

Assessing Vasculitis in Giant Cell Arteritis by Ultrasound: Results of OMERACT Patient-based Reliability Exercises

Valentin S. Schäfer¹, Stavros Chrysidis, Christian Dejaco, Christina Duftner, Annamaria Iagnocco, George A. Bruyn, Greta Carrara, Maria Antonietta D'Agostino, Eugenio De Miguel, Andreas P. Diamantopoulos, Ulrich Fredberg, Wolfgang Hartung, Alojzija Hocevar, Aaron Juche, Tanaz A. Kermani, Matthew J. Koster, Tove Lorenzen, Pierluigi Macchioni, Marcin Milchert, Uffe Møller Døhn, Chetan Mukhtyar, Cristina Ponte, Sofia Ramiro, Carlo A. Scirè, Lene Terslev, Kenneth J. Warrington, Bhaskar Dasgupta, and Wolfgang A. Schmidt

ABSTRACT. Objective. To test the reliability of Outcome Measures in Rheumatology Clinical Trials (OMERACT) consensus-based ultrasound definitions for normal and vasculitic temporal and axillary arteries in patients with giant cell arteritis (GCA) and in controls.

Methods. A preliminary 1-day meeting and a full 3-day meeting fulfilling OMERACT Ultrasound Group guidelines were held. Temporal and axillary arteries were examined at 2 timepoints by 12 sonographers on 4 patients with GCA and 2 controls. The aim was to test inter- and intrareader reliability for normal findings, halo sign, and compression sign. In both meetings, patients had established GCA. Pathology was more recent in the full meeting, which was preceded by 6 h of training. Scanning time was 15–20 min instead of 10–13 min.

Results. In the preliminary exercise, interreader reliabilities were fair to moderate for the overall diagnosis of GCA (Light κ 0.29–0.51), and poor to fair for identifying vasculitis in the respective anatomical segments (Light κ 0.02–0.46). Intrareader reliabilities were moderate (Cohen κ 0.32–0.64). In the main exercise, interreader reliability was good to excellent (Light κ 0.76–0.86) for the overall diagnosis of GCA, and moderate to good (Light κ 0.46–0.71) for identifying vasculitis in the respective anatomical segments. Intrareader reliability was excellent for diagnosis of GCA (Cohen κ 0.91) and good (Cohen κ 0.71–0.80) for the anatomical segments.

Conclusion. OMERACT-derived definitions of halo and compression signs of temporal and axillary arteries are reliable in recent-onset GCA if experienced sonographers (> 300 examinations) have 15–20 min for a standardized examination with prior training and apply > 15 MHz probes. (J Rheumatol First Release July 1 2018; doi:10.3899/jrheum.171428)

Key Indexing Terms:

GIANT CELL ARTERITIS ULTRASOUND RELIABILITY DIAGNOSIS VASCULITIS

From the University Hospital Bonn III, Medical Clinic, Department of Oncology, Hematology and Rheumatology, Bonn; Immanuel Krankenhaus Berlin, Medical Center for Rheumatology Berlin-Buch, Berlin; Asklepios Medical Center, Bad Abbach, Germany; Hospital of Southwest Denmark, Esbjerg; Diagnostic Centre Region Hospital Silkeborg, Silkeborg; Odense University Hospital, Odense; Copenhagen Center for Arthritis Research (COPECARE), Glostrup, Denmark; Medical University Graz, Graz; Medical University Innsbruck, Innsbruck, Austria; Hospital of Bruneck, Bruneck; Università degli Studi di Torino, Turin; Epidemiology Unit – Italian Society for Rheumatology (SIR), Milan; Arcispedale Santa Maria Nuova, Reggio Emilia; University of Ferrara, Italy; MC Groep Hospitals, Lelystad; Leiden University Medical Center, the Netherlands; Hôpital Ambroise Paré, Boulogne-Billancourt, France; University Hospital La Paz, Madrid, Spain; Martina Hansens Hospital, Bærum, Oslo, Norway; University Medical Centre Ljubljana, Ljubljana, Slovenia; University of California at Los Angeles, Los Angeles, California; Mayo Clinic, Rochester, Minnesota, USA; Pomeranian Medical University, Szczecin, Poland; Norfolk and Norwich University Hospital, Norwich; Southend University Hospital, UK National Health Service (NHS) Foundation Trust

and Anglia Ruskin University, Westcliff, UK; Rheumatology Research Unit, Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon; Rheumatology Department, Hospital de Santa Maria - CHLN, Lisbon, Portugal.

