Three-Dimensional Carotid Ultrasound Plaque Texture Predicts Vascular Events

Arna van Engelen, PhD; Thapat Wannarong, MD; Grace Parraga, PhD; Wiro J. Niessen, PhD; Aaron Fenster, PhD; J. David Spence, MD; Marleen de Bruijne, PhD

Background and Purpose—Carotid ultrasound atherosclerosis measurements, including those of the arterial wall and plaque, provide a way to monitor patients at risk of vascular events. Our objective was to examine carotid ultrasound plaque texture measurements and the change in carotid plaque texture during 1 year in patients at risk of events and to compare these with measurements of plaque volume and other risk factors as predictors of vascular events.

Methods—We evaluated 298 patients with carotid atherosclerosis using 3-dimensional (3D) ultrasound at baseline and after 1 year and measured carotid plaque volume and 376 measures of plaque texture. Patients were followed up to 5 years (median [range], 3.12 [0.77–4.66]) for myocardial infarction, transient ischemic attack, and stroke. Sparse Cox regression was used to select the most predictive plaque texture measurements in independent training sets using a 10-fold cross-validation, repeated 5x, to ensure unbiased results.

Results—Receiver operator curves and Kaplan–Meier analysis showed that changes in texture and total plaque volume combined provided the best predictor of vascular events. In multivariate Cox regression, changes in plaque texture (median hazard ratio, 1.4; \( P<0.001 \)) and total plaque volume (median hazard ratio, 1.5 per 100 mm\(^2\); \( P<0.001 \)) were both significant predictors, whereas the Framingham risk score was not.

Conclusions—Changes in both plaque texture and volume are strongly predictive of vascular events. In high-risk patients, 3D ultrasound plaque measurements should be considered for vascular event risk prediction. (Stroke. 2014;45:2695-2701.)

Key Words: carotid arteries ■ stroke ■ ultrasound

There is an urgent need for rapid, reliable, and cost-effective methods to monitor patients who are at high risk for adverse vascular events. Such methods may be used to target treatment to high-risk patients, thereby preventing vascular events.\(^1\) Ultrasound is a relatively inexpensive and widely available imaging method enabling quantitative imaging measurements of the carotid artery wall, including intima-media thickness, vessel wall volume, and plaque burden. It has been shown that carotid plaque burden measures, such as total plaque area or total plaque volume (TPV) and their changes over time, provide strong predictors of adverse events.\(^2\)

Carotid ultrasound, by means of plaque echogenicity or texture, also provides a way to measure plaque composition. Lipid cores and intraplaque hemorrhage are thought to destabilize plaque, whereas calcifications have a stabilizing effect.\(^3,4\) In ultrasound, lipid and hemorrhagic areas are more echolucent, whereas calcified and fibrous areas are echogenic.\(^5\) Ultrasound echogenicity has been shown to differentiate between symptomatic and asymptomatic subjects\(^6\) and has been used to predict events.\(^7,8\) More complex texture measures, with examples given in Table I in the online-only Data Supplement, provide information on the distribution of pixel intensities over the plaque. Incorporating such higher order texture parameters, such as coarseness or contrast, may provide more insight into the underlying tissue properties and has been used in several studies as well.\(^9,10\) In previous studies, these higher order texture measures were shown to differentiate accurately between symptomatic and asymptomatic subjects\(^10\) and performed better than a set of plaque shape parameters.\(^11\) In addition, they were more predictive of events than a combination of a history of events and plaque parameters, such as plaque area and gray scale median.\(^12\)

In addition to single time-point measurements, progression of TPV was shown to be a strong predictor of events,\(^2\) and changes in plaque texture were more sensitive to statin-induced effects than changes in TPV.\(^13\) On the basis of all
these findings, we hypothesized that changes in plaque composition over time may be related to subsequent vascular events. Therefore, here our goal is to develop 3-dimensional ultrasound (3DUS) plaque texture measurements and to determine whether plaque texture and changes in plaque texture could improve the prediction of events in high-risk patients.

