	0.30
	Harmonic	0.30	0.32
	Combined	0.26	0.31
OPSSE	Fundamental	0.29	0.31
	Harmonic	0.28	0.32
	Combined	**0.25**	**0.29**

* Polynomial kernel

6.2. Evaluation of optimal classification systems

Minimum OOB or cross-validation error for each cross-validated model was used for selecting the best combination of input parameters and ML classification model. Each model was trained using its optimal hyperparameters, and the accuracy assessment of each classification system was based on the confusion matrix constructed from the predicted and known values of the test data set. Plaque ROI's were classified into three types: Calcium (Ca); Fibrous/Fibro-Fatty (F); Hemorrhagic and/or Necrotic Core (H/NC). Predictive accuracy, sensitivity, and specificity for each class in the test data set is displayed in Table 3. The overall accuracy and kappa statistic are also listed. Regarding overall performance and tissue-specific predictive power, all models produced similar results.

TABLE 3 Predictive accuracy, H/NC accuracy, sensitivity and specificity, and kappa statistic for select independently evaluated classification models. Tested the best performing of each type of classifier.

STATISTIC†		COMBINED PARAMETERS			
		Random Forest		Support Vector Machine	
		Reference Phantom	OPSSE	Reference Phantom Linear	OPSSE Linear
Predictive Accuracy		0.68 ± 0.09	0.69 ± 0.08	0.65 ± 0.09	0.65 ± 0.09
Fibrous	Accuracy	0.84 ± 0.07	0.84 ± 0.07	0.84 ± 0.07	0.84 ± 0.07
	Sensitivity	0.83 ± 0.12	0.84 ± 0.12	0.73 ± 0.14	0.73 ± 0.15
	Specificity	0.85 ±0.08	0.85 ± 0.08	0.90 ± 0.07	0.90 ± 0.07
Hemorrhagic and/or necrotic core	Accuracy	0.71 ± 0.08	0.74 ± 0.08	0.68 ± 0.09	0.68 ± 0.09
	Sensitivity	0.43 ± 0.16	0.49 ± 0.16	0.43 ± 0.16	0.43 ± 0.17
	Specificity	0.85 ± 0.09	0.86 ± 0.07	0.79 ± 0.09	0.79 ± 0.10
Calcium	Accuracy	0.80 ±0.07	0.81 ± 0.07	0.78 ± 0.08	0.78 ± 0.08
	Sensitivity	0.76 ± 0.13	0.76 ± 0.13	0.78 ± 0.13	0.78 ± 0.12
	Specificity	0.82 ±0.09	0.84 ±0.08	0.78 ± 0.09	0.78 ± 0.11
Kappa Statistic‡		0.52 ± 0.07	0.54 ± 0.06	0.49 ± 0.07	0.49 ± 0.07

† ± margin of error based on 95% confidence interval
‡ ± Standard error

7. DISCUSSION

The primary focus of this work was to determine if the nonlinearly generated second harmonic portion of the backscatter signal is useful for spectral based tissue characterization and to compare the effectiveness of a number of attenuation compensation approaches for improving tissue classification. The attenuation compensation approaches all start with applying the reference phantom data to normalize the backscattered signal from plaque tissue. Then the OPSSE and adventitial based approaches provide small adjustments, $\leq$ 0.15 dB/cm-MHz, and our results demonstrate that there might be a slight advantage in using the OPSSE but it comes at a significant computational cost.

The fact that the reference phantom approach provides reasonable results supports prior *in vivo* spectral based tissue characterization efforts which utilized a 0.5 dB/cm-MHz attenuation compensation for a number of different applications: nerves [43], prostate [44], and carotid plaque [25], [32]. The attenuation compensation approaches here are far from exhaustive but do demonstrate that the relatively simple reference phantom approach is a reasonable choice for spectral analysis-based tissue characterization of the carotid plaque.