The preliminary exercise was funded in connection with the International Symposium on Giant Cell Arteritis, Polymyalgia Rheumatica and Large Vessel Vasculitis in Southend, United Kingdom. The full exercise was funded by a grant from Roche Pharma Germany. Ultrasound equipment was provided for both exercises by Esaote SpA, Genoa, Italy.

V.S. Schäfer, MD, University Hospital Bonn III, Medical Clinic, Department of Oncology, Hematology and Rheumatology, and Immanuel Krankenhaus Berlin, Medical Center for Rheumatology Berlin-Buch; S. Chrysidis, MD, Hospital of Southwest Denmark; C. Dejaco, MD, PhD, Medical University Graz, and Hospital of Bruneck; C. Duftner, MD, PhD, Medical University Innsbruck; A. Iagnocco, MD, Università degli Studi di Torino; G.A. Bruyn, MD, PhD, MC Groep Hospitals; G. Carrara, MSc, Epidemiology Unit – SIR; M.A. D'Agostino, MD, PhD, Hôpital Ambroise Paré; E. De Miguel, MD, University Hospital La Paz; A.P. Diamantopoulos, MD, PhD, MPH, Martina Hansens Hospital;

Early and accurate diagnosis of giant cell arteritis (GCA) is imperative. Failure to accurately diagnose and expeditiously treat GCA may lead to vision loss and other severe ischemic complications, whereas misdiagnosis of non-GCA pathology as GCA leads to inappropriate glucocorticoid use and toxicity. Temporal artery biopsy (TAB) has been the diagnostic test of choice. However, TAB is invasive, and results are not immediately available. Hence it is increasingly being replaced by imaging, which includes ultrasound (US), magnetic resonance imaging (MRI), computed tomography (CT), and 18F-fluorodeoxyglucose positron emission tomography (FDG-PET)¹. FDG-PET and CT facilitate the examination of extracranial arteries to confirm the diagnosis of large-vessel GCA and exclude alternative serious pathology. MRI and particularly US can additionally visualize temporal arteries and other superficial cranial arteries.

US is widely available in rheumatology practice. It is patient-friendly, reproducible, and repeatable. Modern US transducers achieve image resolution of 0.1 mm for superficial arteries, which is higher than that of other imaging techniques¹. US displays a noncompressible, hypoechoic, most commonly concentric arterial wall thickening (“halo sign”) in acute GCA^{2,3}. Alongside medical history and clinical examination it can be used in fast-track clinics offering appointments for patients within 24 h, to rapidly confirm or exclude the diagnosis of suspected GCA. Two studies have shown a decrease of permanent irreversible vision loss after inauguration of fast-track clinics^{4,5}.

US in all patients with suspected GCA is cost-effective compared to biopsy plus clinical judgment without imaging⁶. It has a higher sensitivity than TAB regarding the clinical diagnosis, particularly in patients with large-vessel GCA^{6,7}. Several studies have investigated the accuracy, and construct and criterion validity of US in the diagnosis of GCA,

including 3 metaanalyses^{8,9,10}. There is a trend to higher sensitivities in newer studies because of better technology and increasing experience. A new metaanalysis including studies until February 2017 revealed a pooled sensitivity of 77% and a pooled specificity of 96% with a positive likelihood ratio of 19 and a negative likelihood ratio of 0.2 for the halo sign in temporal arteries compared to the clinical diagnosis of GCA¹¹.

