

Limited Value of Temporal Artery Ultrasonography Examinations for Diagnosis of Giant Cell Arteritis: Analysis of 77 Subjects

CARLA MALDINI, CAROLINE DÉPINAY-DHELLEMMES, THI T.S. TRA, MICHEL CHAUVEAU, YANNICK ALLANORE, LAURE GOSSEC, GENEVIÈVE TERRASSE, LOÏC GUILLEVIN, JOËL COSTE, and ALFRED MAHR

ABSTRACT. *Objective.* Use of TA-US for diagnostic investigation of giant cell arteritis (GCA) has been proposed but remains a matter of debate because of the heterogeneous findings. We retrospectively evaluated operating characteristics of temporal artery ultrasonography (TA-US) in a single teaching hospital. *Methods.* All subjects with suspected GCA had been seen between 2002 and 2008 and had undergone TA-US with continuous-wave Doppler (until 2004) or color duplex ultrasonography (after 2004), followed within 30 days by a temporal artery biopsy (TAB). TA-US findings were compared with TAB-proven GCA and clinically diagnosed GCA. Results are expressed as sensitivities, specificities, and positive (LR+) and negative likelihood ratios (LR-) of stenoses, occlusions, and the halo sign; for the latter, only color duplex TA-US was considered. *Results.* Seventy-seven patients fulfilled the selection criteria; 13 had TAB-proven and 19 had clinically defined GCA. Stenoses/occlusions were seen on 45.5% of TA-US and the halo sign was seen only once (3.2%) in 31 duplex TA-US. Respective sensitivities, specificities, LR+, and LR- for GCA diagnosis (using TAB-proven/clinically defined GCA as reference standards) were 69%/53%, 59%/57%, 1.7/1.2, and 0.5/0.8 for stenoses and/or occlusions, and 17%/10%, 100%/100%, infinite/infinite, and 0.8/0.9 for the halo sign. *Conclusion.* The halo sign showed 100% specificity for GCA but only 10%–17% sensitivity. Stenoses/occlusions were of low diagnostic value. These observations suggest that TA-US is neither an effective substitute for TAB nor a reliable screening test to decide which patients can be safely spared TAB. (First Release September 1 2010; J Rheumatol 2010;37:2326–30; doi:10.3899/jrheum.100353)

Key Indexing Terms:

GIANT CELL ARTERITIS
HALO

TEMPORAL ARTERY ULTRASONOGRAPHY
SENSITIVITY

SPECIFICITY

Ultrasonography of the superficial temporal arteries (TA-US) in the diagnostic investigation for giant cell arteritis (GCA) has generated much interest for almost 3 decades. Early investigators applying Doppler US reported hemo-

dynamic flow abnormalities with stenoses and occlusions of the temporal artery^{1,2}. Later studies using color duplex ultrasonography, which also provided information on vessel anatomy, showed edematous thickening of the arterial wall, called the “halo” sign³. Compared to temporal artery biopsy (TAB), the traditional gold standard, TA-US appeared to be a promising, noninvasive test for diagnosis of GCA⁴.

Despite extensive investigation, the accuracy of TA-US as a diagnostic or screening test for GCA remains uncertain. Studies have given this test a wide range of performance characteristics including 100% sensitivity^{5,6,7} or 100% specificity^{3,8,9}. Moreover, it is unclear whether the halo sign alone or any abnormal imaging feature is of diagnostic value. The notable inconsistencies of findings crystallized in a comprehensive metaanalysis of 23 studies, which found significant between-study variability for almost all pooled estimates of the operating characteristics of TA-US for GCA¹⁰, thereby questioning the generalizability of its results.

Taking advantage of longstanding access to TA-US in our institution, we evaluated its validity for diagnosis of GCA to further clarify this matter.

From the Department of Internal Medicine, National Referral Center for Necrotizing Vasculitides and Systemic Sclerosis; Department of Pathology; Vascular Investigation Unit; Department of Rheumatology A; Department of Rheumatology B, UPRES-EA 4058; Department of Biostatistics, Hospital Cochin, Assistance Publique-Hôpitaux de Paris, University Paris 5-René Descartes, Paris, France.