This work processed both the fundamental and the harmonic portions of the signal, using identical steps. That is, we treated the nonlinearly generated second harmonic as if it was linearly propagating. This treatment can be successful if the amplitude of the nonlinearly generated second harmonic is comparable across all subjects and the reference phantom when backscattered by the target tissue at a given depth. Assuming that density and velocity are sufficiently consistent across subjects and the phantom, then from equation 5, the remaining elements are the coefficient of nonlinearity and the attenuation where differences in the attenuation will have a larger effect arising from its place in the exponent. Our analysis has shown that the phantom attenuation of 0.5 dB/cm-MHz is reasonable for tissue differentiation which indicates the harmonic backscatter from the phantom may be adequate to normalize the harmonic backscatter from the tissue. For the coefficient of nonlinearity ($1 + B/2A$), the variation across soft tissues is less than a factor of two and ranges from roughly 4 to 6 [29]. The nonlinearity parameter for the phantom material (zerdine, CIRS) is not known but was assumed to be in the range of other tissue mimicking materials where the coefficient of nonlinearity ranges from 3.7 to 6.05 [45]. This assumption is supported by the fact that the error rates reported in table 2 are comparable for the fundamental and harmonic bands. Our results indicate that treating the harmonic portion of the bandwidth as a 'linear' signal appears useful for tissue characterization. Thus we are effectively using the nonlinearly generated second harmonic portion the backscatter as an extension of the usable bandwidth.

Our long-term aim is to improve clinical risk stratification for CVE arising from carotid stenosis. At this time, risk stratification is heavily dependent on the degree of stenosis with information regarding the composition and morphology largely unavailable. Ultrasound has the potential to provide this information at the point-of-care in a non-invasive approach, since it is relatively inexpensive and well tolerated by the patient population. Pulse wave imaging [46], ARFI [47]–[49], and grayscale median [6], [8] approaches have all relied on analyzing longitudinal slices of the carotid artery. The strength of this approach is that for most carotid plaques the entire length of the plaque can be sampled in a single frame, and it minimizes issues arising from anisotropy by insonifying only the top and bottom of the carotid artery. A weakness of this approach is that carotid plaque is a three-dimensional disease and dangerous plaque can develop on the lateral portions of the artery and an image that includes only the anterior and posterior sides may miss these eccentric plaques. To make a robust clinical tool, this three-dimensional aspect of the plaque will need to be addressed.

In thinking about a future three-dimensional application, the work presented in this paper, chose to use a transverse imaging approach. This approach enabled the use of ROI's from all sides of the carotid plaque when training the machine learning algorithms so that in the future the algorithms will be applicable to three dimensional applications. An approximate kappa statistic of 0.5 suggests that this combination of input spectral parameters works moderately well at distinguishing between fibrous, hemorrhagic and/or necrotic core, and calcified ROI's as determined by histology. An extended analysis of other independent parameters that are sensitive to the shape of the power spectrum and to the nonlinearly generated second harmonic may improve discrimination.

While not the primary focus of this study, using a machine learning strategy to combine multiple US parameters for the purpose of distinguishing carotid plaque components, namely calcium, collagen, necrosis and hemorrhage, has also been recently addressed by others. A roughly two to eight fold increase in contrast-to-noise ratio (CNR) between soft (hemorrhage and necrosis) and stiff (collagen and calcium) plaque components was seen using a combination of ARFI parameters versus any one parameter [50]. Combining

ARFI data also achieved statistically higher CNRs between tissues of like stiffness than using any one or combination of two parameters [51]. In another study, higher areas under receiver-operating curves (AUC) were achieved with a combination of elastography, homodyned-K and B-mode grayscale features for identifying areas of lipid and calcium and ruptured fibrous caps [32]. While this study did not attempt to compare intensity-based analysis to spectral analysis, the results presented here indicate that a combination of intensity-based parameters, particularly IB, and linear fit spectral parameters may be better than IB alone [12]–[14], [23] at distinguishing hemorrhagic an/or necrotic core tissue from fibrous and calcified tissue, both of which have been associated with more stable plaques [52], [53]. Attenuation coefficient computations similar to the attenuation compensation measurements used in this paper have been used for tissue characterization and provide a significant correlation with whole plaque surgical scores (related to overall percentage) of hemorrhage [25]. It is important to note that these classification features are computed from a single two-dimensional ultrasound acquisition, generally a longitudinal view, which provides a limited sampling of the three-dimensional carotid plaque.