Nevertheless, issues have been raised regarding the diagnostic performance and reliability of US, thus challenging its overall usefulness in GCA. A recent phase III trial enrolled 37% of its patients on the basis of cross-sectional imaging, although US was not included¹². Another phase III trial (ClinicalTrials.gov identifier: NCT02531633) included US as an eligible diagnostic modality. This trial was however prematurely terminated in October 2017 by the sponsor, based on the decision to discontinue development of sirukumab in autoimmune diseases. Recently published European League Against Rheumatism recommendations on imaging in large-vessel vasculitis suggest US as the first imaging modality particularly in patients with suspected predominantly cranial GCA¹³.

An Outcome Measures in Rheumatology Clinical Trials (OMERACT) US subgroup on large-vessel vasculitis was formed. A Delphi survey based on a systematic literature search arrived at US definitions for normal temporal and axillary arteries, the halo sign, and the “compression sign.” These definitions were tested in a Web-based exercise on still images and videos of normal and vasculitic temporal and axillary arteries. The reliability was excellent, with inter-reader agreements of 91–99% and mean κ values of 0.83–0.98 for both interreader and intrareader reliability¹⁴.

The focus of our study, described herein, is the OMERACT validation process, which tested the inter- and intrareader reliability of the aforementioned definitions for both normal and vasculitic arteries. The real-time patient-based exercises required simultaneous data acquisition and interpretation.

MATERIALS AND METHODS

Study design and setting. A preliminary 1-day meeting was held following the International Symposium on Giant Cell Arteritis, Polymyalgia Rheumatica and Large Vessel Vasculitis in Southend, UK, in March 2016 to test the feasibility and study setting for a patient-based exercise. Lessons learned were implemented in a definitive 3-day exercise in Berlin, Germany, in February 2017, modeled on previous OMERACT Ultrasound Working Group studies for testing patient-based reliability of US in rheumatic diseases^{15,16,17}. The methodology and reporting of the Berlin OMERACT reliability exercise adhered to the recommendations from the Enhancing the Quality and Transparency of Health Research network¹⁸ using the Guidelines for Reporting Reliability and Agreement Studies Statement¹⁹.

US examination. At each meeting, 12 sonographers individually examined 6 study subjects. All sonographers were previously involved in the development of the consensus-based US definitions. Each sonographer performed bilateral examinations of the common superficial temporal artery, its frontal and parietal branches and of the axillary arteries (i.e., 8 artery segments per patient) in longitudinal and transverse scans applying a binary score for

U. Fredberg, MD, PhD, Diagnostic Centre Region Hospital Silkeborg, and Odense University Hospital; W. Hartung, MD, Asklepios Medical Center; A. Hocevar, MD, MSc, University Medical Centre Ljubljana; A. Juche, MD, Immanuel Krankenhaus Berlin, Medical Center for Rheumatology Berlin-Buch; T.A. Kermani, MD, University of California at Los Angeles; M.J. Koster, MD, Mayo Clinic; T. Lorenzen, MD, Diagnostic Centre Region Hospital Silkeborg; P. Macchioni, MD, Arcispedale Santa Maria Nuova; M. Milchert, MD, PhD, Pomeranian Medical University; U.M. Døhn, MD, PhD, COPECARE; C. Mukhtyar, MBBS, MSc, MD, Norfolk and Norwich University Hospital; C. Ponte, MD, Rheumatology Research Unit, Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon, and Rheumatology Department, Hospital de Santa Maria – CHLN; S. Ramiro, MD, MSc, PhD, Leiden University Medical Center; C.A. Scirè, MD, PhD, Epidemiology Unit – SIR, and University of Ferrara; L. Terslev, MD, PhD, COPECARE; K.J. Warrington, MD, Mayo Clinic; B. Dasgupta, MBBS, MD, FRCP, Southend University Hospital, NHS Foundation Trust and Anglia Ruskin University; W.A. Schmidt, MD, Immanuel Krankenhaus Berlin, Medical Center for Rheumatology Berlin-Buch.