C. Maldini, MD; L. Guillevin, MD, Department of Internal Medicine, National Referral Center for Necrotizing Vasculitides and Systemic Sclerosis, Hospital Cochin; C. Dépinay-Dhellemmes, MD; G. Terrasse, MD, Department of Pathology, Hospital Cochin; T.T.S. Tra, MD; M. Chauveau, MD, Vascular Investigation Unit, Hospital Cochin; Y. Allanore, MD, PhD, Department of Rheumatology A, Hospital Cochin; L. Gossec, MD, PhD, Department of Rheumatology B, UPRES-EA 4058, Hospital Cochin; J. Coste, MD, PhD, Department of Biostatistics, Hospital Cochin; A. Mahr, MD, PhD, Department of Internal Medicine, National Referral Center for Necrotizing Vasculitides and Systemic Sclerosis, and Department of Biostatistics, Hospital Cochin.

Address correspondence to Dr. A. Mahr, Department of Internal Medicine, Hôpital Cochin, 27 rue du Faubourg Saint-Jacques, 75679 Paris Cedex 14, France. E-mail: alfred.mahr@cch.aphp.fr

Accepted for publication June 1, 2010.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2010. All rights reserved.

MATERIALS AND METHODS

Setting and selection of subjects. Our retrospective study was conducted in a teaching hospital, where all vascular ultrasonography examinations are performed by specialized physicians affiliated with a vascular investigation unit. Routine TA-US for patients with suspected GCA was made available in 1981². Standardized reports for all ultrasonography examinations have been stored in a computerized database since 2002.

For the period January 2002–September 2008, we crosslinked the databases of the vascular investigation unit and the local pathology department to identify all patients who underwent TA-US and/or TAB. Only patients who had both tests were considered eligible for our study's primary objective. These selection criteria were chosen to enable comparison of TA-US features against TAB findings for all patients and relied on the supposition that only patients who underwent TAB had high clinical suspicion of GCA. Moreover, only patients whose TAB was obtained within 30 days after TA-US were included in the primary analyses.

Routine TA-US protocol. Until May 2004, TA-US relied on continuous-wave Doppler ultrasonography using the pencil probe (5 MHz) of an ATL Apogee 800 sonograph. Starting from June 2004, color-coded, pulsed-wave duplex ultrasonography was introduced with the linear array probe (7.5 MHz) of a Toshiba Aplio sonograph.

Routine TA-US protocol included bilateral insonation of the common superficial temporal artery and its frontal branches, followed as far as possible. The following abnormalities were sought: occlusion (defined as a nonexistent Doppler signal), stenosis (localized acceleration with a ≥ 2 -fold peak systolic velocity increase or a dampened velocity curve with a ≥ 2 -fold peak systolic velocity decrease, compared with either the upstream segment or contralateral artery), and, once the duplex sonograph was used, halo sign (defined as an anechoic ring ≥ 0.3 mm thick, separating surrounding tissue from the colored arterial lumen, seen in both transverse and longitudinal planes).

Data collection. TA-US and TAB findings and medical charts were reviewed for all patients selected for the primary analyses. To guarantee high-quality data retrieval, TA-US reports were independently reviewed by 1 clinical investigator and 1 physician of the vascular investigation unit. Similarly, the original TAB reports were assessed by 1 clinical investigator and compared with an *ad hoc* reading of all TAB slides by a pathologist, who remained blinded to the original TAB report findings. Between-investigator discrepancies concerning information extracted from the TA-US reports and between the initial and second TAB readings were resolved by consensus discussion. The medical charts of all selected patients were assessed by 1 investigator to control fulfillment of American College of Rheumatology (ACR) criteria¹¹ and to record clinical manifestations, inflammatory markers (erythrocyte sedimentation rate, C-reactive protein) and glucocorticoid therapy at first consultation and during followup, and final diagnoses.

In addition, 2 independent investigators abstracted, and reached consensus on, the TA-US or TAB reports' principal findings for the patients not eligible for the primary analyses and who underwent TA-US without subsequent TAB or TAB without prior TA-US.

Reference standards definitions. "TAB-proven GCA" was defined based on the presence of a mononuclear cell infiltrate at the intima-media junction or in the media; additional features, e.g., giant cells or disruption of the internal elastic lamina, were considered discretionary. "Clinically defined GCA" referred to the combination of ACR criteria for GCA¹¹, the waning of clinical and laboratory signs under high-dose prednisone, and the absence of other diagnoses over a followup period of ≥ 6 months.