A primary limitation of this work is that it only used histology as a comparison, rather than clinical outcome measures (e.g. symptomatic vs asymptomatic). The limited sampling at roughly 1cm (needed to maintain independence between matched slides) cannot truly sample the extent of the carotid plaque and could easily miss areas of high CVE risk. In addition, matching ROI's in histology to ROI's in the grayscale image is not exact: (1) the histological processing can warp and damage the tissue, (2) the surgical excision cuts the plaque longitudinally and can disturb portions of plaque leading to missing regions within the histological slides, (3) histological sections are microns in thickness as compared to hundreds of microns for the lateral size on ultrasound beam. Thus, a perfect match between histology and an ultrasound image is not possible. Anisotropy was ignored for this effort even though atherosclerotic plaque does exhibit anisotropy [54]. Also, differences in velocity which effects diffraction were ignored. In order to combine types that are associated with high CVE risk, all areas for necrotic core and mild to severe hemorrhage are classified as one tissue type even though their composition is different.

8. CONCLUSION

For spectral based approaches to succeed in determining the composition of carotid plaque, proper attenuation compensation is required. These results indicate that a reference phantom approach with a 0.5 dB/cm-MHz attenuation coefficient works as well as more complicated approaches presented here to account for the total ultrasonic attenuation. This reference phantom approach also enables the use of the nonlinearly generated second harmonic bandwidth for tissue characterization. The best results came from using input spectral parameters derived from both the fundamental and second harmonic bands.

References

[1] "Global Health Estimates: Life expectancy and leading causes of death and disability." https://www.who.int/data/maternal-newborn-child-adolescent/monitor (accessed Jan. 20, 2021).

[2] Virani Salim S. *et al.*, "Heart Disease and Stroke Statistics—2020 Update: A Report From the American Heart Association," *Circulation*, vol. 141, no. 9, pp. e139–e596, Mar. 2020, doi: 10.1161/CIR.0000000000000757.

[3] A. Saxena, E. Y. K. Ng, and S. T. Lim, "Imaging modalities to diagnose carotid artery stenosis: progress and prospect," *Biomed. Eng. OnLine*, vol. 18, May 2019, doi: 10.1186/s12938-019-0685-7.

[4] D. P. Howard *et al.*, "Symptomatic carotid atherosclerotic disease: correlations between plaque composition and ipsilateral stroke risk," *Stroke J. Cereb. Circ.*, vol. 46, no. 1, pp. 182–189, Jan. 2015, doi: 10.1161/STROKEAHA.114.007221.

[5] J. S. McNally *et al.*, "Intraluminal thrombus, intraplaque hemorrhage, plaque thickness, and current smoking optimally predict carotid stroke," *Stroke*, vol. 46, no. 1, pp. 84–90, Jan. 2015, doi: 10.1161/STROKEAHA.114.006286.

[6] U. R. Acharya *et al.*, "Atherosclerotic plaque tissue characterization in 2D ultrasound longitudinal carotid scans for automated classification: a paradigm for stroke risk assessment," *Med. Biol. Eng. Comput.*, vol. 51, no. 5, pp. 513–523, May 2013, doi: 10.1007/s11517-012-1019-0.

[7] T. Elatrozy, A. Nicolaides, T. Tegos, A. Z. Zarka, M. Griffin, and M. Sabetai, "The effect of B-mode ultrasonic image standardisation on the echodensity of symptomatic and asymptomatic carotid bifurcation plaques," *Int. Angiol. J. Int. Union Angiol.*, vol. 17, no. 3, pp. 179–186, Sep. 1998.

[8] H. Hashimoto, M. Tagaya, H. Niki, and H. Etani, "Computer-assisted analysis of heterogeneity on B-mode imaging predicts instability of asymptomatic carotid plaque," *Cerebrovasc. Dis. Basel Switz.*, vol. 28, no. 4, pp. 357–364, 2009, doi: 10.1159/000229554.

[9] R. Virmani, E. R. Ladich, A. P. Burke, and F. D. Kolodgie, "Histopathology of carotid atherosclerotic disease," *Neurosurgery*, vol. 59, no. 5 Suppl 3, pp. S219-227; discussion S3-13, Nov. 2006, doi: 10.1227/01.NEU.0000239895.00373.E4.

[10] Grønholdt Marie-Louise M., Nordestgaard Børge G., Schroeder Torben V., Vorstrup Sissel, and Sillesen Henrik, "Ultrasonic Echolucent Carotid Plaques Predict Future Strokes," *Circulation*, vol. 104, no. 1, pp. 68–73, Jul. 2001, doi: 10.1161/hc2601.091704.