Address correspondence to Dr. V.S. Schäfer, University Hospital Bonn, Department of Rheumatology and Clinical Immunology, Medical Clinic and Policlinic III, Sigmund-Freud-Str. 25, Bonn 53105, Germany.
E-mail: valentin.s.schaefer@hotmail.de

Accepted for publication March 29, 2018.

vasculitis US lesions as defined by OMERACT¹⁴. The subject was lying on an examination couch in supine position. The head was rotated slightly toward the examiner for examining the left temporal artery and away from the examiner for examining the right temporal artery. The probe was placed in the axilla for examining the axillary artery. After a predetermined time, sonographers rotated to the next station until every sonographer examined all patients/controls. The data were collected immediately after each examination to exclude communication between sonographers. Sonographers were blinded to the study subjects' diagnosis. They were not allowed to communicate with the patients about signs or symptoms of the disease. None of the examined patients had visibly swollen temporal arteries. An identical examination sequence was repeated later the same day to assess intrareader reliability.

US equipment and settings. Esaote MyLab Twice/Class systems equipped with 6–18 MHz linear array transducers were used in the exercises. In the Berlin meeting, 2 additional Esaote MyLab 8 machines were used. The following settings were applied for the examination of the temporal arteries (axillary arteries): B-mode frequency 18 MHz (14 MHz), image depth 1.5 cm (3 cm), 1 focus point at 0.5 cm (1.5 cm) below skin surface, color Doppler frequency 9 MHz (6 MHz), and pulse repetition frequency 2.5 KHz (3.5 KHz). Sonographers were advised not to change these predefined settings except for adjusting image depth and focus point position for the examination of the axillary arteries, if necessary.

Preliminary meeting. The sonographers received no training on US machines and settings before the exercise. Thirteen minutes were allocated for scanning and scoring the findings for the first round and 10 min for the second round. The limit was set after a discussion about daily clinical practice conditions, where these time frames seemed to be adequate and realistic.

The examined study subjects were chosen by the convenors (BD, WAS), who did not participate in the reliability exercise. WAS, being unblinded to the history and diagnosis of all study subjects and having performed > 5000 scans in suspected GCA over 23 years, examined all study subjects in addition to the other sonographers (independent sonographer) to decide whether arterial segments were exhibiting clear or ambivalent pathology and to store reference images and videos. The examined study subjects were 63–76 years old (mean age 68 yrs). Four of them were females. Four study subjects had GCA consistent with the revised inclusion criteria of the SIRRESTA trial (NCT02531633). These criteria require age \geq 50 years, erythrocyte sedimentation rate \geq 50 mm/h, and/or C-reactive protein \geq 2.45 mg/dl (24.5 mg/l), unequivocal cranial symptoms of GCA and/or polymyalgia rheumatica, and evidence of large-vessel vasculitis by cross-sectional imaging including US if diagnosis is not confirmed histologically. Further, the diagnosis had remained unchanged until the exercise. By the time of the exercise, patients were receiving glucocorticoids for 5 weeks, 2 years, 2 years, and 6 years. One of the 2 controls had an uncommon finding of arteriosclerosis of both axillary arteries.

All sonographers were rheumatologists except one who was in his last year of rheumatology training. Prior to the exercise, 7 sonographers had performed > 300 scans of temporal and axillary arteries before, 2 had performed 101–300 scans, 2 had performed 51–100 scans, and 1 had performed < 20 scans. Five sonographers used US machine types in their institutions similar in manufacturer and price level to the ones used in the exercise.

Full meeting. The meeting included 6 h of practical US training on healthy individuals and patients with GCA, different from those who participated in the exercise, using the machines and settings used in the exercise. In the exercise, 20 min were allocated for scanning and scoring the findings for the first round and 15 min for the second round.

