NEURO

Annet Waaijer
M. Weber
M. S. van Leeuwen
J. Kardux
W. B. Veldhuis
R. Lo
F. J. A. Beek
M. Prokop

Grading of carotid artery stenosis with multidetector-row CT angiography: visual estimation or caliper measurements?

A. Waaijer (✉) · M. S. van Leeuwen ·
W. B. Veldhuis · R. Lo · F. J. A. Beek ·
M. Prokop
University Medical Centre Utrecht,
Heidelberglaan 100,
3583 CX Utrecht, The Netherlands
e-mail: annetwaaijer@gmail.com
Tel.: +31-88-7556687
Fax: +31-88-2581098

M. Weber
Medical University Wien,
Dr. Karl Luegerring 1,
1010 Innerestad Wien, Austria

J. Kardux
Gelre Ziekenhuis Apeldoorn,
Albert Schweizerlaan 31,
7334 JZ Apeldoorn, The Netherlands

Abstract To assess the optimal method for grading carotid artery stenosis with computed tomographic angiography (CTA), we compared visual estimation to caliper measurements, and determined inter-observer variability and agreement relative to digital subtraction angiography (DSA). We included 46 patients with symptomatic carotid stenosis for whom CTA and DSA of 55 carotids was available. Stenosis quantification by CTA using visual estimation (CTA_{VE}) (method 1) was compared with caliper measurements using subjectively optimized wide window settings (method 2) or predefined contrast-dependent narrow window settings (method 3). Measurements were independently performed by two radiologists and two residents. To determine accuracy and inter-observer variability, we calculated linear weighted kappa, performed a Bland-Altman analysis and calculated mean difference (bias) and standard deviation of differences (SDD). For inter-observer variability, kappa analysis was "very good" (0.85) for expert observers using CTA_{VE} compared with "good" (0.61) for experts using DSA. Compared with DSA, method 1 led to overestimation (bias 5.8–8.0%, SDD 10.6–14.4), method 3 led to underestimation (bias −6.3 to −3.0%, SDD 13.0–18.1). Measurement variability between DSA and visual estimation on CTA (SDD 11.5) is close to the inter-observer variability of repeated measurements on DSA that we found in this study (SDD 11.6). For CTA of carotids, stenosis grading based on visual estimation provides better agreement to grading by DSA compared with stenosis grading based on caliper measurements.

Keywords CT angiography · Carotid stenosis · Measurement · Digital subtraction angiography · Visual estimation

Introduction

Recent investigations have shown that the incidence of stroke and transient ischemic attacks is exceeding the incidence of coronary heart disease [1]. Both carotid endarterectomy and carotid artery stenting are performed to reduce the risk of recurrent stroke and transient ischemic attack [2, 3]. Accurate measurements of the degree of stenosis are important because higher grades of carotid artery stenosis are associated with an increased risk of stroke and because the degree of stenosis together with stenosis-related symptoms will determine how much a patient might profit from carotid endarterectomy or stent placement [3, 4]. Based on the results of the North American Symptomatic Endarterectomy Trial (NASCET), only discrimination between 50–69% and 70–99% stenosis was considered to be important. However, more recent studies made use of different cut-off values for patient selection, using stenosis degree of 50%, 60%, 70% and 80%, dependent on the presence of symptoms and co-morbidity [2, 4–6].

While stenosis grading was primarily based on intra-arterial digital subtraction angiography (DSA), this

technique has gradually been replaced by less invasive techniques such as duplex ultrasound, magnetic resonance angiography (MRA) and computed tomographic angiography (CTA). With the introduction of multidetector-row scanning, CTA has become faster, easier to use and has further gained in spatial resolution [7]. In addition, CTA of the carotid and intra-cerebral arteries has been advocated as part of the work-up of patients with acute stroke [8].

Stenosis grading on DSA is most commonly based on the NASCET criteria [9]. Stenosis grading MRA or CTA can be based on visual estimation or caliper measurements either on the source data or on images that are processed to resemble angiographic projections [10–13]. Despite the good overview that these processed images can provide, axial images have been shown to be most accurate for CTA measurements [14–17].

Compared with DSA, cross-sectional techniques in general suffer from a reduced spatial resolution. This is especially evident when trying to precisely define the borders of a vessel, in particular if this vessel is small or stenosed. Placing calipers for digital measurements may thus become less precise than usually assumed. In an effort to avoid these blurred vessel edges and increase precision, contrast-dependent window settings have been suggested for CTA [18, 19]. However, any caliper measurement of a vessel stenosis in vivo relies on a number of subjective factors, such as precise position of the calipers at the vessel borders as well as the decision where to measure the most stenotic segment and where to locate the reference segment.

Visual estimation of the vessel stenosis, on the other hand, implies one subjective decision that integrates all available imaging information. It is not clear which technique for stenosis grading is most accurate and reproducible. Because of the potential limitations of caliper measurements, we hypothesize that visual estimation may actually perform better than assumed.

In this study we compared visual estimation with caliper measurements in patients with carotid artery disease to assess the optimum method for grading of carotid artery stenosis with CTA, and determined inter-observer variability as well as accuracy of the various grading technique relative to DSA as the reference standard.

Materials and methods

Study population

Between September 2003 and August 2005, 48 consecutive patients were included in this prospective diagnostic study. All patients had a more than 50% symptomatic carotid artery stenosis diagnosed by ultrasound in combination with MRA, and had been scheduled for DSA and stent placement because of participation in the International Carotid Stenting Study (ICSS) (www.cavatas.com, registered under number ISRCTN25337470). In our hospital, the CTA was added to the protocol after permission of the medical ethics committee. Carotid multidetector-row CTA was performed within 2 weeks of DSA. We included all carotid arteries for further evaluation in which there were at least two DSA projections available for comparison. The ICSS trial and the CTA study were performed with institutional review board approval. All patients had given written informed consent before entering the study.