Methods
Study population, image acquisition and plaque volume annotation have previously been described.2,14

Study Population
Patients with a history of risk factors, such as hypertension or hyperlipidemia, or with a history of vascular events, who were being followed up in the Stroke Prevention Clinic or the Premature Atherosclerosis Clinic at the University Hospital, London, Canada, were enrolled in the study. The inclusion criteria included a baseline plaque area between 40 and 600 mm², measured by 2D ultrasound.15 All plaques from the clavicle to the angle of the jaw, including the right subclavian artery and the common, internal and external carotid artery were considered. Participants with a stenosis ≥70% on Doppler ultrasound were excluded, and all subjects provided written informed consent to a protocol approved by the Western University Human Research Ethics Board.

Follow-Up for Outcomes
Participants were followed up for ≤5 years (median [range], 3.12 [0.77–4.66]). At each annual visit, participants were queried about any events in the previous year. Any report of stroke, transient ischemic attack (TIA), or myocardial infarction (MI) was confirmed by the sonographer who evaluated the images was blinded with respect to events.

Ultrasound Acquisition
3DUS scans (most common voxel size, 0.21x0.21x0.21 mm; mean, 0.21x0.21x0.33 mm) of both carotid arteries were acquired at baseline and after 1 year (median [range], 364 [226–897] days). The ultrasound transducer (L12-5; 50 mm; Philips, Bothel, WA) was manually moved along the neck of the patient for 4.0 cm in >8 s, imaging ≥30 slices/s, centered around the bifurcation. Video frames were digitized using 3DEchotec equipment (General Electric Medical Systems, Hallbergmoos, Germany). The acquired 2D images were reconstructed immediately into a 3DUS image (3DQuantify; Robarts Research Institute, London, Canada). It was attempted that the same technician made the baseline and follow-up scan of each patient.

TPV Assessment
For plaque delineation, the 3DUS image was displayed using 3DQuantify. TPV was measured as described previously14 in all scans by 1 observer (T.W.) who was blinded to the time-point (baseline or follow-up), patient characteristics, and vascular events. All plaques between 1.5 cm below the bifurcation and 1.0 cm above the bifurcation into the internal carotid artery were considered. For each plaque, a plane in the longitudinal view which best visualized the plaque was selected for defining the length of the plaque and for placing 2 end points at the proximal and distal ends of the plaque. At 25%, 50%, and 75% of its length, the boundary of the plaque was annotated perpendicular to the long axis. These boundaries and the end points were connected by a computer program to create a volume. TPV per patient was measured as the sum of the volume of all plaques present. For reference, intima-media thickness was measured from longitudinal image planes extracted from the 3DUS images as previously described,2,13 and repeated 3x per image.

Texture Analysis
A set of 376 texture measures based on 9 different texture extraction techniques was calculated. Most of the texture measures have previously been used in studies on 2D10–12 and 3D13 carotid ultrasound. An overview is provided in Table 1 (details can be found in the online-only Data Supplement).

Texture was calculated for the annotated plaque volumes in both arteries. For each measure, the average was calculated by weighting the values per plaque by the plaque volumes, to obtain 1 value per patient for each measure, at both baseline and follow-up. For 7 cases with no plaque in the 3D ultrasound at baseline, baseline texture was taken as the mean of all other subjects at baseline. Texture measures were normalized by setting the mean of all subjects to 0 with a SD of 1 for each measure.

Sparse Cox regression was used to combine the 376 texture measures into 1 texture-based risk indicator, using the glmnet toolbox17,28 for R.29 With sparse regression, a penalty term promotes the reduction of the number of parameters in the model, leaving only the strongest predictors. In our experiments, we fixed the number of remaining parameters in the model to 5 because of the relatively low number of events.

To ensure unbiased parameter selection, experiments were performed by 10-fold cross-validation, so subjects were randomly divided in 10 equally sized groups with the event incidence equally divided over those 10-folds. Within cross-validation experiments the model for combining texture measures was built on 9 of the 10-folds, and used to calculate the hazard ratio (HR) for the subjects in the 10th-fold. This was repeated 10x with each fold left out once, to obtain an HR for each subject. This HR was used as the texture-based risk indicator. This was performed both for the 376 baseline texture parameters and for the 376 texture change parameters, which were calculated by subtracting baseline texture from follow-up texture.

To combine ≥2 parameters (texture/TPV baseline/changes), a Cox model was built using only those parameters. This was done using the same 10-fold cross-validation: both the model to combine the texture measures into 1 risk parameter, and the model to combine parameters was developed on 9 folds and evaluated on the 10th. All cross-validation experiments were repeated 5x, to evaluate stability of the procedure.