Statistical analyses. Using each of the 2 reference standards, sensitivities, specificities, and positive (LR+) and negative likelihood ratios (LR-) were computed for stenoses, occlusions, and the halo sign; for the halo sign, calculations concerned only the subset of duplex TA-US examinations. LR+/LR- values were calculated based on the respective formulas LR+ = sensitivity/(100 - specificity) and LR- = (100 - sensitivity)/specificity. LR+ values > 10 and LR- values < 0.1 are generally considered to allow a

conclusive increase or decrease, respectively, of the posttest likelihood of the presence of the disease under investigation¹². Confidence intervals of 95% for sensitivities/specificities were calculated applying the Wilson score method without continuity correction¹³; for LR+/LR-, 95% CI were computed as described¹⁴.

To assess the potential influence of workup bias, we applied the method described by Diamond¹⁵, which provides sensitivity and specificity estimates that incorporate information from the patients who underwent the test under evaluation but not the gold standard test. The formula for corrected sensitivity was: $PPV/[PPV + (100 - NPV) \times (100 - PA)/PA]$ where PPV is the positive predictive value, NPV the negative predictive value, and PA the proportion of abnormal TA-US test results in the entire population. The formula for corrected specificity was: $100 \times NPV/[NPV + (100 - PPV) \times (PA/[100 - PA])]$ ¹⁵.

RESULTS

Patient selection and characteristics. Figure 1 shows patient selection; 77 subjects (mean age at TA-US 72.2 ± 9.1 yrs; 64% women), all referred for suspected GCA, were retained for the primary analyses. For the 2 patients who had sequential bilateral TAB within the selected 30-day window, the second examination findings were used for the analyses. The mean TA-US–TAB interval was 4.8 ± 4.7 days (range 0–21). A total of 13 (16.9%) had a positive TAB and 19 (24.7%) fulfilled the clinical criteria for GCA diagnosis; 1 TAB-positive patient failed to meet ACR criteria. For these 20 patients, the mean time from TA-US to initiation of therapy with high-dose glucocorticoids was 0.4 ± 5.3 days (range -10 to 12).

Computed TA-US operating characteristics. The 77 TA-US subjects included 46 who had Doppler ultrasonography and 31 who had duplex ultrasonography. The same operator (MC) performed 62 (80.5%) studies, and 4 other staff physicians, the remaining 15 (19.5%). Flow abnormalities were identified in 35/77 (45.5%), while an additional (bilateral) halo sign was seen in 1 (3.2%) of the 31 duplex ultrasonographies.

Figure 2 and Table 1 show the breakdown of TA-US findings and the computed TA-US operating characteristics for the study sample for both reference standards. Table 1 also reports the results of a sensitivity analysis, assessing the diagnostic accuracy of bilateral TA-US abnormalities. Sensitivity analyses addressing the operating characteristics of stenoses and/or occlusions evaluated in continuous-wave versus duplex TA-US or stratified by TA-US operators (the most experienced assessor vs the 4 others) did not yield substantially different findings (data not given).

Additional estimates to correct workup bias were calculated for the operating characteristics of stenoses and/or occlusions. These corrected estimates used a 23.8% proportion of abnormal TA-US, i.e., showing stenoses and/or occlusions, in the combined samples of the 77 subjects used for the primary analyses and the 204 subjects who had TA-US without TAB (Figure 1). When assessed against TAB-proven GCA, the "crude" sensitivity and specificity of stenoses and/or occlusions were 69% and 59%, respectively (Table 1).