One patient was excluded because of renal failure and one patient had to be excluded because of a technical failure of contrast injection during CTA. In the remaining 46 patients, correlation of CTA with at least two diagnostic DSA series was available for 55 carotid arteries.

The group of patients included 29 males and 17 females with a mean age of 68 years (range, 44–84 years). Symptoms were stroke in 23 patients, transient ischemic attack in 20 patients and amaurosis fugax in 14 patients. Eleven patients had two symptoms.

CTA

CTA was performed using a 16-detector-row CT system (MX 8000 IDT, Philips Medical Systems, Cleveland, Ohio) on 37 patients, a 40-detector-row CT system on eight patients and a 64-detector-row CT system on three patients (Brilliance-40 and Brilliance-64, Philips Medical Systems, Cleveland, Ohio). Patients underwent CT supine with the head tilted so that the mandible was perpendicular to the table in order to minimize dental artefacts. The volume covered started 3 cm under the vertex and ended just below the aortic arch, to avoid artefacts from the subclavian vein. We used 16×0.75-mm or 40×0.625-mm or 64×0.625-mm collimation with a pitch between 0.77 and 0.85 (dependent on the CT system options) and a rotation time of 0.42 s. In order to keep the differences between CT systems as small as possible, we reconstructed overlapping sections of 1.0-mm slice thickness (16 detector-rows) or 0.9-mm thickness (40 and 64 detector-rows) at a reconstruction interval of 0.5 mm and a field of view of 160 mm. The resulting pixel size on axial images was 0.32 mm. A moderately smoothing filter was applied (filter B) on all CT systems. We employed 120 kVp tube voltage and 180 mAs (effective) with all CT system scanners.

CT was performed after intravenous injection of 80 ml of contrast material (Ultravist 300, Schering, Berlin, Germany) at 5 ml/s followed by a saline chaser bolus of 50 ml injected at the same flow rate. The delay before CT data acquisition was determined from a test bolus that consisted of an injection of 40 ml contrast material that was used for a brain perfusion study.

The resulting images were sent to a dedicated CT workstation (Extended Brilliance Workspace, software version 2.2, Phillips Medical Systems, Cleveland, Ohio) for further evaluation.

DSA

Intra-arterial DSA was performed on a Philips Integris V3000 angiographic unit (Philips Medical Systems, Best, The Netherlands). A 5-F catheter was introduced using the Seldinger technique (femoral arteriotomy) and was selectively positioned in one or both common carotid arteries. At least two projections (postero-anterior and lateral) were acquired from each carotid artery. For patients undergoing stent placement, frequently only the symptomatic side was selectively catheterized to keep intervention times low. For each projection, 6 ml of contrast material (Ultravist 300, Schering, Berlin, Germany) was injected at a flow rate of 3 ml/s. An image intensifier size of 16 cm was used with a matrix size of $1{,}024 \times 1{,}024$. The resulting pixel size at DSA was 0.16 mm.

Grading of carotid stenosis

We included readers with a wide spectrum of experience to be able to estimate the influence of knowledge and training on measurement results. Therefore, two radiologists with extensive experience in reading CTA studies (observer 1 >15 years; observer 2 >10 years), and two radiology residents (observer 3 at the end of his residency period and observer 4 at the beginning of his residency period) were asked to determine the degree of carotid artery stenosis on CTA using three different methods (Fig. 1). Every observer was

blinded to clinical information. All evaluations were performed interactively on the CT workstation and were based on the NASCET criteria [9]: degree of stenosis = (1 − minimal residual lumen diameter/distal lumen diameter) × 100%.

In case the observers did not consider it possible to grade the stenosis using a specific grading technique, they were asked to state the reason for it. For each method we determined the number of carotids in which grading was not considered possible.

Method 1 was visual estimation (CTA_{VE}) of the degree of stenosis based on interactive reformations perpendicular and parallel to the internal carotid artery. Visual estimation was performed interactively on the CT workstation. This workstation allows for interactively scrolling through the data set in a cine-like display and adapting the view plane interactively so as to be perpendicular to the vessel axis or parallel to the carotid bifurcation. The observers were asked to adapt the window setting so that calcifications in the region of stenosis would not be rendered completely 'white' and the background soft tissue would not be rendered completely 'black', starting with a wide window typically at a width/level of 600/150. They were free to interactively adapt the sectional plane using multiplanar reformatting to optimally display the region of maximum stenosis as well as a distal reference segment in a longitudinal fashion.

