Sensitive three-dimensional ultrasound assessment of carotid atherosclerosis by weighted average of local vessel wall and plaque thickness change

Jieyu Cheng
Department of Electronic Engineering, City University of Hong Kong, Kowloon Tong, Hong Kong

Eranga Ukwatta
Department of Systems and Computer Engineering, Carleton University, Ottawa, Ontario, Canada

Shadi Shavakh
Imaging Research Laboratories, Robarts Research Institute, London, Ontario, Canada

Tommy W. S. Chow
Department of Electronic Engineering, City University of Hong Kong, Kowloon Tong, Hong Kong

Grace Parraga
Imaging Research Laboratories, Robarts Research Institute, London, Ontario, Canada

J. David Spence
Stroke Prevention and Atherosclerosis Research Centre, Robarts Research Institute, London, Ontario, Canada

Bernard Chiua
Department of Electronic Engineering, City University of Hong Kong, Kowloon Tong, Hong Kong

(Received 3 March 2017; revised 17 July 2017; accepted for publication 19 July 2017; published 31 August 2017)

Purpose: Vitamin B deficiency has been identified as a risk factor for vascular events. However, the reduction of vascular events was not shown in large randomized controlled trials evaluating B-Vitamin therapy. There is an important requirement to develop sensitive biomarkers to be used as efficacy targets for B-Vitamin therapy as well as other dietary treatments and lifestyle regimes that are being developed. Carotid vessel-wall-plus-plaque thickness change (VWT-Change) measured from 3D ultrasound has been shown to be sensitive to atorvastatin therapies in previous studies. However, B-Vitamin treatment is expected to confer a smaller beneficial effect in carotid atherosclerosis than the strong dose of atorvastatin. This paper introduces a sensitive atherosclerosis biomarker based on the weighted mean VWT-Change measurement from 3D ultrasound with a purpose to detect statistically significant effect of B-Vitamin therapy.

Methods: Of the 56 subjects analyzed in this study, 27 were randomized to receive a B-Vitamin tablet daily and 29 received a placebo tablet daily. Participants were scanned at baseline and 1.9/C6/0.8 yr later. The 3D VWT map at each scanning session was computed by matching the outer wall and lumen surfaces on a point-by-point basis. The 3D annual VWT-Change maps were obtained by first registering the 3D VWT maps obtained at the baseline and follow-up scanning sessions, and then taking the point-wise difference in VWT and dividing the result by the years elapsed from the baseline to the follow-up scanning session. The 3D VWT-Change maps constructed for all patients were mapped to a 2D carotid template to adjust for the anatomic variability of the arteries. A weight at each point of the carotid template was assigned based on the degree of correlation between the VWT-Change measurements exhibited at that point and the treatment received (i.e., B-Vitamin or placebo) quantified by mutual information. The weighted mean of VWT-Change for each patient, denoted by D_{VWT\text{Weighted}}, was computed according to this weight. T-tests were performed to compare the sensitivity of D_{VWT\text{Weighted}} with existing biomarkers in detecting treatment effects. These biomarkers included changes in intima-media thickness (IMT), total plaque area (TPA), vessel wall volume (VWV), unweighted average of VWT-Change (ΔVWT) and a previously described biomarker, denoted by ΔVWT_S, that quantifies the mean VWT-Change specific to regions of interest identified by a feature selection algorithm.

Results: Among the six biomarkers evaluated, the effect of B Vitamins was detected only by ΔVWT_{Weighted} in this cohort (P = 4.4 × 10⁻³). The sample sizes per treatment group required to detect an effect as large as exhibited in this study were 139, 178, 41 for ΔVWV, ΔVWT and ΔVWT_{Weighted} respectively.

Conclusion: The proposed weighted mean of VWT-Change is more sensitive than existing biomarkers in detecting treatment effects. This measurement tool will allow for many proof-of-principal
1. INTRODUCTION

Stroke is among the leading causes of death and disability worldwide, with 17.3 million people suffering a stroke each year. Over two-thirds of stroke deaths occurred in developing countries. China, as the most populous developing country, has an annual stroke mortality of 1.6 million, and the mortality rate is more than seven times higher than in the United States. Carotid atherosclerosis is a major source of emboli, which may block one of the cerebral arteries, causing ischemic stroke. Fortunately, 75–80% of vascular events can be prevented through lifestyle/dietary changes and medical therapies. Effects of these treatment and management strategies on the development of carotid atherosclerosis are required to be validated in clinical trials. Thus, parallel to the development of new therapies or management strategies, there is an important requirement for the development of treatment-specific measurement tools or biomarkers for longitudinal assessment of disease progression and regression.

Ultrasound imaging has played an important role in the development of a variety of image-based biomarkers. Carotid intima-media thickness (IMT) measured from 2D B-mode ultrasound has been shown to correlate with the risk of cardioembolic events and used in clinical trials. However, the annual rate of change of IMT is smaller than the resolution of ultrasound. Therefore, a large sample size and long duration of clinical observations are required to identify statistically significant changes. In addition, IMT is different from direct atherosclerotic plaque phenotypes as it measures the thickness in the distal far wall of the common carotid artery that has no plaque. The increase in IMT reflects mostly hypertensive medial hypertrophy that is not related to atherosclerosis and contributes to a variety of risk factors such as hypertension and blood flow dynamics. Direct plaque measurements have been introduced to address issues related to IMT. Total plaque area (TPA) measures the area of each plaque in a longitudinal view in which it appears largest. Both TPA and IMT measurements are made from 2D longitudinal images. The requirement for localizing a specific 2D image plane for measurements introduces variability to these measurements, making them suboptimal for serial monitoring of plaque burden. The development of 3D ultrasound imaging techniques has allowed for vessel wall volume (VWV) measurements, which have been demonstrated to be sensitive to atherosclerosis progression/regression. Due to the larger dynamic ranges for annual changes as compared to IMT and TPA, VWV has markedly reduced the sample size and duration (therefore cost) required to demonstrate the effect of strong-dose atorvastatin to 41 subjects in 3 months in a previous clinical study.