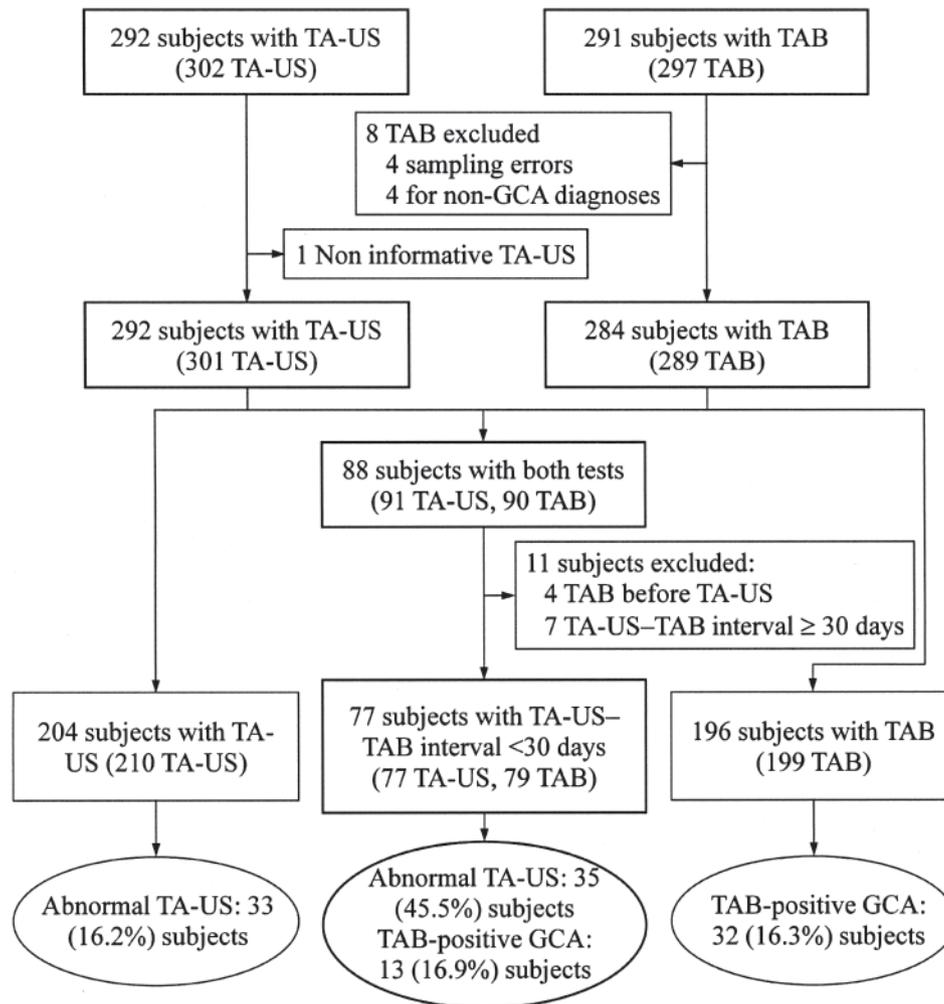


Figure 1. Selection of 77 study subjects extracted from all patients who underwent ultrasonography of the superficial temporal arteries (TA-US) and/or temporal artery biopsy (TAB) from 2002 to 2008. Some patients had multiple TA-US and/or TAB. GCA: giant cell arteritis.

After accounting for workup bias, sensitivity decreased to 46% and specificity increased to 80%. However, the corresponding “corrected” LR+ and LR– values changed little, with adjusted values of 2.24 and 0.68, respectively. Similar findings with small changes were observed for LR+ and LR– values corrected for workup bias, with respective values of 1.33 and 0.91, for stenoses and/or occlusions assessed against clinically defined GCA.

TA-US and TAB findings in patients not eligible for the primary analyses. Among the 204 patients with TA-US but no TAB (210 TA-US: 104 Doppler ultrasonography and 106 duplex ultrasonography), 33 (16.2%) patients had abnormal TA-US findings. Those abnormalities were stenoses and/or occlusions in 32 patients and an isolated (unilateral) halo in 1. Compared to the 77 subjects considered for the primary analyses, the lower abnormal TA-US rate in these patients indicated the possibility of workup bias. In contrast, review of the TAB reports of the 196 patients with no prior TA-US yielded a

16.3% TAB-positivity rate very similar to that found for the 77 subjects used for the primary analyses (Figure 1).

DISCUSSION

Our analysis of 77 subjects with suspected GCA who underwent Doppler or duplex TA-US before TAB credited the halo sign with 100% specificity (with a corresponding infinite LR+) but with very low sensitivity. Also, stenoses and occlusions detected by TA-US appeared to be of little clinical significance, as indicated by their modest LR+ and LR– values. These findings corroborate that, in our institution and in everyday clinical settings, TA-US unfortunately contributes little to diagnosis of GCA.

Surprisingly, a halo sign was seen in only 1 GCA case among the 31 subjects examined by duplex ultrasonography. This resulted in low 17%/10% sensitivities for TAB-proven/clinically defined GCA; however, these estimates must be interpreted with their broad 95% CI in mind.