[11] M.-L. M. Grønholdt *et al.*, "Macrophages are associated with lipid-rich carotid artery plaques, echolucency on B-mode imaging, and elevated plasma lipid levels," *J. Vasc. Surg.*, vol. 35, no. 1, pp. 137–145, Jan. 2002.

[12] Urbani M P *et al.*, "In vivo radiofrequency-based ultrasonic tissue characterization of the atherosclerotic plaque.," *Stroke*, vol. 24, no. 10, pp. 1507–1512, Oct. 1993, doi: 10.1161/01.STR.24.10.1507.

[13] H. Waki *et al.*, "Ultrasonic tissue characterization of the atherosclerotic carotid artery: histological correlates or carotid integrated backscatter," *Circ. J. Off. J. Jpn. Circ. Soc.*, vol. 67, no. 12, pp. 1013–1016, Dec. 2003, doi: 10.1253/circj.67.1013.

[14] M. Kawasaki *et al.*, "Noninvasive quantitative tissue characterization and two-dimensional color-coded map of human atherosclerotic lesions using ultrasound integrated backscatter: comparison between histology and integrated backscatter images," *J. Am. Coll. Cardiol.*, vol. 38, no. 2, pp. 486–492, Aug. 2001, doi: 10.1016/s0735-1097(01)01393-6.

[15] N. Katakami *et al.*, "Ultrasonic Tissue Characterization of Carotid Plaque Improves the Prediction of Cardiovascular Events in Diabetic Patients," *Diabetes Care*, vol. 35, no. 12, pp. 2640–2646, Dec. 2012, doi: 10.2337/dc12-0331.

[16] E. J. Feleppa, F. L. Lizzi, D. J. Coleman, and M. M. Yaremko, “Diagnostic spectrum analysis in ophthalmology: A physical perspective,” *Ultrasound Med. Biol.*, vol. 12, no. 8, pp. 623–631, Aug. 1986, doi: 10.1016/0301-5629(86)90183-3.
[17] F. L. Lizzi, M. Greenebaum, E. J. Feleppa, M. Elbaum, and D. J. Coleman, “Theoretical framework for spectrum analysis in ultrasonic tissue characterization,” *J. Acoust. Soc. Am.*, vol. 73, no. 4, pp. 1366–1373, Apr. 1983, doi: 10.1121/1.389241.
[18] M. F. Insana, R. F. Wagner, D. G. Brown, and T. J. Hall, “Describing small-scale structure in random media using pulse-echo ultrasound,” *J. Acoust. Soc. Am.*, vol. 87, no. 1, pp. 179–192, Jan. 1990, doi: 10.1121/1.399283.
[19] W. D. O’Brien, “Quantitative acoustical assessment of wound maturation with acoustic microscopy,” *J. Acoust. Soc. Am.*, vol. 69, no. 2, pp. 575–579, Feb. 1981, doi: 10.1121/1.385432.
[20] M. O’Donnell and J. G. Miller, “Quantitative broadband ultrasonic backscatter: An approach to nondestructive evaluation in acoustically inhomogeneous materials,” *J. Appl. Phys.*, vol. 52, no. 2, pp. 1056–1065, Feb. 1981, doi: 10.1063/1.328803.
[21] Nair Anuja, Kuban Barry D., Tuzcu E. Murat, Schoenhagen Paul, Nissen Steven E., and Vince D. Geoffrey, “Coronary Plaque Classification With Intravascular Ultrasound Radiofrequency Data Analysis,” *Circulation*, vol. 106, no. 17, pp. 2200–2206, Oct. 2002, doi: 10.1161/01.CIR.0000035654.18341.5E.
[22] A. Nair, B. D. Kuban, N. Obuchowski, and D. G. Vince, “Assessing spectral algorithms to predict atherosclerotic plaque composition with normalized and raw intravascular ultrasound data,” *Ultrasound Med. Biol.*, vol. 27, no. 10, pp. 1319–1331, Oct. 2001, doi: 10.1016/s0301-5629(01)00436-7.
[23] S. Takiuchi *et al.*, “Quantitative ultrasonic tissue characterization can identify high-risk atherosclerotic alteration in human carotid arteries,” *Circulation*, vol. 102, no. 7, pp. 766–770, Aug. 2000, doi: 10.1161/01.cir.102.7.766.
[24] T. Erlöv *et al.*, “Determining carotid plaque vulnerability using ultrasound center frequency shifts,” *Atherosclerosis*, vol. 246, pp. 293–300, Mar. 2016, doi: 10.1016/j.atherosclerosis.2016.01.019.
[25] C. N. Steffel *et al.*, “Attenuation Coefficient Parameter Computations for Tissue Composition Assessment of Carotid Atherosclerotic Plaque in Vivo,” *Ultrasound Med. Biol.*, vol. 46, no. 6, pp. 1513–1532, Jun. 2020, doi: 10.1016/j.ultrasmedbio.2020.02.015.
[26] F. D. Kolodgie, G. Nakazawa, G. Sangiorgi, E. Ladich, A. P. Burke, and R. Virmani, “Pathology of Atherosclerosis and Stenting,” *Neuroimaging Clin. N. Am.*, vol. 17, no. 3, pp. 285–vii, Aug. 2007, doi: 10.1016/j.nic.2007.03.006.
[27] A. Nair, M. P. Margolis, B. D. Kuban, and D. G. Vince, “Automated coronary plaque characterisation with intravascular ultrasound backscatter: ex vivo validation,” *EuroIntervention J. Eur. Collab. Work. Group Interv. Cardiol. Eur. Soc. Cardiol.*, vol. 3, no. 1, pp. 113–120, May 2007.
[28] L. Yao, “Reference Phantom Method for Acoustic Backscatter Coefficient and Attenuation Coefficient Measurements.,” *PhD Thesis*, Aug. 1990, Accessed: Jul. 02, 2019. [Online]. Available: https://ui.adsabs.harvard.edu/abs/1990PhDT.....39Y/abstract
[29] R. T. Beyer, “The parameter B/A,” in *Nonlinear acoustics*, D. T. Blackstock and M. F. Hamilton, Eds. San Diego, CA: Academic Press, 1998.
[30] K. Samimi and T. Varghese, “Optimum Diffraction-Corrected Frequency-Shift Estimator of the Ultrasonic Attenuation Coefficient,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 63, no. 5, pp. 691–702, 2016, doi: 10.1109/TUFFC.2016.2538719.
[31] H. Kim and T. Varghese, “Hybrid Spectral Domain Method for Attenuation Slope Estimation,” *Ultrasound Med. Biol.*, vol. 34, no. 11, pp. 1808–1819, Nov. 2008, doi: 10.1016/j.ultrasmedbio.2008.04.011.