The cost reduction conferred by the 3D ultrasound-based biomarkers allows many proof-of-principal studies to demonstrate the efficacy of new therapies, before a large-scale study is performed to validate the finding in the pilot studies. However, as carotid atherosclerosis is a focal disease, quantifying local vessel wall thickness change may give rise to a more sensitive biomarker for treatment monitoring, which is required in the current investigation on the effect of B Vitamins as it is expected that B Vitamins would confer a smaller beneficial effect than strong-dose atorvastatin. Our group pioneered a 3D vessel-wall-plus-plaque thickness change (VWT-Change) map to display point-wise VWT-change over a 3-month period for subjects enrolled in a randomized, placebo-controlled atorvastatin trial. However, since the geometry of the 3D VWT-Change map is highly subject specific, the comparison between VWT-Change distributions observed in different subjects required visual matching, which is qualitative, subjective, and prone to observer variability. The variability in vessel geometry among individuals was adjusted in our recent study by projecting the 3D VWT-Change maps of all subjects onto a 2D carotid template. The 2D standardized map allows quantitative group-wise comparison between the VWT-Change distributions of subjects treated with atorvastatin and placebo. However, drawing clinical conclusions based on the complex VWT-Change distributions on the 2D maps was difficult as each map consists of thousands of vertices, each equipped with a VWT-Change measurement. To address this issue, we developed a concise biomarker based on a feature selection algorithm that identified regions of interest (ROI) in the 2D maps on which the VWT-Change distributions tend to be different in atorvastatin and placebo subjects. Average VWT-Change computed over the identified regions of interest was more sensitive for demonstrating the effect of atorvastatin treatment than that computed over the whole 2D map.

However, there are several limitations associated with this biomarker. First, feature selection returns a binary distribution on the 2D map, which identifies whether each point in the 2D map is included in the ROI. Points not selected were not involved in the computation of the region-based biomarker, but may still provide important information in the difference of VWT-Change distributions in the two groups. In this study, we aim to develop a biomarker based on a weighted average, which, while emphasizing regions with more prominent difference between treatment groups, takes all points in the 2D map into consideration. We hypothesize that this new biomarker will be more sensitive to treatment effects than the feature-selection-based biomarker. Second, the mutual-information-based feature selection algorithm is a forward
sequential searching algorithm that adds features one at a time without an objectively defined stopping criterion. The lack of a termination criterion produces variability in the feature-selection-based biomarker as the average \( VWT-Change \) is a function of the size of the ROI over which the average was taken. In a previous study,\(^\text{17}\) we constrained the size of the selected region to be between 10–50\% of the total number of vertices in the 2D map, and chose the selected region size to be the one that produced an average \( VWT-Change \) that discriminated the atorvastatin and placebo groups with the greatest sensitivity. The rationale for choosing these upper and lower thresholds was to select an ROI that is large enough to detect “representative” local patterns, while small enough so that we were still focusing on “local” instead of “global” patterns. However, the line between local and global patterns is hard to draw and the choices of lower and upper thresholds were based on empirical observation of the data. A weighted averaging approach can be considered as a “soft thresholding” approach in which a “hard threshold” is not required to cut off points outside the chosen ROI, thereby eliminating the variability due to the choices of these thresholds.

In this paper, we propose a weighted average measurement of the point-wise \( VWT-Change \) represented on the 2D \( VWT-Change \) map. We validate this biomarker in subjects recruited for the Diabetic Intervention with Vitamin to Improve Nephropathy (DIVINe) trial. These subjects were exposed to an increased risk of cardiovascular diseases due to diabetic nephropathy and diabetic microvascular diseases.\(^\text{18,19}\) DIVINe was a randomized, placebo-controlled trial performed to evaluate the effect of B Vitamins, which was shown to reduce plasma homocysteine, a risk factor in vascular events.\(^\text{20–22}\) We compared changes quantified using the new weighted average \( VWT-Change \) biomarker, the feature-selection-based biomarker introduced previously\(^\text{17}\) and the existing one-, two-, and three-dimensional ultrasound phenotypes, such as IMT, TPA, and VWV, to test our hypothesis that the new biomarker is the most sensitive in detecting the difference between vessel changes demonstrated by the B-Vitamin and placebo groups. Table I lists the acronyms and notations used throughout the paper for easy reference.

<table>
<thead>
<tr>
<th>Acronym/notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMT</td>
<td>Intima-media thickness</td>
</tr>
<tr>
<td>TPA</td>
<td>Total plaque area</td>
</tr>
<tr>
<td>TPV</td>
<td>Total plaque volume</td>
</tr>
<tr>
<td>VWV</td>
<td>Vessel wall volume</td>
</tr>
<tr>
<td>VWT</td>
<td>Vessel-wall-plus-plaque thickness</td>
</tr>
<tr>
<td>DIVINe</td>
<td>Diabetic intervention with vitamins to improve nephropathy</td>
</tr>
<tr>
<td>CCA</td>
<td>Common carotid artery</td>
</tr>
<tr>
<td>ICA</td>
<td>Internal carotid artery</td>
</tr>
<tr>
<td>MI</td>
<td>Mutual information</td>
</tr>
<tr>
<td>( \Delta VWT )</td>
<td>Mean ( VWT-Change ) computed over the entire 2D carotid map</td>
</tr>
<tr>
<td>( \Delta VWT_S )</td>
<td>Mean ( VWT-Change ) computed over the ROI identified by feature selection</td>
</tr>
<tr>
<td>( \Delta VWT_{Weighted} )</td>
<td>Weighted mean ( VWT-Change ) computed over the entire 2D carotid map</td>
</tr>
</tbody>
</table>