They were asked to assign the degree of luminal stenosis to one of the 11 following categories: 0–9%, 10–19%, ...,
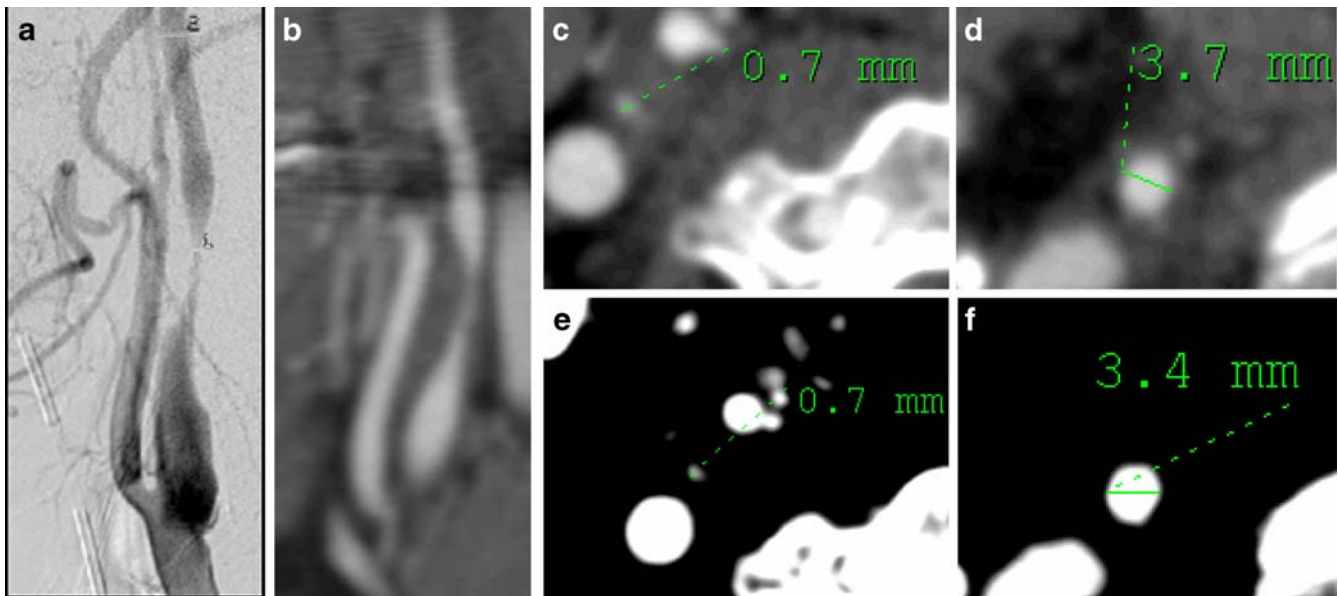
Fig. 1 Comparison of DSA (**a**) and CTA (**b**–**f**) of a high-grade carotid artery stenosis. The luminal diameter in the region of the stenosis (**c, e**) was related to a distal reference diameter (**d, f**). This was done using individually adapted wide window settings (here: width/level=600/150) and using the predefined contrast-dependent window settings (here: width/level=98/163). On DSA the two observers measured a 82 and 88% stenosis, respectively. Visual estimation on CTA yielded a degree of stenosis between 75 and 85%. Caliper measurements using a wide window setting resulted in 58–83% stenosis while caliper measurements with the predefined window setting resulted in 69–81% stenosis

90–99%, 100% (= occlusion). For further analysis of visual estimation data, these categories were substituted by the middle 5% of each range, thus 5%, 15%, 25%, 35%, 45%, 55%, 65%, 75%, 85%, 95% and 100%.

The other two methods were based on measurements using the caliper function of the workstation on transverse cuts perpendicular on the central lumen line, resulting in MPR images. This central lumen line was semi-automatically acquired using the "Advanced Vessel Analysis" program (Phillips Medical Systems, Cleveland, Ohio) on the CT workstation. Measurements were only performed on the CT workstation since PACS does not offer such techniques. Care was taken to manually adapt the central lumen line whenever necessary so that it avoided crossing calcified plaques. For determining the location of the stenotic segment and the distal reference segment for further measurements according to the NASCET criteria, the observers used interactive window settings as described for method 1. Each observer could freely define the level of the axial cut plane he deemed most appropriate for measuring the stenosis and the distal reference segment. The distal reference segment was always chosen such that the carotid walls were parallel and beyond the tapering after the carotid bulb. All observers used equal magnification (4-cm display field of view on a 4-on-1 monitor subdivision).

Method 2 was based on caliper measurements using a wide window setting (CTA_{WW}) as described for method 1. This window setting could be individually adapted by the observer if deemed appropriate.

Method 3 was based on the predefined window settings (CTA_{PW}) proposed by Liu et al. [18]. Using phantom experiments in which they varied luminal contrast, degree of stenosis and window/level settings, they suggested optimum window/level settings based on the luminal contrast enhancement in the carotid arteries. In their article eight categories of arterial enhancement were selected and appropriate narrow window settings are given for each category. Therefore, we determined the individual CT numbers by drawing an ROI of 10 mm^2 in the centre of the common carotid artery (CCA) just proximal to the carotid bifurcation and determined from Liu and co-workers' publication, which preset window settings to use for measuring the diameter of the carotid artery at the level of the stenosis and in the distal reference segment.

Two experienced interventional radiologists (observers A and B) performed all diameter measurements on DSA images using calipers. They could choose image enlargement at the viewing workstation according their personal preference. They chose the projection with maximum stenosis and determined the degree of stenosis using the NASCET criteria [9]. If the degree of stenosis could not be measured due to near-occlusion, i.e. severe collapse of the distal lumen, the observers assigned a 95% degree of stenosis. Next, a consensus reading was performed to serve as standard of reference for comparison with CTA measurements.

Statistical analysis

The consensus reading on DSA served as a reference standard for CTA. If the degree of stenosis could not be measured due to severe distal collapse, due to the presence of calcifications or due to the presence of high-grade stenosis resulting in reduction of intra-luminal contrast enhancement, these measurements were excluded from the analysis of caliper measurements on CTA (methods 2 and 3).

We assessed inter-observer variability for the two experienced interventional radiologists for DSA as well as for the two most experienced (observers 1 and 2) and for the two less experienced observers (observers 3 and 4) for CTA using linear weighted kappa statistics (http://faculty.vassar.edu/lowry/kappa.html). Therefore, all stenosis gradings were placed into four groups: group 1, 0–49% stenosis; group 2, 50–69% stenosis; group 3, 70–99% stenosis; group 4, 100% stenosis. A kappa value <0.20 is considered poor, fair for 0.21–0.40, moderate for 0.41–0.60, good for 0.61–0.80 and very good for 0.81–1.00. Second, we applied Bland-Altman analysis to show the absolute differences between repeated measurements. Therefore, we determined the mean difference between pairs of repeated stenosis gradings to assess mean difference (bias) and calculated the standard deviation of the differences (SDD) to assess variability. Upper and lower limits of agreement (LoA; 95% confidence interval for two repeated measurements) were calculated from the mean difference ±1.96 × SDD.