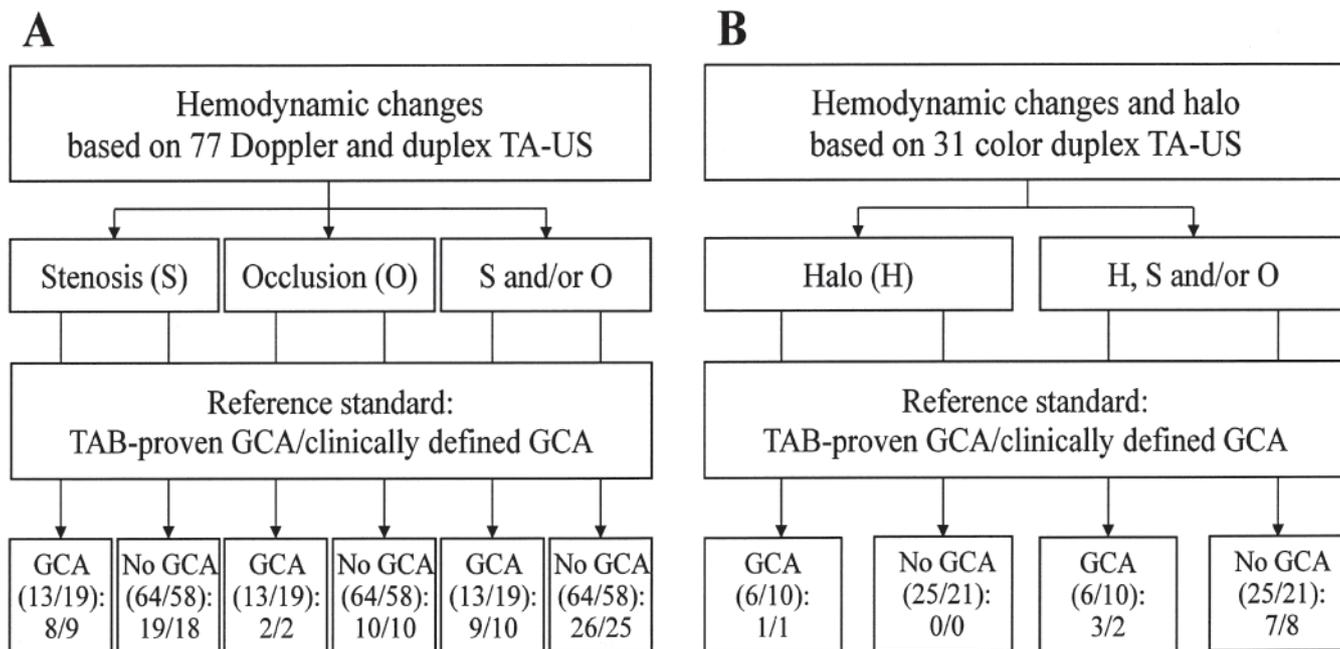


Figure 2. Ultrasonography findings for the superficial temporal arteries (TA-US) of 77 subjects with suspected giant cell arteritis (GCA). A. Assessment of hemodynamic changes. B. Subanalysis of the 31 color duplex TA-US also allowing detection of the halo sign. GCA diagnosis is based on positive histology (proven by temporal artery biopsy, TAB) or clinically defined. Values in parentheses refer to numbers of cases meeting those standards, and values below the parentheses are the number of cases for which the corresponding TA-US findings were observed in each category.

Table 1. Sensitivities, specificities, positive (LR+), and negative likelihood ratios (LR-) for temporal artery ultrasonography (TA-US) findings versus TAB-proven or clinically defined GCA. Values in parentheses are the 95% CI.