2.A. Study subjects and ultrasound image acquisition

Subjects with diabetic nephropathy were recruited by the Nephrology Clinics and Diabetes Clinics at the London Health Science Centre (London, Canada), one of the sites in the DIVINe trial, and provided written informed consent to the study protocol approved by the local institutional review board as described previously.\(^\text{23}\) Subjects required a diagnosis of type I or II diabetes and a clinical or histologic diagnosis of diabetic nephropathy. Of the 71 subjects analyzed previously,\(^\text{24}\) the carotid surfaces segmented at the baseline and follow-up ultrasound images for 15 subjects could not be accurately aligned using our existing registration technique.\(^\text{16,17}\) The remaining 56 subjects were analyzed in this study, with 27 randomized to receive a combination tablet of 2.5 mg folic acid (vitamin B9), 25 mg pyridoxine (vitamin B6) and 1 mg of cobalamin daily, and 29 randomized to receive a placebo tablet daily.

High-resolution 3D ultrasound images were obtained by translating an ultrasound transducer (L12-5, Philips, Bothel, WA, USA) with a mechanical assembly along the neck of the subject for approximately 4.0 cm. The 2D ultrasound frames from the ultrasound machine (ATL HDI 5000, Philips, Bothel, WA, USA) were digitalized at 30 Hz, and reconstructed into a 3D image.\(^\text{25,26}\) Participants were scanned at baseline and a follow-up session, which took place at a maximum of 3.5 yr after the initial scan.

2.B. 3D \( VWT-Change \) maps and 2D standardized map

The algorithms for constructing the 3D \( VWT-Change \) map and the 2D standardized map have been described elsewhere,\(^\text{16,17}\) and are briefly summarized here. First, the outer wall and the lumen boundaries of the common and internal carotid arteries (CCA and ICA, respectively) were segmented on each resliced 2D transverse plane from a 3D ultrasound image\(^\text{14}\) and reconstructed into 3D surfaces. The point-wise \( VWT \) of a carotid artery was computed by matching the outer wall and the lumen surfaces on a point-by-point basis, resulting in a 3D \( VWT \) map. The 3D \( VWT \) maps obtained at baseline and the follow-up scanning sessions for the same artery were registered using a previously described algorithm.\(^\text{16,30}\) The 3D \( VWT-Change \) map was constructed by computing the \( VWT \) difference between the baseline and follow-up 3D \( VWT \) maps for each pair of corresponding points [Fig. 1(a)]. The CCA and ICA segments of a 3D \( VWT-Change \) map were each cut, unfolded and projected onto a rectangular domain [Figs. 1(b) and 1(c)]. The area of the resulting 2D standardized map was determined by the average surface area of the
3D VWT-Change maps of the whole group of subjects involved in the study as described previously.17

2.C. Feature selection based on mutual information (MI)

Chiu et al.17 described a feature selection technique used to identify a subset of locations with VWT-Change that was most correlated to the treatment status of patients, and defined a biomarker based on the feature selection result. Since this biomarker was used as a benchmark to evaluate the biomarkers proposed in this paper, this technique was briefly described here.

Considering the VWT-Change at each location \( p_i \) on the 2D standardized map as a feature, denoted by the random variable \( F(p_i) \) associated with a realization \( f_j(p_i) \) for each subject \( j \). The feature selection method employs a sequential forward searching strategy selecting a feature in each iteration. Suppose \( S \) represents the selected feature set, which was initialized to be empty, and \( C \) be the random variable representing whether or not a subject belongs to the treatment group (i.e., \( C = 0 \) and \( C = 1 \) for subjects in the placebo and treatment groups, respectively), in each iteration, the algorithm chose the feature with maximum mutual information (MI) between the feature set \( S + F(p_i) \) and \( C \), denoted by \( I(S + F(p_i); C) \), where \( S + F(p_i) \) was constructed by adding \( F(p_i) \) to the existing feature set \( S \), which contains all features selected in previous iterations. High \( I(S + F(p_i); C) \) indicates that the knowledge of the values of the feature set \( S + F(p_i) \) has removed much uncertainty in which treatment group an artery belongs to, and therefore, important to be characterized in a sensitive biomarker. The average VWT-Change was computed over the region of interest \( S \) for each subject \( j \), resulting in the following biomarker:

\[
\Delta VWT_{Sj} = \frac{1}{\#S} \sum_{F(p_i) \in S} f_j(p_i),
\]

(1)

where \( \#S \) represents the total number of features in the set \( S \). \( \Delta VWT_{Sj} \) was demonstrated to be more sensitive in differentiating the placebo and atorvastatin groups than the subject-based average VWT-Change computed over all points on the 2D standardized map,17 denoted by \( \overline{\Delta VWT} \). To simplify the notations, we drop the \( j \) subscript hereafter with the understanding that \( \overline{\Delta VWT} \) and \( \Delta VWT_S \) indicate a scalar parameter computed for each subject.