For determining the agreement of the various CTA measures for each observer separately relative to the grading of carotid stenoses with DSA, we also used linear weighted kappas using the consensus reading as standard of reference grouping data, as we did for inter-observer variability.

We also performed a normalized Bland-Altman analysis in which the consensus DSA reading served as the standard of reference. [20]. We calculated the bias and SDD for each method and each observer relative to DSA measurements. For each observer separately, significance of differences between the SDDs of three methods was tested with Levene's test for homogeneity of variances (Statistical Software Package SPSS for Windows, version 12).

Results

Standard of reference

In all 48 patients, stenosis ≥50% in the internal carotid artery was found. Stenosis grading with DSA revealed a mean degree of stenosis of 76% (range, 46–100%) on the symptomatic side ($n=46$) and 34% (range, 0–82%) on the asymptomatic side ($n=9$).

Carotid enhancement

The mean common carotid enhancement on CTA was 343 HU (range 207–501 HU) with a standard deviation of 89 HU.

Carotid stenosis assessment

With DSA the exact degree of stenosis could not be measured in two cases because of severe narrowing of the distal lumen and extensive calcifications. In these cases, a 95% stenosis was assigned.

The diameter of the stenosis at its narrowest point varied from 0.2 to 3.3 mm (mean 1.1, SD 0.9), with 29/55 carotids having a diameter of less than 1.0 mm at the site of maximum stenosis. The diameter of the distal reference segment varied between 2.3 and 5.8 mm (mean 3.8, SD 0.9).

All observers could use visual estimation (method 1) for grading every carotid artery included in the study. With method 2 (caliper measurements using a wide window setting), observer 1 could not assess the stenosis degree in four cases, observer 2 in five, observer 3 in three and observer 4 in seven cases. On average, this amounted to 12% of carotid arteries. Reasons given for non-evaluability were the presence of distal collapse ($n=8$), calcifications ($n=5$) or high-grade stenosis resulting in reduction of intraluminal contrast enhancement ($n=6$). With method 3 (caliper measurements using a predefined window setting), the number of non-evaluable segments was nine, 12, 13

and five cases for observers 1–4, respectively (on average 24% of carotid arteries). For this method, the main reasons for non-evaluability were calcifications that could not be differentiated from lumen using the predefined window settings ($n=19$), or because the stenosis was so severe that the stenotic lumen was not visualized using the predefined window settings ($n=8$). Distal collapse was named as a reason in 12 cases. Figure 1 shows a case in which all methods could be applied, Figs. 2 and 3 provide examples for high-grade and calcified stenoses that could not be evaluated using the predefined window setting.

Inter-observer variability

Linear weighted kappa analysis showed that inter-observer variability was best for CTA_{VE}, compared with DSA as well as with the other caliper measurements with CTA. Especially, expert observers showed a very good agreement, with kappa 0.85 (95% CI: 0.73–0.96), compared with DSA (kappa 0.61, 95% CI: 0.44–0.78) and CTA_{WW} (kappa 0.54, 95% CI: 0.36–0.72). All kappa values are shown in Table 1. For CTA_{PW}, the kappa value was also very good, but 24% of carotid arteries were not analysed because of non-evaluability.

Results of the Bland-Altman analysis for inter-observer variability are shown in Table 2. Variability for visual estimation on CTA was comparable with DSA with bias <5% and SDD of 10–11%. For caliper measurements, however, variability was significantly higher with bias 5–10% and SDD of 13–16% ($p<0.05$).