Statistical Analysis
Because of non-normality of the studied parameters, continuous variables are given as median and interquartile range, and Kruskal–Wallis testing was used to test for differences between groups. For categorical variables, a χ² test was used to compare groups.

The predictive value of the studied parameters was evaluated by receiver-operating characteristic (ROC) analysis, Kaplan–Meier analysis, and multivariate Cox regression. For ROC analysis, the area under the curve was determined for baseline texture, texture change, baseline TPV and TPV change, and combinations of those measures. Further analyses focused on texture change and TPV change, and

Table 1. Texture Measures Used in the Analysis

<table>
<thead>
<tr>
<th>Method</th>
<th>No. of Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray-level distribution</td>
<td>34</td>
</tr>
<tr>
<td>Gray-level co-occurrence matrix</td>
<td>78</td>
</tr>
<tr>
<td>Gray-level run length matrix</td>
<td>60</td>
</tr>
<tr>
<td>Gray-level difference matrix</td>
<td>12</td>
</tr>
<tr>
<td>Neighborhood gray-tone difference matrix</td>
<td>10</td>
</tr>
<tr>
<td>Law’s texture</td>
<td>105</td>
</tr>
<tr>
<td>Local binary pattern</td>
<td>27</td>
</tr>
<tr>
<td>Gaussian filter bank</td>
<td>24</td>
</tr>
<tr>
<td>Structure tensor</td>
<td>20</td>
</tr>
</tbody>
</table>

A more detailed description can be found in the online-only Data Supplement.
their combination. Kaplan–Meier curves were made after dividing the participants in 3 equally sized groups of increasing TPV or texture change (HR). The differences between these groups were tested for statistical significance by log-rank tests. In addition, the HR of the high-risk group with respect to the other 2 groups was determined by Cox regression.

To evaluate the potential effects of the ultrasound-derived measures in combination with other covariates, Cox regression was performed with change in TPV, change in plaque texture, and the extended Framingham risk score. The extended Framingham risk score predicts general cardiovascular disease risk, based on age, sex, high-density lipoprotein and total cholesterol, systolic blood pressure (treated or untreated), smoking, and diabetes mellitus. Because of skewness of the texture measure, the logarithm of this measure was used. The full model was compared with the model without texture change, and the model without TPV change, using a log-likelihood ratio test. In addition, Cox regression was performed using a stepwise backward Wald approach, with baseline TPV and texture added as covariates to verify their predictive value in combination with plaque changes and clinical parameters.

In all experiments, participants who experienced MI, TIA, or stroke during follow-up were considered as positive for vascular events. Statistical analysis was performed using both R (Sparse Cox regression, ROC, and Kaplan–Meier analysis) and SPSS software (full Cox regression models).

### Results

#### Patient Characteristics

In total, 298 patients were included for 3DUS analysis. During follow-up, 27 of these subjects experienced a vascular event, of which 9 had a stroke, 11 a TIA, and 7 a MI. Two subjects died as a consequence of the event. Baseline characteristics are provided in Table 2.

#### Texture Measures

In Figure 1, 2 of the strongest texture measures (when considering texture change) are shown for a single patient. These 2 measures were most often selected by Cox regression, 49 and 48 of 50 experiments, respectively. The 2 plaques in this vessel are different in appearance, which is reflected by the 2 texture images. Table I in the online-only Data Supplement provides the 6 most often selected texture measures as selected by sparse Cox regression, alongside with texture change and TPV change, and their values in patients who do or do not experience a vascular event for reference. In practice, only these strongest texture parameters need to be calculated, on which the risk parameter can be determined and compared with reference values.