In addition to the shortcomings already described in the introduction, the MI criterion used in the feature selection approach relates to the absolute difference in VWT-Change at a point between treatment groups, but does not indicate which group had a larger progression/regression. In the previously described study,17 although the progression of VWT in the placebo group was larger than in the treatment group at most points within the region of interest \( S \) described above, there existed points at which VWT progression in the treatment group was larger. To facilitate later discussion, we define the pointwise treatment group difference of VWT-Change, denoted by \( GD\Delta VWT(p_i) \), as follows:

\[
GD\Delta VWT(p_i) = \Delta VWT_0(p_i) - \Delta VWT_1(p_i),
\]

(2)

where \( \Delta VWT_c(p_i) = \frac{1}{N_c} \sum_{j \in Group_c} f_j(p_i) \) with \( c = 0 \) or \( 1 \) representing the placebo and treatment groups, respectively, and \( N_0 \) and \( N_1 \) denote the total number subjects in the placebo and treatment groups, respectively. To simplify the notation, \( GD\Delta VWT \) is understood as a function of \( p_i \) and the argument \( p_i \) is omitted in the following discussion.

2.D. Biomarkers based on mutual-information-weighted average of VWT

To address the issues described above related to the previously described biomarkers, we considered two important factors in designing the point-wise weighting factor used in biomarker definition. First, the weights must be signed, reflecting the signed point-wise VWT-Change between treatment groups [i.e., \( GD\Delta VWT \) in Eq. (2)]. Second, although MI fails to indicate the sign of GD, it was able to identify the difference of VWT-Change between treatment groups with intersubject VWT-Change variability taken into consideration.17 Thus, \( I(F(p_i); C) \) was kept to be a major component in the definition of the weighting factor. With these two factors considered, we arrived at a definition of the point-wise weighting factors with the form:

\[
W_i = \text{sgn}(GD\Delta VWT(p_i))I(F(p_i); C),
\]

(3)

where \( \text{sgn} \) represents the sign function. For each subject \( j \), the new subject-based biomarker was defined as:

[Image 48x496 to 284x730]
\[ \Delta VWT_{\text{Weighted},j} = \sum_{i=1}^{F_{\text{total}}} W_i \Delta f_j(p_i), \]  

(4)

where \( F_{\text{total}} \) is the total number of features in the 2D standardized map. As shown in Eq. (3), computation of \( W_i \) requires an estimate of \( h(F(p_i); C) \). As the estimation of this MI was performed independently on a point-by-point basis, we omit the argument \( p_i \); when representing the random variable \( F(p_i) \) and its realizations \( f(p_i) \) thereafter. Shannon’s MI is widely used to quantify the relationship between two random variables. In our application, it is defined as:

\[
I(F; C) = \sum_{x \in \mathbb{R}} p(f, c) \log \frac{p(f, c)}{p(f)P(c)} df,
\]

(5)

where \( p(f, c) \) is the joint probability density function (pdf) of two random variables, \( F \) and \( C \). \( p(f) \) and \( P(c) \) are the marginal probabilities of these two variables. Since computation of Shannon’s MI involves integration of pdfs, many computational approaches have been described to alleviate the computational burden.\(^ {31} \) MI can be understood as the Kullback–Leibler (KL) divergence that quantifies how dependent \( F \) is on \( C \) by evaluating the dissimilarity of \( p(f, c) \) and \( p(f)P(c) \). If the two variables are independent, \( p(f, c) = p(f)P(c) \) and the KL divergence is minimized. As pointed out previously,\(^ {32} \) Shannon’s MI is by no means the only measure of dependence of the two random variables. Another such measurement easier to be computed is based on the Cauchy–Schwarz inequality and referred to as the quadratic mutual information (QMI):

\[
I_Q(F; C) = \log \frac{V_{(f,x)}^2 V_{(c)}^2 V_{(f)}^2}{V_{(c)}^2},
\]

(6)

where

\[
V_{(c)}^2 = \sum_{c=0}^{1} p^2(c) \int_{f \in \mathbb{R}} p^2(f) df
\]

(7)

\[
V_{(f)}^2 = \sum_{c=0}^{1} \int_{f \in \mathbb{R}} p^2(f, c) df
\]

\[
V_{(f,c)}^2 = \sum_{c=0}^{1} \int_{f \in \mathbb{R}} p^2(f, c) df
\]

The conditional and marginal pdfs were obtained using the Parzen window estimator:\(^ {33} \)

\[
p(f) = \frac{1}{N} \sum_{j=1}^{N} G_\sigma(f - f_i),
\]

(8)

\[
p(f|c) = \sum_{f_i \in \text{Group } c} p(f_i)G_\sigma(f - f_i),
\]

(9)

\[
P(c) = \sum_{f_i \in \text{Group } c} p(f_i),
\]

(10)

where \( N \) is the total number of subjects in both groups, and \( G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp(-\frac{x^2}{2\sigma^2}) \) is the Gaussian kernel with bandwidth \( \sigma \), which was optimized by the Silverman’s rule.\(^ {34} \) Applying the following property of the Gaussian kernel:

\[
\int_{x \in \mathbb{R}} G_\sigma(x - x_i)G_\sigma(x - x_j) dx = G_\sigma(\sqrt{2}\sigma(x_i - x_j)),
\]

(11)

quantities defined in Eq. (7) can be obtained:

\[
V_{(c)}^2 = \left( \sum_{f_i \in \text{Group } 0} p(f_i) \right)^2 + \left( \sum_{f_i \in \text{Group } 1} p(f_i) \right)^2
\]

(12)

\[
V_{(f)}^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} p(f_i)p(f_j)G_\sigma(f_i - f_j)
\]

(13)

\[
V_{(f,c)}^2 = \sum_{f_i \in \text{Group } 0} \sum_{f_j \in \text{Group } 0} p(f_i)p(f_j)G_\sigma(f_i - f_j)
\]

(14)

\[
V_{(f,c)}^2 = \sum_{f_i \in \text{Group } 1} \sum_{f_j \in \text{Group } 1} p(f_i)p(f_j)G_\sigma(f_i - f_j)
\]

(15)

2.E. Statistical analyses

Intima-media thickness, TPA, and VWV were obtained for a population of 56 subjects as described previously,\(^ {24} \) from which \( \Delta IMT, \Delta TPV \) and \( \Delta VWV \) per year were calculated. The 2D \( VWT-\text{Change} \) maps of the left and right arteries of each patient were obtained as described in Section 2.B with \( VWT-\text{Change} \) values divided by the number of years between the baseline and the follow-up scans to obtain annual \( VWT-\text{Change} \) maps. \( \Delta VWT \), \( \Delta VWT_S \) [Eq. (1)] and \( \Delta VWT_{\text{Weighted}} \) [Eq. (4)] of each patient were computed based on the annual 2D \( VWT-\text{Change} \) maps of the left and the right arteries.