Fig. 2 Example of a high-grade stenosis (on DSA: 95%) that cannot be evaluated using the predefined window (width/level: 142/ 238) because the lumen is no longer visible at this window setting (**b**). Using the wide window setting an 81–92% stenosis was measured by the four observers (**a**)
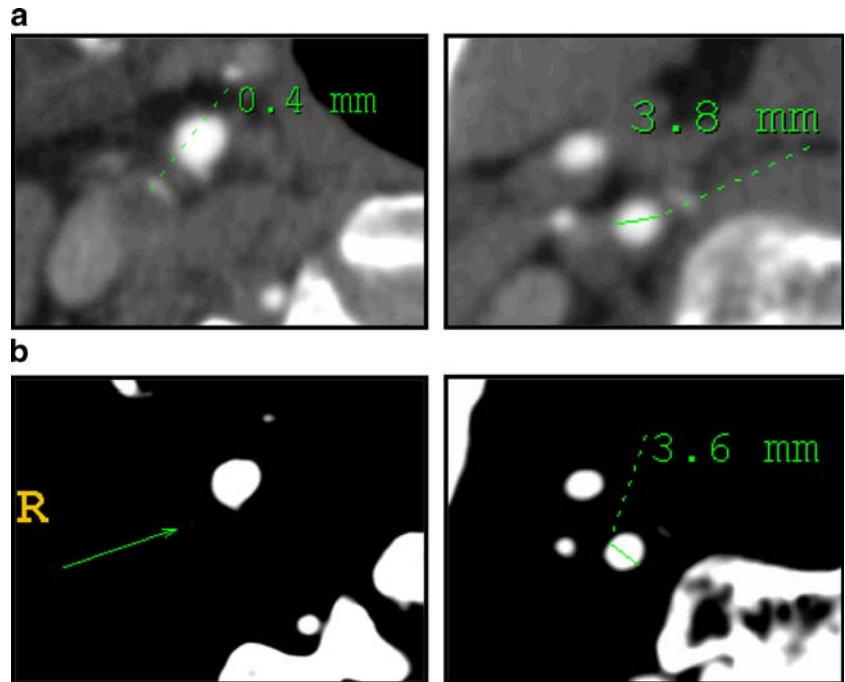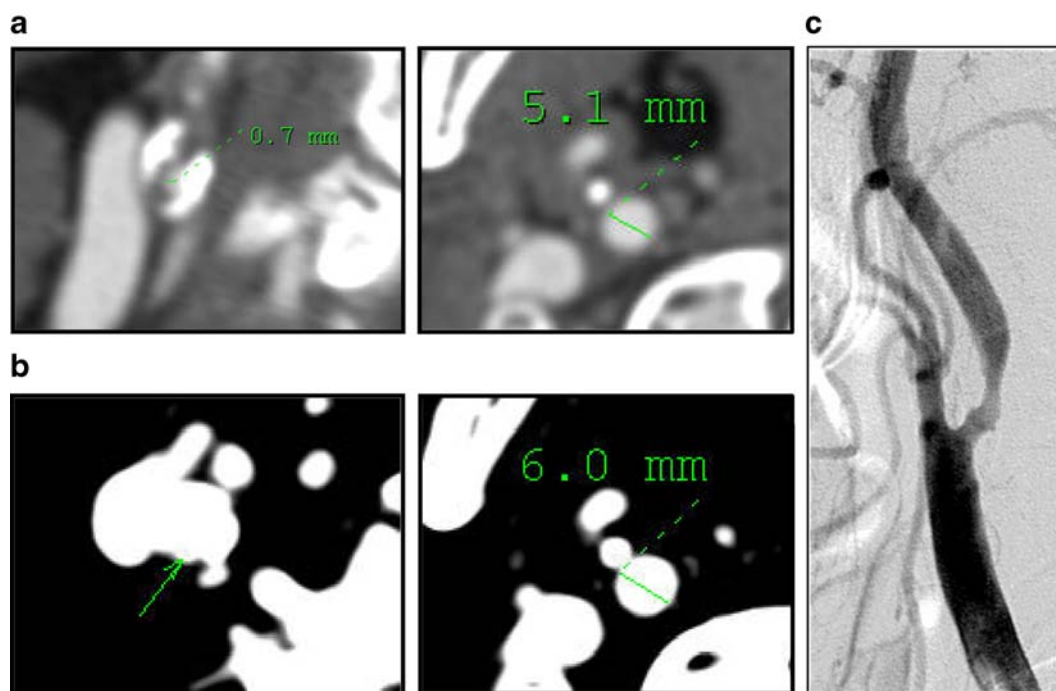
Fig. 3 Example of a calcified plaque [on DSA: 57% (**c**)] that cannot be evaluated using the predefined window (width/level: 83/138) because lumen and plaque can no longer be discriminated (**b**). Using the wide window setting a 62–86% stenosis was measured by the four observers (**a**)

Agreement of CTA grading relative to DSA

The agreement of CTA measurement compared to DSA, was best for visual estimation with a good agreement for both experienced and non-experienced observers. Individual kappa's for CTA_{VE} ranged from 0.67 to 0.76, compared to a moderate result for CTA_{WW} (kappa ranging from 0.54 to 0.61) and CTA_{PW} (kappa ranging from 0.46 to 0.54). These results are shown in Table 3.

Compared with the consensus reading on DSA as standard of reference, visual estimation on average resulted in a slight overestimation of stenoses (bias 5.8–8.0%), caliper measurements using wide window setting resulted on average in slight over- or underestimation (bias −0.4 to 8.8%), depending on the observer. Caliper measurements using the preset window settings resulted on average in a slight underestimation that varied from −6.3 to −3.0% (Table 4). The differences between the various techniques were not significant.

Table 1 Inter-observer variability for measurement of carotid artery stenosis for two experienced interventional radiologists on DSA, in comparison with inter-observer variability for grading of carotid stenosis by the two most experienced observers (observers 1 and 2) and the two less experienced observers (observers 3 and 4) using various CTA grading techniques (*VE* visual estimation, *WW* wide window caliper measurement, *PW* preset window caliper measurement)

Method	Observer	Linear weighted kappa	95% CI
DSA	Expert	0.61	0.44–0.78
CTA_{VE}	Expert	0.85	0.72–0.96
	Non-expert	0.75	0.60–0.89
CTA_{WW}	Expert	0.54	0.36–0.72
	Non-expert	0.51	0.33–0.68
CTA_{PW}	Expert	0.84	0.70–0.98
	Non-expert	0.65	0.49–0.83

Table 2 Inter-observer variability for measurement of carotid artery stenosis, analysed with Bland-Altman method. The table shows results of two experienced interventional radiologists on DSA in comparison with inter-observer variability for grading of carotid stenosis by the two most experienced observers (observers 1 and 2) using various CTA grading techniques [*LLoA/ULoA* lower/upper 95% level of agreement (% stenosis)]

	Bland-Altman analysis			
	Bias	SDD[a]	LloA	UloA
DSA	−4.8	11.6	−27.5	17.9
CTA$_{VE}$	−2.2	10.3	−22.4	18.0
CTA$_{WW}$	9.3	13.4	−17.1	35.6
CTA$_{PW}$	5.9	15.8	−25.1	36.9

[a]SDD was significantly different between techniques (Levene's test for homogeneity of variances, $p < 0.05$)

Variability between CT and DSA measurements was least for visual estimation (SDD 10.6–14.4%), followed by caliper measurements using wide window settings (SDD 12.0–16.7%). The largest variability between CT and DSA measurements was seen for method 3 (SDD 13.0–18.1%) (illustrated in Fig. 4).