### Table 2. Baseline Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Without Event (n=271)*</th>
<th>With Event (n=27)*</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, y</strong></td>
<td>70 (64–77)</td>
<td>75 (64–80)</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Men, %</strong></td>
<td>58</td>
<td>52</td>
<td>0.54</td>
</tr>
<tr>
<td><strong>Systolic blood pressure, mmHg</strong></td>
<td>134 (122–149)</td>
<td>131 (114–124)</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>Diastolic blood pressure, mmHg</strong></td>
<td>74 (66–82)</td>
<td>69 (62–78)</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>IMT, mm</strong></td>
<td>0.92 (0.82–1.04)</td>
<td>0.93 (0.84–0.98)</td>
<td>0.74</td>
</tr>
<tr>
<td><strong>TPV, mm³</strong></td>
<td>273 (191–437)</td>
<td>253 (119–422)</td>
<td>0.21</td>
</tr>
<tr>
<td><strong>Stenosis, %</strong></td>
<td>40 (40–50)</td>
<td>40 (40–40)</td>
<td>0.26</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td>27.9 (25.5–31.5)</td>
<td>29.0 (25.3–31.8)</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>Smoking (never, quit, and still smoking)</strong></td>
<td>36%, 56%, and 8%</td>
<td>33%, 48%, and 19%</td>
<td>0.74, 0.45, and 0.06</td>
</tr>
<tr>
<td><strong>Smoking pack-years</strong></td>
<td>5 (0–24)</td>
<td>15 (0–28)</td>
<td>0.34</td>
</tr>
<tr>
<td><strong>Total cholesterol, mmol/L</strong></td>
<td>3.9 (3.4–4.7)</td>
<td>4.1 (3.5–4.9)</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>HDL cholesterol, mmol/L</strong></td>
<td>1.3 (1.1–1.7)</td>
<td>1.3 (1.1–1.8)</td>
<td>0.85</td>
</tr>
<tr>
<td><strong>LDL cholesterol, mmol/L</strong></td>
<td>1.9 (1.5–2.5)</td>
<td>2.1 (1.5–2.6)</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Triglycerides, mmol/L</strong></td>
<td>1.2 (0.9–1.7)</td>
<td>1.1 (0.9–1.8)</td>
<td>0.77</td>
</tr>
<tr>
<td><strong>Framingham risk score</strong></td>
<td>16 (13–18)</td>
<td>15 (13–19)</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>Diabetes mellitus, %</strong></td>
<td>21</td>
<td>19</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Previous stroke, %</strong></td>
<td>22</td>
<td>33</td>
<td>0.19</td>
</tr>
<tr>
<td><strong>Previous TIA, %</strong></td>
<td>42</td>
<td>56</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Previous MI, %</strong></td>
<td>14</td>
<td>26</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>Previous atrial fibrillation, %</strong></td>
<td>9</td>
<td>15</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>Previous endarterectomy, %</strong></td>
<td>7</td>
<td>4</td>
<td>0.48</td>
</tr>
<tr>
<td><strong>Previous carotid angioplasty, %</strong></td>
<td>1</td>
<td>4</td>
<td>0.26</td>
</tr>
<tr>
<td><strong>Previous peripheral artery angioplasty, %</strong></td>
<td>1</td>
<td>4</td>
<td>0.26</td>
</tr>
<tr>
<td><strong>Previous coronary angioplasty, %</strong></td>
<td>7</td>
<td>11</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>CABG, %</strong></td>
<td>10</td>
<td>15</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>Taking antihypertensive drug, %</strong></td>
<td>89</td>
<td>93</td>
<td>0.52</td>
</tr>
</tbody>
</table>

*Continuous variables are given as median (interquartile range); categorical variables are given as percentages.

†P values using Kruskal–Wallis testing for continuous variables and χ² testing for categorical variables.
ROC Analysis

In Figure 2A, ROC curves are shown for baseline TPV and texture change in intima-media thickness, TPV, and texture. Figure 2B shows ROC curves when ≥2 features are combined.

Kaplan–Meier Analysis

Figure 3 shows Kaplan–Meier curves for TPV change, texture change, and their combination. Threshold values of the HRs for separating the medium-risk tertile from the low-risk and high-risk tertiles were −21 and 59 mm³ for TPV change, 0.9 and 1.1 for texture change, and 0.7 and 1.3 for their combination. The difference between risk tertiles was significant for all experiments (P<0.039 for TPV change; P≤0.01 for texture change; and P<0.001 for the combination).

Comparing the high-risk tertile with the combined low-risk and medium-risk tertiles with Cox regression, showed a HR of 2.3 (P<0.03) for TPV change, 4.3 (2.7–6.4) with P<0.001 to 0.01 for texture change, and 6.2 (4.2–7.9) with P<0.001 for the combination.

Cox Regression

Table 3 shows the results for Cox regression with TPV change, texture change, and the Framingham risk score. Both TPV change and texture change were significant predictors of stroke, TIA, and MI, whereas in this data set the Framingham risk score was not. The log-likelihood ratio tests indicated that the full model was significantly better than the model excluding texture change and the model excluding TPV change (all P<0.01).