In contrast with the previous study,\(^ {17} \) in which the same set of data were used in feature selection and biomarker evaluation, a leave-one-out strategy was adopted in this study to remove potential bias introduced in computing \( \Delta VWT_S \) and \( \Delta VWT_{\text{Weighted}} \). More specifically, before evaluating \( \Delta VWT_S \) and \( \Delta VWT_{\text{Weighted}} \) for each subject, the feature set (or the ROI) had to be selected in the \( \Delta VWT_S \) case or the weighted map had to be trained in the \( \Delta VWT_{\text{Weighted}} \) case using the remaining 55 subjects. In evaluating \( \Delta VWT_S \) and \( \Delta VWT_{\text{Weighted}} \), the sampled points on the left and right
arteries were pooled together and served as feature points. In computing $\Delta VWT_s$, the previously described stopping criterion of the feature selection algorithm was applied. Briefly, starting from the time when 10% of total feature points were selected, a two-sample $t$-test involving of the B-Vitamin and placebo groups was performed each time after 1% additional feature points were collected, thereby obtaining a $P$-value each time, which reflects the sensitivity of $\Delta VWT_s$ in discriminating between the two groups. The algorithm terminated when a local minimum of the $P$-value was attained.

The ability of the six response biomarkers in differentiating B-Vitamin and placebo patients was evaluated by the two-sample $t$-tests and receiver-operating characteristic (ROC) analysis. The $P$-values obtained in the two-sample $t$-tests performed for the six biomarkers quantify how sensitive each biomarker was to the difference in wall and plaque burden change between the B-Vitamin and the placebo groups. The area under the ROC curve (AUC) indicates how well a particular biomarker discriminates B-Vitamin and placebo patients across a series of cutoff values. The effect of B-Vitamin treatment was more sensitively measured by a biomarker with a larger AUC.

To assess the ability of $\Delta VWT_{\text{Weighted}}$ in detecting the actual difference between the B-Vitamin and placebo groups, it is important to check whether the detected difference was a result of overtraining the weighting map, which is manifested by a significant difference between any two groups. This evaluation was performed by randomly assigning B-Vitamin and placebo subjects into two groups. To avoid bias in a truly random assignment, the two groups were composed of equal numbers of B-Vitamin and placebo subjects. This experiment was repeated 30 times, and a $t$-test was performed in each trial to evaluate whether the difference between the two groups was statistically significant.

For each biomarker, the sample size required to detect a specific effect size $\delta$ was calculated using the following equation:

$$n = \left(\frac{z_{\alpha/2} + z_{\beta}}{\sigma_0 + \sigma_1}\right)^2 \frac{\sigma_0^2}{\rho_0^2} + \frac{\sigma_1^2}{\rho_1^2} \cdot \frac{r^2}{\delta^2},$$

where $p(Z > z_{\beta}) = \beta$ with $Z \sim N(0,1)$. $\sigma_0$ and $\sigma_1$ are the standard deviations of the biomarker associated with the placebo and the B-Vitamin groups, respectively. The sample size were computed at 90% statistical power (i.e., $\beta = 0.1$) and a significance level of $\alpha = 0.05$.

3. RESULTS

Measurements were obtained at baseline and 1.9±0.8 yr (range: 0.3 to 3.5 yr). Changes for all measurements were normalized to their annual rates in all statistical analyses. Figure 2 shows the $VWT$ maps and the 3D ultrasound images obtained at the baseline and follow-up scans for a subject who received B Vitamins. The annual $VWT$-Change map is

![VWT maps](image1.png)

**FIG. 2.** $VWT$ maps (a, b) and the 3D ultrasound images (c, d) of an example subject from the B-Vitamin group at baseline and follow-up. The black lines in (a) and (b) correspond to the locations of the transverse imaging slices shown in (c) and (d). The white lines in (c) and (d) indicate the location where the CCA was cut in the generation of the 2D maps. (e) shows the annual $VWT$-Change map. The time between the baseline and the follow-up visits was 365 days. The annual rates of change of the six biomarkers are listed below: $\Delta IMT = -0.05$ mm, $\Delta TPA = -0.03$ mm$^2$, $\Delta VWV = -418.46$ mm$^3$, $\Delta VWT = -0.37$ mm, $\Delta VWT_s = -0.13$ mm, $\Delta VWT_{\text{Weighted}} = -0.36$ mm. [Color figure can be viewed at wileyonlinelibrary.com]
also displayed. The annual rates of changes of six biomarkers are provided in the caption with all biomarkers indicating regression. The magnitude of regression quantified by $\Delta VWT_{\text{Weighted}}$ is larger than $\Delta VWT_S$, suggesting a higher sensitivity for $\Delta VWT_{\text{Weighted}}$ in detecting changes. Figure 3 shows the $VWT$ maps and the 3D ultrasound images obtained at the baseline and follow-up scans, as well as the annual $VWT$-Change map for a subject in the placebo group. Of particular interest is that while the $VWT$-based biomarkers and $VWV$ showed progression, $\Delta IMT$, $\Delta TPA$, $\Delta VWT_S$ suggested regression. $\Delta VWT_S$ was not sensitive to the difference between $VWT$-Changes in the example B-Vitamin and placebo subjects. This is caused by the high dependence of $\Delta VWT_S$ on the number of points selected in the feature selection algorithm. The effect of this dependence will be illustrated in detail later in this section. The small negative changes in $\Delta IMT$ and $\Delta TPA$ may be caused by measurement variability. As $\Delta IMT$ and $\Delta TPA$ were measured in 2D longitudinal images, which are difficult to reproduce in different scanning sessions, these two measurements are not optimal for serial monitoring of small changes. Furthermore, the $\Delta IMT$ change in a single subject cannot be reliably monitored as the annual change is smaller than the ultrasound resolution ($\sim 0.2$–$0.3$ mm). Therefore, to confidently monitor changes in $\Delta IMT$ and $\Delta TPA$, a large sample size and long duration are required, which pilot studies cannot afford. The development of 3D-ultrasound-based and the currently proposed biomarkers all works towards making proof-of-principal studies more affordable.