Discussion

To our knowledge, this is the first study showing that visual estimation can outperform the use of caliper measurements for determining the degree of carotid artery stenosis on CTA exams with respect to agreement and reproducibility. Linear weighted kappa values were good for CTA$_{VE}$ in comparison with DSA, while these kappas were only moderate for caliper measurements in comparison with DSA. For inter-observer variability the linear weighted kappa was even "very good" for expert observers using CTA$_{VE}$ compared with "good" for experts using DSA.

In fact, the measurement variability between DSA and visual estimation on CTA (95% limits of agreement, −16% to 30%, SDD 11.5) is close to the inter-observer variability of repeated measurements on DSA that we found in this study (95% limits of agreement −28% to 18%, SDD 11.6), which is comparable with a previous study by Young and co-workers (95% LoA −22 to 22%, SDD 11) [21, 22].

At first sight these results appear surprising because objective measurements are usually considered more precise and reproducible than subjective estimation. The results become less surprising, however, when one examines the way caliper measurements are performed: first, the reference region and the region of the maximum stenosis have to be identified visually, then the precise position of the calipers at the vessel borders has to be identified. Since two regions are involved, four such decisions about where to place the caliper have to be made. Together with the decision about the location of the site of measurement, six subjective decisions have to be made to come up with one stenosis grade. All decisions are potentially problematic: in complex stenoses especially, the precise location of the maximum stenosis is difficult to determine. If the vessel distal to the stenosis does not have a uniform diameter, variations in the position of the reference measurement will affect stenosis quantification. Finally, positioning the cursors at the vessel borders is known to be difficult and has been the main reason why alternative techniques have been suggested [10, 11].

Many methods have been proposed to achieve more sharply defined vessel edges by adjusting window width and level. However, they were only tested in phantoms, thereby excluding such clinical problems as calcifications of high-grade stenoses. Dix et al. [19] were the first to demonstrate that the use of binominal criteria (also called "full width at half maximum method") resulted in better reproducibility of carotid measurements compared to wider window settings. They selected a fixed level, set

Table 3 Comparison of linear weighted kappa for various methods on CTA compared to the consensus reading of DSA (*VE* visual estimation, *WW* wide window caliper measurement, *PW* preset window caliper measurement)

CTA method	Observer	Linear weighted kappa	95% CI
VE	1	0.69	0.53–0.84
	2	0.68	0.53–0.84
	3	0.76	0.62–0.90
	4	0.67	0.51–0.84
WW	1	0.61	0.44–0.79
	2	0.54	0.35–0.73
	3	0.56	0.38–0.74
	4	0.56	0.38–0.74
PW	1	0.50	0.32–0.68
	2	0.50	0.31–0.70
	3	0.54	0.34–0.74
	4	0.46	0.29–0.64

Table 4 Bland-Altman analysis showing the agreement between DSA and CTA for four different observers (observers 1–4) and three measurement methods. The consensus DSA reading was used as standard of reference. SDD was significantly better for visual estimation compared with the other techniques for observer 1 and observer 3 ($p < 0.05$, Levene's test) (*LoA* limits of agreement)

CTA method	Feature	Observer 1	Observer 2	Mean of observers 1 and 2	Observer 3	Observer 4
VE	Bias	5.8	8.0	6.9	7.5	6.0
	SDD	10.6	12.5	11.5	10.9	14.4
	LoA	(26.5, −14.9)	(32.4, −16.4)	(29.5, −15.6)	(28.9, −13.8)	(34.2, −22.9)
	No. of patients not evaluable	0	0		0	0
WW	Bias	8.8	−0.4	4.1	−2.5	2.9
	SDD	12.0	16.4	14.2	15.7	16.7
	LoA	(32.3, −14.7)	(31.7, −32.4)	(32.0, −23.6)	(28.3, −33.2)	(35.6, −29.8)
	No. of patients not evaluable	4	5		3	7
PW	Bias	−4.8	−6.3	−5.5	−3.0	−3.4
	SDD	14.4	13.0	13.8	12.8	18.1
	LoA	(23.5, −33.1)	(19.8, −31.9)	(21.4, −32.5)	(22.1, −28.1)	(32.1, −38.8)
	No. of patients not evaluable	9	12		13	5

at halfway between the density within the vessel lumen and the surrounding tissue, combined with a window width (HU) of one, thus creating a black and white image with sharply delineated vessel edges. However, this principle does not work in cases with a subtotal stenosis: the contrast enhancement within the lumen decreases due to partial volume effect, and the level for the binary window will be set too low, resulting in an overestimation of the residual lumen and an underestimation of the stenosis.

Liu et al. [18] presented a method with preset narrow window settings that depend on the CT number in the carotid artery. This method also suffers from shortcomings in the presence of subtotal stenoses: when intra-luminal contrast decreases in high-grade stenoses, the window level set according to the predefined criteria will be too high, and the lumen is no longer visualized. Calcifications constitute another limiting factor for this method: because they have a higher CT number than the vessel lumen, lumen and calcifications both appear 'white' at the predefined window settings and can therefore no longer be differentiated. Even in the absence of intra-luminal contrast decrease or calcifications, the method was not as good as visual estimation, even for less-experienced observers.

These limitations can be overcome if a wide window width is used. A wide window offers the observer the opportunity to correctly interpret misleading CT values caused by volume averaging, presenting as decreased intra-luminal contrast density in the region of a severe stenosis, or apparent narrowing of the lumen adjacent to calcifications due to blooming artefacts. At the level of a narrow stenosis, however, blurring of vessel edges is almost

inevitable and may hamper proper selection of the point where the caliper must be placed.