The examined study subjects were chosen and also examined by the convenor (WAS), who did not participate in the reliability exercise. Subjects' age ranged from 56 to 80 years (mean 68 yrs). Four of them were females. Four study subjects had GCA fulfilling the above-mentioned inclusion criteria. They had been receiving glucocorticoids for 4, 7, and 8 months. The

fourth patient had a persistent halo sign of temporal arteries for 4 years despite discontinuation of glucocorticoids. Two controls never had any signs or symptoms of GCA.

Eight of the 12 sonographers had participated in the preliminary exercise. All sonographers were rheumatologists. Eleven sonographers had performed > 300 scans of temporal and axillary arteries before. Two of them had indicated an experience of 101–300 scans at the time of the preliminary meeting. One sonographer had performed 51–100 scans at the time of each meeting. Six sonographers used US machines in their institutions similar in manufacturer and price level to the ones used in the exercise.

Ethics committee approval was obtained from the Berlin Medical Association (Berliner Ärztekammer, Eth-04-17). All patients provided written informed consent prior to participation in our study.

Definitions. The definitions obtained by the Delphi exercise and applied at the Web-based reliability exercise¹⁴ were also used in the patient-based reliability exercises:

- Normal temporal artery: Pulsating, compressible artery with anechoic lumen surrounded by mid- to hyperechoic tissue. Using US equipment with high resolution, the intima-media complex (IMC) presents as a homogeneous, hypo-, or anechoic echo structure delineated by 2 parallel hyperechoic margins (double-line pattern) may be visible.
- Normal axillary artery: Pulsating, hardly compressible artery with anechoic lumen; the IMC presents as a homogeneous, hypoechoic, or anechoic echo structure delineated by 2 parallel hyperechoic margins (double-line pattern), which is surrounded by mid- to hyperechoic tissue.
- Halo sign: Homogeneous, hypoechoic wall thickening, well delineated toward the luminal side, visible both in longitudinal and transverse planes, most commonly concentric in transverse scans.
- Compression sign: The thickened arterial wall remains visible upon compression; the hypoechogenic vasculitic vessel wall thickening contrasts with the mid-echogenic to hyperechogenic surrounding tissue.

Figures explaining these definitions can be found in the article describing the Delphi process in more detail¹⁴ and in another review article²⁰.

Statistical analysis. All sonographers (1–12) evaluated all study subjects ($n = 6$) in 2 rounds, in a total of 8 anatomical positions (common temporal artery, parietal branch, frontal branch, and axillary artery), taking both sides of the body (right, left) into account. Intra- and interreader reliabilities were calculated using the kappa coefficient (κ). Intrareader reliability was assessed by Cohen κ . Interreader reliability was studied by calculating the mean κ on all pairs (i.e., Light κ)²¹. Kappa coefficients and the corresponding 95% CI were interpreted according to Landis and Koch: κ values of 0–0.2 were considered poor, 0.2–0.4 fair, 0.4–0.6 moderate, 0.6–0.8 good, and 0.8–1 excellent²². The percentage of observed agreement (i.e., percentage of observations that obtained the same score), prevalence of the observed lesions, and prevalence-adjusted bias-adjusted κ (PABAK) were also calculated^{23,24}. Analyses were performed using R Statistical Software (Foundation for Statistical Computing).

RESULTS

Preliminary meeting. The mean interrater agreement for the overall diagnosis of GCA was 0.73 in round 1 and 0.83 in round 2. It was 0.79 in round 1 and 0.77 in round 2 for identifying vasculitis in the respective anatomical segments. The mean intrarater agreements were 0.82 (0.50–1) for the overall diagnosis of GCA and 0.84 (range 0.58–1) for identifying vasculitis in the respective anatomical segments.

The mean interrater reliabilities were fair to moderate for the overall diagnosis of GCA (Light κ 0.29–0.51) and poor to fair for identifying vasculitis in the respective anatomical segments (Light κ 0.02–0.46). Mean intrarater reliabilities were moderate (Cohen κ 0.32–0.64).

The independent sonographer rated 21 of 36 temporal

Table 1. Interrater reliability and agreement in the full exercise (Round 1).