	TAB-Proven GCA			LR-	Clinically Defined GCA			
	Sensitivity, %	Specificity, %	LR+		Sensitivity, %	Specificity, %	LR+	LR-
Doppler and duplex TA-US, n = 77								
Stenosis	62 (36-82)	70 (58-80)	2.1 (1.2-3.7)	0.5 (0.3-1.1)	47 (27-68)	69 (56-79)	1.5 (0.8-2.8)	0.8 (0.5-1.2)
Occlusion	15 (4-42)	84 (74-91)	1.0 (0.2-4.0)	1.0 (0.8-1.3)	11 (3-31)	83 (71-90)	0.6 (0.1-2.5)	1.1 (0.9-1.3)
Any stenosis and/or occlusion	69 (42-87)	59 (47-71)	1.7 (1.1-2.7)	0.5 (0.2-1.2)	53 (32-73)	57 (44-69)	1.2 (0.7-2.1)	0.8 (0.5-1.4)
Any stenosis and/or occlusion (seen bilaterally)	23 (8-50)	88 (77-94)	1.8 (0.6-6.0)	0.9 (0.6-1.2)	16 (6-38)	86 (75-93)	1.1 (0.3-3.9)	1.0 (0.8-1.2)
Duplex TA-US, n = 31								
Halo sign	17 (3-56)	100 (87-100)	Infinite (NA)	0.8 (0.6-1.2)	10 (2-40)	100 (85-100)	Infinite (NA)	0.9 (0.7-1.1)
Any stenosis, occlusion, and/or halo	50 (19-81)	72 (52-86)	1.8 (0.6-4.9)	0.7 (0.3-1.6)	20 (6-51)	62 (41-79)	0.5 (0.1-2.0)	1.3 (0.8-2.0)
Any stenosis, occlusion, and/or halo (seen bilaterally)	17 (3-56)	96 (80-99)	4.2 (0.3-57.5)	0.9 (0.6-1.3)	10 (2-40)	95 (77-99)	2.1 (0.1-30.3)	0.9 (0.8-1.2)

NA: not applicable. TAB: temporal artery biopsy; GCA: giant cell arteritis.

Previous studies reported 9%–100% sensitivity for the halo sign¹⁰, and these variations probably mirror interoperator variability of stringency for defining this feature and/or differences in patient selection. In light of the previously described strong concordance between physical examination abnormalities and halo sign¹⁶, higher observed halo rates might be explained by study samples skewed toward patients with suspected GCA and clinically prominent temporal artery involvement. Our observed 16.9% prevalence of TAB-positive GCA in a population of biopsied patients agrees with previously reported observations^{17,18,19,20} and

seems to indicate that the selected study sample was representative of the actual target population for which TAB is usually obtained.

In our study, TA-US findings of stenoses and/or occlusions yielded moderate 53%/69% sensitivities and 59%/57% specificities. Those sensitivities are roughly in keeping with the metaanalysis estimate of 68%/66% sensitivities for stenoses and/or occlusions¹⁰. In contrast, the metaanalysis showed that visualization of any flow abnormality, when assessed against clinical GCA criteria, reached 95% specificity, but was based on the pooling of only 4

studies¹⁰. The limited specificity of stenoses/occlusions found in our study likely reflects that intimal proliferation and fibrosis can also occur in degenerative temporal artery disease²¹.

A workup bias may have occurred because our primary analyses used only a subset of all the patients undergoing TA-US in our institution over the 7-year study period (Figure 1). This bias, which is common in evaluations of diagnostic tests for which an invasive procedure is the reference standard, originates from the preferential application of the reference standard (here, TAB) to patients with a positive result of the diagnostic test under evaluation (here, TA-US)^{15,22}. In the end, a workup bias inflates the test's true sensitivity and decreases its true specificity^{15,22}. Thus, because the likelihood measures integrate both sensitivity and specificity, we think that our LR+/LR- estimates were not substantially distorted by a potential workup bias. Adjusting for workup bias shifted uncorrected-to-corrected sensitivity and specificity for stenoses and/or occlusions (against TAB) from 69% to 46% and from 59% to 80%, respectively, but LR+/LR- values were only minimally affected by these changes.

Our observations challenge the usefulness of TA-US as a diagnostic or screening tool for GCA. The failure of studies to consistently demonstrate either high sensitivity or high specificity (or more accurately, high LR+ or low LR-) continues to prevent use of TA-US findings to conclusively rule a diagnosis of GCA in or out. Even if we were to accept our finding and that of other studies^{3,8,9}, suggesting that the halo sign be considered pathognomonic of GCA, in our setting, adopting a diagnostic algorithm based on TA-US would have eventually spared only 1 of 31 TAB. In general, the apparent difficulties in reproducing TA-US performance characteristics from one setting to another probably highlight its high level of operator dependency and entail a considerable drawback to use of this test for diagnosis of GCA.

ACKNOWLEDGMENT

The authors are indebted to Peter Villiger, MD, PhD, for valuable discussions, and to Janet Jacobson for editorial assistance.