Visual estimation, in contrast to caliper measurements, is a simple and quick technique. As mentioned by previous authors, visual estimation gives the observer the freedom to take decreasing luminal attenuation or calcifications into account when estimating the degree of a stenosis [23–25]. Our results suggest that good correlation with arterial DSA can be achieved with visual estimation, and that even inexperienced radiologists usually perform better with visual estimation than with caliper measurements, both regarding accuracy and reproducibility of measurements.

This study suffered from the following limitations.

We determined the degree of stenosis in a pre-selected group of symptomatic patients. Contralateral arteries were selectively included and thus the study comprises an inhomogeneous group of symptomatic and asymptomatic carotids. The results may potentially be different in patient groups with less severe stenoses [26, 27]. However, since CTA is rapidly replacing DSA as a tool for work-up of patients with carotid symptoms, our results should hold true for the most important clinical indication.

We analysed grading accuracy and variability by measuring the absolute differences in stenosis grades instead of using stenosis categories. Because cut-off values for taking clinical decisions vary [2, 4–6], we opted for this approach. As a consequence, we did not calculate sensitivity and specificity of detecting "significant" or "actionable" stenoses. Our aim was to determine the exact measurement error for the three methods compared with angiography. For that purpose the Bland-Altman method used in this study is well suited.

Estimation

Caliper Wide Window

Caliper Preset Window
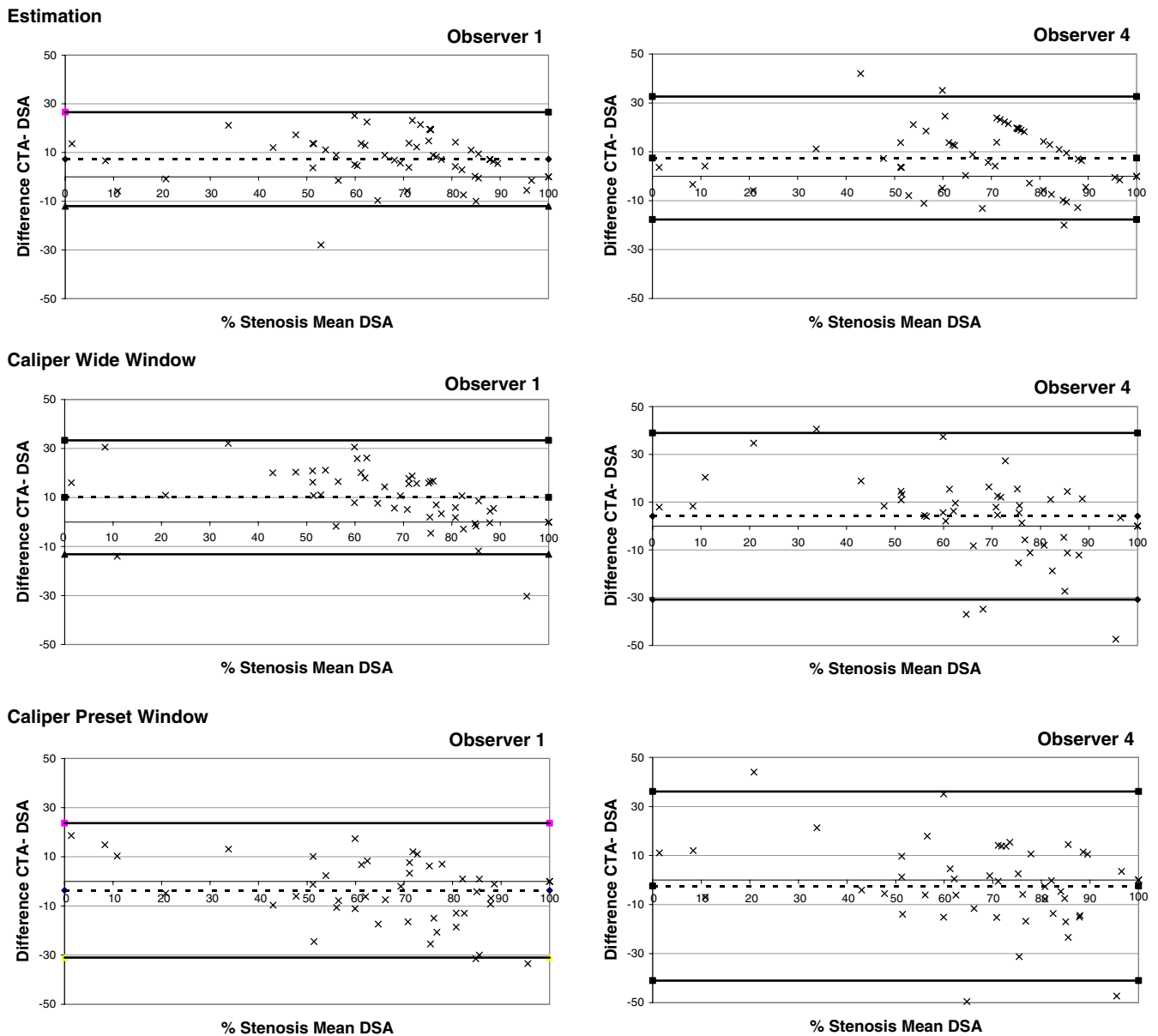
Fig. 4 Bland Altman plots for comparison of DSA to CTA$_{VE}$, CTA$_{WW}$ and CTA$_{PW}$ for observer 1 with the longest experience in reading CTA (>15 years) and observer 4 who had the least experience (<1 year). Note that independent of experience the best agreement with DSA was found for visual estimation

The fact that visual estimation was performed using 10% categories is a potential further shortcoming. However, we considered a finer or even continuous scale not to be in line with the limited spatial resolution of CT and the capabilities of human observers. The 10% categories correspond well to the recently suggested cut-off values in the group of >50% stenosis [4–6]. In addition, we did not further discriminate between near-occlusion and high-grade stenosis in the group of stenoses between 90 and 99%, which can be important in clinical practice because these two entities imply

different treatment strategies [4]. Despite this lower intrinsic precision of our visual grading system, however, results were best with visual estimation.