Table II in the online-only Data Supplement shows 3 additional subanalyses. When only stroke and TIA are considered as events, TPV change (HR, 1.5; P≤0.007) and texture change (HR, 1.4; P<0.003) remain the only significant parameters in the model. Moreover, after exclusion of 7 patients without plaque at the 3DUS baseline image (TPV change: HR, 1.5; P≤0.001 and texture change: HR, 1.6; P<0.001) and after exclusion of 28 patients with atrial fibrillation at baseline (TPV change: HR, 1.7; P<0.001 and texture change: HR, 1.5; P<0.003), results were similar as well.
For backward Wald analysis, including baseline TPV, baseline texture, TPV change, texture change, and the Framingham risk score in all 5 experiments, TPV change ($P \leq 0.001$; HR, 1.4–1.5) and texture change ($P \leq 0.003$; HR, 1.3–1.5) were the only 2 remaining parameters.

**Discussion**

This study compared the predictive value for vascular events of texture and TPV derived from 3D carotid ultrasound. Changes in 3DUS texture characteristics and TPV over time were stronger predictors than baseline variables. Both texture and TPV change remained significant predictors after adjustment for the Framingham risk score. The full model was significantly better than the model including only TPV or texture change and the Framingham risk score.

The predictive value of TPV change over time\(^2\) and baseline texture\(^1\) have been evaluated previously, but to our knowledge this is the first study to consider changes in texture and to combine texture change and TPV change for risk stratification of patients over time. Texture change was more predictive of events than baseline texture, which suggests that plaques that are changing faster impose a higher risk on patients than plaques that are stable in texture and by analogy, stable in composition. This notion is also supported by the observation that in our analysis, for all texture measures that were selected, the median change is closer to zero for patients not experiencing an event than for those who do (Table I in the online-only Data Supplement). Moreover, this is analogous to recent findings that plaque changes, such as TPV progression\(^2\) or fastest intima-media thickness progression over several vessel segments,\(^3\) are predictive of events.

Framingham risk score was not predictive of events in this data set. This could be explained by the inclusion that was based on the presence of carotid plaque and risk factors or symptoms. All included patients had an increased Framingham risk score (interquartile range, 13–18). Here, we show that in a population with increased risk, plaque measurements using ultrasound can further stratify patients.

In a previous study where statin-induced changes were better reflected by texture change than by TPV change,\(^13\) Law’s texture measures performed best, which is similar to our findings. Moreover, neighborhood gray-tone difference matrix coarseness was important in our study and was among the best measures in previous work that showed the discrimination between symptomatic and asymptomatic patients.\(^11\)

Our findings can be used both for patient monitoring and evaluation of therapies. The present study shows that including plaque texture in the monitoring of patients contributes to improved risk assessment using ultrasound. A yearly follow-up, including ultrasound in an atherosclerosis clinic, is practically feasible.\(^32\) More regular follow-up of high-risk patients would be possible to enable adjustment of therapy in a more timely fashion. Adjusting patient treatment based on changes in plaque area instead of traditional risk factors was already shown to reduce the number of vascular events significantly.\(^33\)

Including changes in TPV and texture could further improve effectiveness of patient management. In addition, cost-effective measurements are needed to evaluate newly developed
Table 3. Results for Cox Regression

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Individual Model</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P Value</td>
<td>HR</td>
</tr>
<tr>
<td>Texture change (per 0.1 change in log [HR])</td>
<td>&lt;0.001</td>
<td>1.4 (1.3–1.5)</td>
</tr>
<tr>
<td>TPV change (per 100 mm³)</td>
<td>&lt;0.001</td>
<td>1.5</td>
</tr>
<tr>
<td>Framingham risk score</td>
<td>0.31</td>
<td>1.1</td>
</tr>
</tbody>
</table>

For texture change and for the full model, the HRs and P values are given with median and range of the 5 repetitions of calculating the texture-based risk indicator. HR indicates hazard ratio.

Conclusions

Changes in ultrasound plaque texture and volume are predictors of vascular events in patients in whom traditional risk factors measured by the Framingham risk score are not. These measures can be used in clinical practice as a cost-effective way to monitor high-risk subjects or as an evaluation measure in the development of new therapies.

Acknowledgments

Drs A.A. Hackam and D.G. House were coinvestigators, in the study in which the images and data on risk factors and outcomes were obtained.2 Dr Spence was the Principal Investigator of that study.