Figure 4 shows the average annual $\Delta VWT$ maps obtained for the left and right arteries for B-Vitamin and placebo subjects. The first row [i.e., Figs. 4(a) and 4(b)] shows the average annual $VWT$-Change maps constructed for the B-Vitamin subjects. The second row [i.e., Figs. 4(c) and 4(d)] shows the average annual $VWT$-Change maps constructed for the placebo subjects. The third row [i.e., Figs. 4(e) and 4(f)] shows the difference in the average annual $VWT$-Change between the placebo and B-Vitamin groups. Figure 4(f) shows a substantial difference in annual $\Delta VWT$ between the two groups near the carotid bulb at the same longitudinal level of the bifurcation as pointed to by the arrows in Figs. 4(b) and 4(d). In contrast, the difference between annual $\Delta VWT$ experienced by subjects in the two groups in the left arteries was not as pronounced as in the right arteries, except at the regions pointed to by the white arrows in Figs. 4(a) and 4(c) in which the mean $VWT$-Change for the B-Vitamin subjects was higher than that of the placebo subjects. The elevated $VWT$-Change in this region was mainly contributed by the three arteries shown in Fig. 5 and may be related to vascular wall or adjacent plaque remodeling caused by large thickness decrease in the vicinity. As only 3 of the 27 subjects exhibited this elevation in $VWT$, they can be considered as outliers in the population. As $MI$ is less sensitive to outliers than first-order statistics such as the mean, the dark region highlighting the difference between the B-Vitamin and the placebo groups pointed to by the white arrow in the weighted map trained by the entire cohort of 56 subjects shown in Fig. 5(g) was much more sensitive to the difference in $VWT$.
less diffused than that appeared in the corresponding region shown in the mean group difference map shown in Fig. 4(e). For this reason, the sensitivity of $D_{VWT}$ remains high as described in the next paragraph despite the elevation of $VWT$ in the B-Vitamin subjects mentioned herein.

The means and standard deviations of the new biomarker, $\Delta VWT_{Weighted}$, for the placebo and B-Vitamin groups were tabulated in Table II and compared with values obtained using five existing biomarkers. A $t$-test was performed to evaluate the sensitivity of each biomarker in discriminating the two groups. At the significance level of 5%, only $D_{VWT}$ was able to detect differences. Of particular interest is that $D_{VWTS}$ was insignificant ($P = 0.71$). As described in Section 2.E, the number of feature points selected was determined in 56 leave-one-out trials. The difference in the number of feature points selected in these 56 trials contributed to the variability of $\Delta VWT_{S}$, rendering it insensitive to the $VWT-Change$ difference between the two groups. Even if the number of feature points was fixed in all 56 feature selection trials, the difference between the two groups was not significant. Nine experiments were performed in which $\Delta VWT_{S}$ was computed with the number of features fixed to be from 10 to 50% with a 5% increment. $P$-values ranged from 0.06 to 0.59 in these experiments. These experiments demonstrated the high sensitivity of the $P$-values to the number of feature points used, which is a major deficiency of $\Delta VWT_{S}$ addressed by $\Delta VWT_{Weighted}$ developed in this study.

To evaluate whether the weighting map used for computing $\Delta VWT_{Weighted}$ was overtrained, $\Delta VWT_{Weighted}$ values were computed in 30 random group reassignment trials as described in Section 2.E. No significant difference was detected, and $P$-values in these 30 trials ranged from 0.21 to 1.0.

Figure 6 shows the individual responses of patients in the B-Vitamin and placebo groups. Receiver-operating characteristic (ROC) analysis was performed to evaluate how well each biomarker was able to discriminate individual patients belonging to the two treatment groups. Figure 7 shows the ROC curves obtained using the five existing and the proposed biomarkers, with the area under each ROC curve (AUC) tabulated in the fourth column of Table II. The AUCs for $\Delta IMT$ and $\Delta TPA$ were under 0.5, which would have been obtained by randomly guessing the treatment group identity of each patient. In the ROC analysis, a subject with an annual rate of change measurement smaller than the threshold was classified to the B-Vitamin group. Otherwise, it was classified to the placebo group. As IMT and TPA progressions for placebo subjects tended to be smaller than those for B-Vitamin subjects, classification results obtained in the ROC analyses of IMT and TPA-Change measurements were expected to be less accurate than random guessing.