To our knowledge this is the first comparison of measurement techniques for determining carotid artery stenosis using multidetector-row CTA in a clinical setting. Our results show that neither DSA nor CTA provide a perfectly reproducible measurement. Using the caliper measurement on DSA as standard of reference, the technique that is considered most subjective (visual estimation) had the least measurement variability, independent of observer

experience. These results show that visual estimation on CTA provides the best correlation with angiographic measurements in patients with symptomatic carotid artery stenosis and outperforms manual caliper measurements.

References

1. Rothwell PM, Coull AJ, Silver LE et al (2005) Population-based study of event-rate, incidence, case fatality, and mortality for all acute vascular events in all arterial territories (Oxford Vascular Study). Lancet 366:1773–1783

2. North American Symptomatic Carotid Endarterectomy Trial Collaborators (1991) Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. N Engl J Med 325:445–453

3. Roubin GS, Iyer S, Halkin A et al (2006) Realizing the potential of carotid artery stenting: proposed paradigms for patient selection and procedural technique. Circulation 113:2021–2030

4. Rothwell PM, Eliasziw M, Gutnikov SA et al (2003) Analysis of pooled data from the randomised controlled trials of endarterectomy for symptomatic carotid stenosis. Lancet 361:107–116

5. Halliday A, Mansfield A, Marro J et al (2004) MRC Asymptomatic Carotid Surgery Trial (ACST) Collaborative Group. Prevention of disabling and fatal strokes by successful carotid endarterectomy in patients without recent neurological symptoms: randomised controlled trial. Lancet 363:1491–1502

6. Yadav JS, Wholey MH, Kuntz RE et al (2004) Protected carotid-artery stenting versus endarterectomy in high-risk patients. N Engl J Med 351:1493–1501

7. Prokop M (2000) Multislice CT angiography. Eur J Radiol 36:86–96

8. Josephson SA, Bryant SO, Mak HK et al (2004) Evaluation of carotid stenosis using CT angiography in the initial evaluation of stroke and TIA. Neurology 63:457–460

9. North American Symptomatic Carotid Endarterectomy Trial (1991) Methods, patient characteristics, and progress. Stroke 22:711–720

10. Dillon EH, van Leeuwen MS, Fernandez MA et al (1993) CT angiography: application to the evaluation of carotid artery stenosis. Radiology 189:211–219

11. Cumming MJ, Morrow IM (1994) Carotid artery stenosis; a prospective comparison of CTA and conventional angiography. AJR Am J Roentgenol 163:517–523

12. Marks MP, Napel S, Jordan JE et al (1993) Diagnosis of carotid artery disease: preliminary experience with MIP spiral CT angiography. AJR Am J Roentgenol 160:1267–1271

13. Anderson GB, Ashforth R, Steinke DE et al (2000) CT angiography for the detection and characterization of carotid artery bifurcation disease. Stroke 31:2168–2174

14. Randoux B, Marro B, Koskas F (2001) Carotid artery stenosis: prospective comparison of CT, 3D gadolinium enhanced MR and conventional angiography. Radiology 220:179–185

15. Verhoek G, Costello P, Khoo EW (1999) Carotid bifurcation CTA: assessment of interactive volume rendering. J Comput Assist Tomogr 23:590–596

16. Leclerc X, Godefroy O, Pruvo JP (1995) Computed tomographic angiography for the evaluation of carotid artery stenosis. Stroke 26:1577–1581

17. Leclerc X, Godefroy O, Lucas C et al (1999) Internal carotid arterial stenosis: CT angiography with volume rendering. Radiology 210:673–682

18. Yi Liu MD, Hopper KD Mauger DT et al (2000) CT angiographic measurement of the carotid artery: optimizing visualization by manipulating window and level settings and contrast material attenuation. Radiology 217:494–500

19. Dix JE, Evans AJ, Kallmes DF et al (1997) Accuracy and precision of CT angiography in a model of carotid artery bifurcation stenosis. AJNR Am J Neuroradiol 18:409–415

20. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310

21. Rothwell PM, Gibson RJ, Slattery J et al (1994) Prognostic value and reproducibility of measurements of carotid stenosis. A comparison of three methods on 1001 angiograms. European Carotid Surgery Trialists' Collaborative Group. Stroke 25:2440–2444

22. Young GR, Humphrey PR, Nixon TE et al (1996) Variability in measurement of extracranial internal carotid artery stenosis as displayed by both digital subtraction and magnetic resonance angiography: an assessment of three caliper techniques and visual impression of stenosis. Stroke 27:467–473

23. Chen CJ, Lee TH, Hsu HL et al (2004) Multi-Slice CT angiography in diagnosing total versus near occlusions of the internal carotid artery: comparison with catheter angiography. Stroke 35:83–85

24. Zhang Z, Berg MH, Ikonen AE et al (2004) Carotid artery stenosis: reproducibility of automated 3D CT angiography analysis method. Eur Radiol 14:665–672

25. Prokop M, Shin HO, Schanz A et al (1997) Use of maximum intensity projections in CT angiography: a basic review. Radiographics 17:433–451

26. Rothwell PM (2000) Analysis of agreement between measurements of continuous variables: general principles and lessons from studies of imaging of carotid stenosis. J Neurol 247:825–834

27. Rothwell PM, Pendlebury ST, Wardlaw J et al (2000) Critical appraisal of the design and reporting of studies of imaging and measurement of carotid stenosis. Stroke 31:1444–1450