[32] M.-H. Roy-Cardinal, F. Destrempes, G. Soulez, and G. Cloutier, “Assessment of Carotid Artery Plaque Components With Machine Learning Classification Using Homodyned-K Parametric Maps and Elastograms,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 66, no. 3, pp. 493–504, 2019, doi: 10.1109/TUFFC.2018.2851846.
[33] “Create bag of decision trees - MATLAB.” https://www.mathworks.com/help/stats/treebagger.html (accessed Aug. 27, 2019).
[34] “Fit multiclass models for support vector machines or other classifiers - MATLAB fitcecoc.” https://www.mathworks.com/help/stats/fitcecoc.html (accessed Sep. 12, 2020).
[35] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[36] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.
[37] “Multiclass model for support vector machines (SVMs) and other classifiers - MATLAB.” https://www.mathworks.com/help/stats/classificationecoc.html;jsessionid=3c6e64e61e72b0a26488b5673b26#bug0_3g-1 (accessed Sep. 12, 2020).
[38] T. G. Dietterich and G. Bakiri, “Solving multiclass learning problems via error-correcting output codes,” *J. Artif. Intell. Res.*, vol. 2, no. 1, pp. 263–286, Jan. 1995.
[39] R. J. Fedewa, S. James, S. Lyden, and D. G. Vince, “Spectral Analysis of Nonlinearly Generated Second Harmonic Backscatter for Characterization of Human Carotid Plaque,” in *2019 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2019, pp. 48–51. doi: 10.1109/ULTSYM.2019.8925709.
[40] T. Hofmann, B. Schölkopf, and A. J. Smola, “Kernel methods in machine learning,” *Ann. Stat.*, vol. 36, no. 3, pp. 1171–1220, Jun. 2008, doi: 10.1214/009053607000000677.
[41] A. J. Viera and J. M. Garrett, “Understanding interobserver agreement: the kappa statistic,” *Fam. Med.*, vol. 37, no. 5, pp. 360–363, May 2005.
[42] J. R. Landis and G. G. Koch, “The Measurement of Observer Agreement for Categorical Data,” *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977, doi: 10.2307/2529310.
[43] J. D. Klingensmith *et al.*, “Spectral Analysis of Ultrasound Radiofrequency Backscatter for the Detection of Intercostal Blood Vessels,” *Ultrasound Med. Biol.*, vol. 44, no. 7, pp. 1411–1422, Jul. 2018, doi: 10.1016/j.ultrasmedbio.2018.03.007.
[44] E. J. Feleppa *et al.*, “Ultrasonic spectral-parameter imaging of the prostate,” *Int. J. Imaging Syst. Technol.*, vol. 8, no. 1, pp. 11–25, 1997, doi: https://doi.org/10.1002/(SICI)1098-1098(1997)8:1<11::AID-IMA3>3.0.CO;2-W.
[45] F. Dong, E. L. Madsen, M. C. MacDonald, and J. A. Zagzebski, “Nonlinearity parameter for tissue-mimicking materials,” *Ultrasound Med. Biol.*, vol. 25, no. 5, pp. 831–838, Jun. 1999, doi: 10.1016/s0301-5629(99)00016-2.
[46] R. X. Li *et al.*, “Pulse Wave Imaging in carotid artery stenosis human subjects in vivo,” *Ultrasound Med. Biol.*, vol. 45, no. 2, pp. 353–366, Feb. 2019, doi: 10.1016/j.ultrasmedbio.2018.07.013.
[47] T. J. Czernuszewicz *et al.*, “Noninvasive In Vivo Characterization of Human Carotid Plaques with Acoustic Radiation Force Impulse (ARFI) Ultrasound: Comparison with Histology Following Endarterectomy,” *Ultrasound Med. Biol.*, vol. 41, no. 3, pp. 685–697, Mar. 2015, doi: 10.1016/j.ultrasmedbio.2014.09.016.
[48] T. J. Czernuszewicz *et al.*, “Performance of acoustic radiation force impulse ultrasound imaging for carotid plaque characterization with histologic validation,” *J. Vasc. Surg.*, vol. 66, no. 6, pp. 1749-1757.e3, Dec. 2017, doi: 10.1016/j.jvs.2017.04.043.