Table III tabulates the sample sizes required for $\Delta VWV$, $\Delta VWT$ and $\Delta VWT_{Weighted}$ for various effect sizes, which are represented as a percentage of the mean difference between the placebo and B-Vitamin groups obtained in this study. Sample-size calculation was not performed for $\Delta IMT$ and $\Delta TPA$ because the effects detected by these two biomarkers were reversed for B-Vitamin and placebo patients, and for $\Delta VWT_{S}$ due to its dependence on the number of selected features. For a given effect size, the sample sizes required for $\Delta VWV$ and $\Delta VWT$ were comparable and the weighted...
Vessel wall volume is a scaled version of $VWT$ and the constant of proportionality depends on the diameter of the carotid vessel. The dependence on vascular diameters may render VWV insensitive to the difference in VWT between two groups if there was a large variation in vascular diameters in the study population as demonstrated previously.\textsuperscript{37} The same statements hold for $DVWV$ and $DVWT$. Figure 8 shows the correlation between $DVWV$ and $DVWT$. In this study, the average $DVWT$ was higher in the placebo than in the B-

**TABLE II.** The means and standard deviations (in parentheses) of six ultrasound-based measurements computed for the B-Vitamin and placebo groups. $\Delta IMT$ measurements were available only for 24 patients. A $t$-test was performed for each measurement and the associated $P$-value is tabulated. The area under the ROC curve (AUC) for each measurement is also tabulated.

<table>
<thead>
<tr>
<th>Measurements</th>
<th>B-Vitamin ($n = 27$)</th>
<th>Placebo ($n = 29$)</th>
<th>$P$-value</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta IMT$ (mm/y)</td>
<td>0.01 (0.04)</td>
<td>$-0.01$ (0.06)</td>
<td>0.10</td>
<td>0.38</td>
</tr>
<tr>
<td>$\Delta TPA$ (cm$^2$/y)</td>
<td>0.11 (0.22)</td>
<td>$-0.02$ (0.36)</td>
<td>0.11</td>
<td>0.41</td>
</tr>
<tr>
<td>$\Delta VWV$ (mm$^3$/y)</td>
<td>$-9.1$ (123.6)</td>
<td>37.9 (91.1)</td>
<td>0.11</td>
<td>0.59</td>
</tr>
<tr>
<td>$\Delta VWT$ (mm/y)</td>
<td>0.01 (0.11)</td>
<td>0.05 (0.10)</td>
<td>0.16</td>
<td>0.54</td>
</tr>
<tr>
<td>$\Delta VWTS$ (mm/y)</td>
<td>$-0.127$ (0.046)</td>
<td>$-0.130$ (0.020)</td>
<td>0.71</td>
<td>0.55</td>
</tr>
<tr>
<td>$\Delta VWT_{Weighted}$ (mm/y)</td>
<td>$-0.06$ (0.23)</td>
<td>0.13 (0.05)</td>
<td>4.4$\times10^{-3}$</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Vitamin group, and the average ratio between $\Delta VWV$ and $\Delta VWT$ in the placebo group was slightly higher than that of the B-Vitamin group, resulting in an expansion of the group difference as measured by $\Delta VWV$ as compared to $\Delta VWT$ as shown in Table II. However, it is important to point out that $\Delta VWV$ measurement would have obscured the thickness change difference between the two groups if the ratio between $\Delta VWV$ and $\Delta VWT$ in the placebo was lower than that of the B-Vitamin group. Our introduction of $\Delta VWT$ in Refs. [16,17] has the advantage of removing the vascular size dependence when measuring vessel wall and plaque burden.

4. DISCUSSION AND CONCLUSION

High concentration of total homocysteine has been shown to accelerate atherosclerosis.\textsuperscript{38–40} Many investigations suggested that Vitamin B12 deficiency as a major cause of elevated homocysteine concentration.\textsuperscript{41} However, reduction of vascular events was not demonstrated by three large randomized, controlled trials evaluating the effect of vitamin therapy,\textsuperscript{42–44} leading many to conclude that vitamin therapy is ineffective in lowering the risk of vascular events. Although many one-, two-, and three-dimensional ultrasound image-based biomarkers (i.e., IMT, TPA, and VWV, respectively) have been developed and shown to be effective in serial monitoring of various therapies,\textsuperscript{14,15,24,45} none were able detect a statistically significant effect of the B-Vitamin therapy for the subjects involved in this study. For this reason, there was an important requirement in developing more sensitive biomarkers to quantify the treatment effect. As carotid atherosclerosis is a focal disease predominantly occurring at the carotid bulbs, we reported in a previous study\textsuperscript{17} that quantifying the mean $VWT$-Change in the ROI identified by a feature selection algorithm, denoted by $\Delta VWT_s$, was more sensitive in the evaluation of treatment effect than measuring the mean $VWT$-Change over the entire carotid surface. However, the previous algorithm identified the ROI by “hard thresholding” and excluded regions that may contribute to an improved discrimination.
Medical Physics, 44 (10), October 2017

5290 Cheng et al.

Effect size (%) those obtained in the current placebo-controlled study. The effect sizes are expressed as a percentage of sample sizes tabulated below give a 90% statistical power at a significance level of 0.05 (two-tailed). The effect sizes are expressed as a percentage of those obtained in the current placebo-controlled study.

<table>
<thead>
<tr>
<th>Effect size (%)</th>
<th>ΔVWV</th>
<th>ΔVWT</th>
<th>ΔVWTWeighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>139</td>
<td>178</td>
<td>41</td>
</tr>
<tr>
<td>75</td>
<td>247</td>
<td>317</td>
<td>73</td>
</tr>
<tr>
<td>50</td>
<td>555</td>
<td>712</td>
<td>164</td>
</tr>
<tr>
<td>25</td>
<td>2220</td>
<td>2849</td>
<td>658</td>
</tr>
</tbody>
</table>

The results show that only the proposed weighted average of VWT-Change, ΔVWTWeighted, were able to identify a statistically significant difference between the B-Vitamin and placebo groups. The sample size requirement was reduced by more than 70% as compared to ΔVWV and ΔVWT.