Sources of Funding

This study was funded by (1) CTMM, the Center for Translational Molecular Medicine (www.ctmm.nl), project PARisk (grant 01C-202), and the Dutch Heart Foundation and (2) the Heart & Stroke Foundation of Canada (Ontario), grant number NA5912. Arna van Engelen received additional funding from the Erasmus trust fund.

Disclosures

Dr Spence is a principal of Vascularis Inc and has received research grants from the Canadian Institutes of Health Research and the Heart & Stroke Foundation of Canada (Ontario). Dr Niessen is cofounder and has stocks in Quantib B.V. The other authors report no conflicts.

References

Three-Dimensional Carotid Ultrasound Plaque Texture Predicts Vascular Events
Arna van Engelen, Thapat Wannarong, Grace Parraga, Wiro J. Niessen, Aaron Fenster, J. David Spence and Marleen de Bruijne

Stroke. 2014;45:2695-2701; originally published online July 17, 2014;
doi: 10.1161/STROKEAHA.114.005752

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/45/9/2695

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2014/07/15/STROKEAHA.114.005752.DC1
Three-dimensional Carotid Ultrasound Plaque Texture Predicts Vascular Events

Arna van Engelen¹, Thapat Wannarong²,³, Grace Parraga⁴, Wiro J. Niessen¹,⁵, Aaron Fenster⁴, J. David Spence²,⁴ and Marleen de Bruijne¹,⁶

1. Biomedical Imaging Group Rotterdam, Departments of Medical Informatics and Radiology, Erasmus MC, the Netherlands
2. Stroke Prevention & Atherosclerosis Research Centre, Robarts Research Institute, Western University, London, Canada
3. Department of Internal Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand
4. Imaging Research Laboratories, Robarts Research Institute, Western University, London, Canada
5. Department of Imaging Science and Technology, Faculty of Applied Sciences, Delft University of Technology, the Netherlands
6. Department of Computer Science, University of Copenhagen, Denmark

1. Supplemental methods: texture descriptors

Here we provide a more detailed description of the texture measures that have been included in the study. Most of these texture measures have previously been used in studies on 2D¹-³ and 3D⁴ carotid US. Our parameter settings were based on these studies, but texture measures are calculated for all three imaging planes, or for the complete plaque volumes. The texture measures were implemented in Matlab (Matlab R2011B, The MathWorks Inc., Natick, MA).

For each group of measures we summarize the settings that were used and provide a complete list of these measures. Before calculating texture measures, the images were normalized. For each image the 10th percentile of the imaged volume was fixed at 10 and the 90th percentile at 150, and all intensities were linearly scaled with these values. Texture was calculated for all 3D plaque regions in both carotid arteries, and the texture of the left and right artery were averaged by taking the weighted average with respect to plaque size, to obtain one value per patient for each measure.

Gray level distribution (GLD, 34 measures)
GLD measures were calculated from all voxels within the plaque regions. Our selection of measures was based on Awad et al⁴. The selected measures were mean, standard deviation, median, minimum, maximum, entropy, mode, energy, the first 7 standardized moments around the mean, a normalized histogram of 20 bins ranging from 0-500, and the histogram bin with the highest count.

Gray-level co-occurrence matrix (GLCM, 78)
GLCM measures as introduced by Haralick et al⁵ measure the joint probability of pairwise combinations of image gray levels. In this paper co-occurrences with a distance of 1 pixel were calculated in horizontal, vertical, and diagonal orientations (θ = 0, 45, 90, 135) for intensity bins with width 10. For the axial, coronal and sagittal
plane one co-occurrence matrix was made that included all slices, and the mean and standard deviation of the 4 orientations were calculated per plane for 13 measures\textsuperscript{5,6}: autocorrelation, contrast, correlation, cluster prominence, dissimilarity, energy, entropy, homogeneity, maximum probability, sum average, sum entropy and two information measures of correlation.

**Gray level run length (GLRL, 66)**
GLRL considers ‘runs’, which are defined as a set of consecutive pixels with the same gray level value\textsuperscript{7} and of which the length is the number of pixels in that run. For the present paper, runs with $\theta = 0, 45, 90$ and 135 were calculated in the axial, coronal and sagittal plane. For each plane the following measures were calculated and averaged over the four orientations: Short run emphasis, long run emphasis, gray level non-uniformity, run length non-uniformity, run percentage, low gray level run emphasis, high gray level run emphasis, short run low gray level emphasis, short run high gray level emphasis, long run low gray level emphasis and long run high gray level emphasis. All measures were calculated for both intensity bins of width 5 and 20 and implemented using the toolbox by Wei\textsuperscript{8}.