REFERENCES

1. Dany F, Liozon F, Adenis JP, Michel JP, Vidal E, Kim M. Doppler examination in temporal arteritis. Preliminary results [French]. *J Mal Vasc* 1981;6:273-7.
2. Menkes CJ, Branche I, Feldmann JL, Chauveau M, Delbarre F. Application of the Doppler effect to the detection of Horton's temporal arteritis [French]. *Nouv Presse Méd* 1981;10:2371.
3. Schmidt WA, Kraft HE, Vorpahl K, Volker L, Gromnica-Ihle EJ. Color duplex ultrasonography in the diagnosis of temporal arteritis. *N Engl J Med* 1997;337:1336-42.
4. Pipitone N, Versari A, Salvarani C. Role of imaging studies in the diagnosis and follow-up of large-vessel vasculitis: an update. *Rheumatology* 2008;47:403-8.
5. LeSar CJ, Meier GH, DeMasi RJ, Sood J, Nelms CR, Carter KA, et al. The utility of color duplex ultrasonography in the diagnosis of temporal arteritis. *J Vasc Surg* 2002;36:1154-60.
6. Stammer F, Ysermann M, Mohr W, Kuhn C, Goethe S. Value of color-coded duplex ultrasound in patients with polymyalgia rheumatica without signs of temporal arteritis [German]. *Dtsch Med Wochenschr* 2000;125:1250-6.
7. Venz S, Hosten N, Nordwald K, Lemke AJ, Schroder R, Bock JC, et al. Use of high resolution color Doppler sonography in diagnosis of temporal arteritis [German]. *Rofo* 1998;169:605-8.
8. Reinhard M, Schmidt D, Hetzel A. Color-coded sonography in suspected temporal arteritis — experiences after 83 cases. *Rheumatol Int* 2004;24:340-6.
9. Schmid R, Hermann M, Yannar A, Baumgartner RW. Color duplex ultrasound of the temporal artery: replacement for biopsy in temporal arteritis [German]. *Ophthalmologica* 2002;216:16-21.
10. Karassa FB, Matsagas MI, Schmidt WA, Ioannidis JP. Meta-analysis: test performance of ultrasonography for giant-cell arteritis. *Ann Intern Med* 2005;142:359-69.
11. Hunder GG, Bloch DA, Michel BA, Stevens MB, Arend WP, Calabrese LH, et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis Rheum* 1990;33:1122-8.
12. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;329:168-9.
13. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;17:857-72.
14. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol* 1991;44:763-70.
15. Cecil MP, Kosinski AS, Jones MT, Taylor A, Alazraki NP, Pettigrew RI, et al. The importance of work-up (verification) bias correction in assessing the accuracy of SPECT thallium-201 testing for the diagnosis of coronary artery disease. *J Clin Epidemiol* 1996;49:735-42.
16. Salvarani C, Silingardi M, Ghirarduzzi A, Lo Scocco G, Macchioni P, Bajocchi G, et al. Is duplex ultrasonography useful for the diagnosis of giant-cell arteritis? *Ann Intern Med* 2002;137:232-8.
17. Boyev LR, Miller NR, Green WR. Efficacy of unilateral versus bilateral temporal artery biopsies for the diagnosis of giant cell arteritis. *Am J Ophthalmol* 1999;128:211-5.
18. Chakrabarty A, Franks AJ. Temporal artery biopsy: is there any value in examining biopsies at multiple levels? *J Clin Pathol* 2000;53:131-6.
19. Davies C, Frost B, Eshan O, McLain AD, Shandall A. Temporal artery biopsy...who needs one? *Postgrad Med J* 2006;82:476-8.
20. Mahr A, Saba M, Kambouchner M, Polivka M, Baudrimont M, Brocheriou I, et al. Temporal artery biopsy for diagnosing giant cell arteritis: the longer, the better? *Ann Rheum Dis* 2006;65:826-8.
21. Diaz FG, Chason J, Shrontz C, Ausman JJ, Dujovny M. Histological structural abnormalities of superficial temporal arteries used for extracranial-intracranial anastomosis. *J Neurosurg* 1982;57:328-33.
22. Sox HC. The evaluation of diagnostic tests: principles, problems, and new developments. *Annu Rev Med* 1996;47:463-71.