Another important observation is that the mean annual ΔIMT was not able to detect the effect of B Vitamins, and that the mean ΔIMT measurements for both the B-Vitamin and placebo groups were approximately equal to 0 as shown in Table II. The IMT was measured for a longitudinal segment on the far wall of the CCA 5 mm away from the bifurcation. In our standardized map representation, the far wall region of CCA is located directly below the bifurcation. Figures 4(b) and 4(d) show that there was an elevation in mean ΔVWT in the right arteries for both the B-Vitamin and the treatment groups within 4 mm of the bifurcation on the CCA. Such an elevation was also exhibited in our prior placebo-controlled study evaluating the effect of atorvastatin.17 ΔIMT was not able to detect vessel-wall-and-plaque thickness burden at this region since IMT measurements were taken starting from 5 mm proximal from the CCA. Even if the IMT measurements were made closer to the bifurcation, according to the difference map shown in Fig. 4(f), the mean difference between the placebo and the B-Vitamin groups approached 0 in the CCA far wall region. A major difference between the VWT-Change between the two groups occurred at the carotid bulb pointed to by the arrow, which would not have been detected by IMT measurements, which were just taken in the CCA far wall region.

A limitation of this study is related to the inability of the current framework to analyze 15 of the 71 subjects investigated in a previous study.24 This limitation arises from the inability of the iterative closest point rigid registration technique employed in the ΔVWT map construction framework16,17,30 to align the carotid surfaces of a patient acquired at the baseline and follow-up sessions accurately due to the difference in the bifurcation angles of the two surface models. Two such examples are shown in Fig. 9. We are currently validating the performance of a nonrigid registration framework to address this limitation with preliminary results presented in Ref. [46]. Following rigid registration, this approach uses an iterative convex optimization strategy to solve for a nonrigid deformation field that minimizes the sum of the absolute intensity difference between the baseline and follow-up images and the smoothness of the deformation field as quantified by its total variation.16,47 Nonetheless, our sample size calculation shown in Table III revealed that it requires 139 and 178 subjects for ΔVWV and ΔVWT, respectively to demonstrate significance between the B-Vitamin and the placebo groups, which could not have been achieved in this study even if all 71 subjects were included in the analyses. Although ΔIMT and ΔTPA could not detect the effect of B Vitamins in the previous study,24 ΔVWV measurements were significantly higher in the placebo group (P = 0.03). For the subgroup of 56 patients involved in this study, none of the above three biomarkers detected a significant difference between the two groups. Therefore, detecting the treatment effect in the subgroup was more difficult than in the entire between the treatment and placebo groups. In addition, the hard thresholding approach in selecting ROI required the specification of the total area of the ROI (i.e., a termination condition in the feature selection algorithm), thereby introducing variability of the biomarker due to the choice of this area. As 56 feature selection trials were required in the current leave-one-out setting, ΔVWT was not sensitive to the difference between the B-Vitamin and placebo groups (P = 0.71) mainly due to the variability in the ROI areas obtained in different leave-one-out trials. The weighted averaging approach proposed in this paper has addressed the above two limitations.

Fig. 7. ROC curves of six biomarkers used in classifying B-Vitamin and placebo patients.

Fig. 8. Relationship between ΔVWV and ΔVWT.
The leave-one-out strategy employed in this study required the construction of a weighted map in the calculation of $\Delta VWT_{\text{weighted}}$ for each subject. For future studies involving another population, there is a possibility of using the weighted map obtained in this study for $\Delta VWT_{\text{weighted}}$ calculation. We do not expect there is a large difference in sensitivity if the new investigation involves diabetic patients receiving B-Vitamin treatment. However, the VWT-Change distribution pattern may be very different for another cohort with other conditions and/or receiving different treatments. Since the weighted map can be constructed within 0.1 s, it is advisable to generate the weighted map using the leave-one-out approach in future studies in order to appropriately capture the trend demonstrated in a new cohort.

Although $\Delta VWT$ can be measured for 56 subjects through rigid registration of the carotid surfaces obtained in the baseline and follow-up imaging sessions, and the resulting 3D $\Delta VWT$ map for each subject can be subsequently represented as a 2D $\Delta VWT$ map, the point correspondence relationship between different subjects implied by the 3D to 2D template mapping algorithm was not optimized in any sense. We proposed an algorithm to optimize the correspondence relationship for a cohort of subjects by minimizing the description length, and hypothesize that if mismatches of correspondence points among 2D $\Delta VWT$ maps were minimized by the algorithm, $\Delta VWT$, $\Delta VWT_S$ and $\Delta VWT_{\text{weighted}}$ would become more sensitive in detecting the effect of the B-Vitamin treatment. We will perform a study to assess this hypothesis.

A combination of lifestyle changes and medical treatments would lead to a 75–80% risk reduction for patients with high stroke risk. With these treatment options available, there is a requirement for the development of sensitive biomarkers to assess the effect of various treatment strategies and monitor the progression/regression of carotid atherosclerosis. The proposed biomarker is more sensitive than existing biomarkers and requires fewer subjects to demonstrate the statistical significance, thereby reducing the cost of clinical studies. The increase of cost-effectiveness will allow more pilot clinical studies to be performed before a larger trial is held to confirm treatment efficacy.

**ACKNOWLEDGMENTS**

Dr. Chiu is grateful for funding support from the Research Grant Council of the HKSAR, China (Project No. CityU 139713), the National Natural Science Foundation of China (Grant No. 81201149) and the City University of Hong Kong Strategic Research Grants (Nos. 7004425, 7004617).

**CONFLICT OF INTEREST**

J. David Spence is an officer of Vascularis Inc. All other authors have no conflicts.

**REFERENCES**