**Gray level difference matrix (GLDM, 12)**
GLDM\textsuperscript{9} calculates the absolute difference between neighboring pixels. We used a distance of 1 pixel, and calculated the mean over 4 orientations ($\theta = 0, 45, 90, 135$), for 3 planes (axial, coronal, sagittal). For each plane the mean, contrast, angular second moment and entropy were calculated.

**Neighborhood gray tone difference matrix (NGTDM, 10)**
NGTDM was developed to calculate five measures related to human perception: coarseness, contrast, busyness, complexity and strength\textsuperscript{10}. We calculated the mean difference with the surrounding voxels in a 3D 26-neighborhood, for intensity bins of 1 and 10, and calculated the five measures for both these settings.

**Laws texture (105)**
Laws developed texture energy measures by convolution of the image with 1D kernels\textsuperscript{11}. We combined five 1D kernels to study 3D texture: $[1 4 6 4 1]$ (Level, L), [-1 -2 0 2 1] (Edge, E), [-1 0 2 0 -1] (Spot, S), $[1 -4 6 -4 1]$ (Ripple, R) and $[-1 2 0 -2 1]$ (Wave, W), which created 125 3D kernels: LLL, LLE, etc. After convolution, we calculated the mean, absolute mean and standard deviation over the region of interest, and averaged measures of rotated kernels, such as LLE, LEL and ELL, or LES, LSE, ELS, ESL, SEL and SLE. This gave 35 kernels with three measures each.

**Local Binary Pattern (LBP, 27)**
The original 2D implementation of LBP measures the homogeneity of texture by determining the number of transitions from intensities higher than each central pixel to intensities lower than that central pixel\textsuperscript{12,13}. We applied a 3D adaptation where we determined the number of ‘regions’ higher or lower in intensity around each voxel in a 26- and a 98-neighborhood. Measures were the mean and standard deviation of the number of regions in both the 26- and the 98-voxel neighborhood, a normalized histogram of 1 up to 6 and > 6 areas for the 26-neighborhood, and 1 up to 15 and > 15 areas for the 98-neighborhood.

**Gaussian filter bank (24)**
We applied 3D Gaussian filters to calculate blurred intensity, gradient magnitude, Laplacian and curvature at 3 scales: 0.16mm, 0.32 mm and 0.64mm. As measures we used the mean and standard deviation of these measures over the regions of interest\textsuperscript{14}.

**Structure tensor (20)**
The structure tensor\textsuperscript{15} measures coherence in an image. From the 3D structure tensor we calculated the fractional anisotropy, the 3 eigenvalues and the determinant, and calculated the mean and standard deviation over the regions of interest\textsuperscript{16}. We applied two different scales: an inner scale of 0.16 mm with an outer scale of 0.48 mm, and an inner scale of 0.32 mm with an outer scale of 0.80 mm.
2. Supplemental results

Here we show some additional results. Table I presents the strongest Texture change measures, together with TPV change and the combined Texture change risk indicator, for the patients with and without events, and the statistical difference between the two groups. These parameters are determined to be the strongest since these were most often selected in sparse Cox regression. Out of the five repetitions of 10-fold cross-validation (50 times performing sparse Cox regression), these were selected 49 (Laws EER), 48, (NGTDM Coarseness and Laws SSR), 41 (Laws SSS) and 15 (NGTDM Contrast, and GLCM standard deviation of correlation in axial plane) times. This means that for use in practice a Texture risk parameter based on these texture measures can be determined, such that for new patients only those few measures need to be calculated, combined and compared with the reference value.

Table II presents three different subanalyses for Cox regression: 1) with only stroke and TIA as event, so considering subjects with myocardial infarction as no event, 2) excluding patients without plaque on 3DUS at baseline, and 3) excluding patients with atrial fibrillation at baseline.