[49] T. J. Czernuszewicz *et al.*, “In vivo carotid plaque stiffness measurements with ARFI ultrasound in endarterectomy patients,” in *2015 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2015, pp. 1–4. doi: 10.1109/ULTSYM.2015.0004.
[50] G. Torres, T. J. Czemuszewicz, J. W. Homeister, M. A. Farber, and C. M. Gallippi, “A Machine Learning Approach to Delineating Carotid Atherosclerotic Plaque Structure and Composition by ARFI Ultrasound, In Vivo,” in *2018 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2018, pp. 1–4. doi: 10.1109/ULTSYM.2018.8579957.
[51] G. Torres, K. Anand, J. W. Homeister, M. A. Farber, and C. M. Gallippi, “Combination of ARFI Excitation Powers and Acquisitions at Diastole and Systole for Improving Automatic Segmentation of Vulnerable Carotid Plaque Features,” in *2020 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2020, pp. 1–3. doi: 10.1109/IUS46767.2020.9251815.
[52] Nandalur Kiran R., Hardie Andrew D., Raghavan Prashant, Schipper Matthew J., Baskurt Erol, and Kramer Christopher M., “Composition of the Stable Carotid Plaque,” *Stroke*, vol. 38, no. 3, pp. 935–940, Mar. 2007, doi: 10.1161/01.STR.0000257995.74834.92.
[53] D. Butcovan, V. Mocanu, D. Baran, D. Ciurescu, and G. Tinica, “Assessment of vulnerable and unstable carotid atherosclerotic plaques on endarterectomy specimens,” *Exp. Ther. Med.*, vol. 11, no. 5, pp. 2028–2032, May 2016, doi: 10.3892/etm.2016.3096.
[54] C.-K. Chai *et al.*, “Local anisotropic mechanical properties of human carotid atherosclerotic plaques - characterisation by micro-indentation and inverse finite element analysis,” *J. Mech. Behav. Biomed. Mater.*, vol. 43, pp. 59–68, Mar. 2015, doi: 10.1016/j.jmbbm.2014.12.004.