Table I: Most relevant (texture change) measures and their difference between patients with and without events

<table>
<thead>
<tr>
<th>Measure Description</th>
<th>Without event (n=271) Median (IQR)</th>
<th>With event (n=27) Median (IQR)</th>
<th>p-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPV change (mm^3)</td>
<td>21 (-53 – 77)</td>
<td>80 (29 – 170)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Texture change: mean of Laws EER</td>
<td>-0.13 (-0.61 – 0.41)</td>
<td>0.50 (0.08 – 1.26)</td>
<td>0.0005</td>
</tr>
<tr>
<td>Texture change: NGTDM Coarseness (bins of 1)</td>
<td>-0.06 (-0.63 – 0.53)</td>
<td>0.49 (0.02 – 1.20)</td>
<td>0.0013</td>
</tr>
<tr>
<td>Texture change: mean of Laws SSR</td>
<td>0.11 (-0.31 – 0.42)</td>
<td>-0.35 (-0.85 – 0.05)</td>
<td>0.0004</td>
</tr>
<tr>
<td>Texture change: mean of Laws SSS</td>
<td>0.00 (-0.38 – 0.47)</td>
<td>-0.53 (-0.96 – 0.06)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Texture change: NGTDM Contrast (bins of 1)</td>
<td>0.02 (-0.27 – 0.29)</td>
<td>-0.32 (-1.33 – 0.15)</td>
<td>0.0061</td>
</tr>
<tr>
<td>Texture change: GLCM in axial plane, standard deviation of correlation (bins of 10)</td>
<td>0.13 (-0.53 – 0.58)</td>
<td>-0.56 (-1.08 – 0.04)</td>
<td>0.0014</td>
</tr>
<tr>
<td>Texture change: risk parameter†</td>
<td>0.97 (0.82-1.14)</td>
<td>1.24 (1.05-1.46)</td>
<td>0.00005</td>
</tr>
</tbody>
</table>

* Kruskal-Wallis test
† HR after combining 5 measures of Texture change. The given result is the median over the five experiments with different cross-validation folds.
Table II: Three subanalyses for Cox regression. *The hazard ratio and p-value for all parameters are given with median and range over 5 repetitions of calculating the Texture change Risk Indicator.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Hazard Ratio*</th>
<th>Hazard Ratio*</th>
<th>Hazard Ratio*</th>
<th>Hazard Ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texture change (per 0.1 change in log(HR))</td>
<td>1.5 (1.3-1.5)</td>
<td>1.7 (1.6-1.7)</td>
<td>1.1 (1.0-1.2)</td>
<td>1.1 (1.0-1.2)</td>
</tr>
<tr>
<td>TPV change (per 100 mm³)</td>
<td>≤0.003</td>
<td>≤0.001</td>
<td>≤0.001</td>
<td>≤0.001</td>
</tr>
<tr>
<td>Framingham risk score</td>
<td>0.87 (0.83-0.94)</td>
<td>0.87 (0.83-0.94)</td>
<td>0.87 (0.83-0.94)</td>
<td>0.87 (0.83-0.94)</td>
</tr>
</tbody>
</table>

Excluding patients with atrial fibrillation at baseline (270 patients, 23 events)
- Hazard Ratio*: 1.5 (1.3-1.5)
- p-value*: ≤0.003

Excluding patients without plaque on 3DUS at baseline (291 patients, 26 events)
- Hazard Ratio*: 1.5 (1.3-1.7)
- p-value*: ≤0.001

Excluding myocardial infarction as positive end point for events (298 patients, 20 events)
- Hazard Ratio*: 1.4 (1.3-1.5)
- p-value*: ≤0.003

Excluding myocardial infarction as positive end point for events (298 patients, 20 events)
- Hazard Ratio*: 1.5 (1.4-1.5)
- p-value*: ≤0.007

Excluding patients with atrial fibrillation at baseline (270 patients, 23 events)
- Hazard Ratio*: 1.7 (1.6-1.7)
- p-value*: ≤0.001

Excluding patients without plaque on 3DUS at baseline (291 patients, 26 events)
- Hazard Ratio*: All 1.6
- p-value*: ≤0.001

Excluding myocardial infarction as positive end point for events (298 patients, 20 events)
- Hazard Ratio*: All 1.1
- p-value*: ≤0.001

Excluding myocardial infarction as positive end point for events (298 patients, 20 events)
- Hazard Ratio*: All 1.1
- p-value*: ≤0.001

Excluding myocardial infarction as positive end point for events (298 patients, 20 events)
- Hazard Ratio*: All 1.1
- p-value*: ≤0.001
3. Supplemental References