Variables	Mean Prevalence, %	Mean Agreement	Mean κ	Mean PABAK
US positive for GCA	61.1	0.88	0.76	0.77
Halo sign				
Temporal arteries, all segments	31.3	0.77	0.47	0.55
Common superficial temporal artery	34.7	0.78	0.49	0.55
Frontal branch	31.2	0.77	0.46	0.54
Parietal branch	27.8	0.78	0.46	0.56
Axillary arteries	52.4	0.83	0.66	0.66
Compression sign				
Temporal arteries, all segments	32.6	0.78	0.49	0.55
Common superficial temporal artery	35.4	0.78	0.51	0.57
Frontal branch	33.3	0.75	0.44	0.49
Parietal branch	29.2	0.80	0.53	0.60

PABAK: prevalence-adjusted bias-adjusted kappa; GCA: giant cell arteritis; US: ultrasound.

Table 2. Interrater reliability and agreement in the full exercise (Round 2).

Variables	Mean Prevalence, %	Mean Agreement	Mean κ	Mean PABAK
US positive for GCA	62.5	0.93	0.86	0.86
Halo sign				
Temporal arteries, all segments	33.3	0.82	0.60	0.65
Common superficial temporal artery	41.0	0.86	0.71	0.72
Frontal branch	31.3	0.79	0.50	0.58
Parietal branch	27.8	0.82	0.54	0.65
Axillary arteries	52.8	0.81	0.64	0.63
Compression sign				
Temporal arteries, all segments	33.8	0.83	0.60	0.65
Common superficial temporal artery	38.2	0.83	0.63	0.65
Frontal branch	34.0	0.81	0.57	0.62
Parietal branch	29.2	0.84	0.60	0.68

GCA: giant cell arteritis; PABAK: prevalence-adjusted bias-adjusted kappa; US: ultrasound.

artery segments (58%) as ambivalent because of minor pathology, such as very small halo size of about < 0.5 mm and incomplete compressibility in some subsegments because of chronic changes in longstanding disease (Figure 1). He considered 4 of 12 axillary arteries (33%) ambivalent including both axillary arteries of 1 control with unusually pronounced arteriosclerosis showing heterogeneous and in part hyperechoic, irregularly delineated, eccentric vessel wall alteration with a diameter of up to 1.7 mm. Only 3 experienced sonographers (> 300 scans) considered the findings in these patients non-GCA in both rounds. There were 65% of

sonographers who felt that unfamiliarity with the equipment might have hampered their results of false-positive or negative diagnosis and of intrareader reliability.

Full meeting. The mean interrater agreement for the overall diagnosis of GCA was 0.88 in round 1 and 0.93 in round 2. It was 0.78 (range 0.75–0.83) in round 1 and 0.82 (range 0.79–0.86) in round 2 for identifying vasculitis in the respective anatomical segments. The mean intrarater agreements were 0.96 (range 0.83–1) for the overall diagnosis of GCA and 0.89 (range 0.58–1) for identifying vasculitis in the respective anatomical segments.

The interrater reliability was good to excellent. The mean Light κ was 0.76 in round 1 and 0.86 in round 2 for the overall diagnosis of GCA. The mean PABAK was 0.77 and 0.86 in rounds 1 and 2, respectively. For identifying vasculitis in the respective anatomical segments, the reliability was moderate for the temporal arteries (mean κ 0.46–0.53, mean PABAK 0.49–0.66) in round 1, moderate to good in round 2 (mean κ 0.5–0.71, mean PABAK 0.58–0.72), and moderate for the axillary arteries in both rounds (mean κ 0.64–0.66). The intrareader reliability was excellent for the diagnosis of GCA (Cohen κ 0.91, PABAK 0.92) and good (Cohen κ 0.71–0.80, PABAK 0.73–0.81) for the respective anatomical segments.

The independent sonographer rated 14 of 36 temporal artery segments (39%) and none of the 12 axillary arteries as ambivalent due to minor pathology because of chronic changes in longstanding disease. All sonographers agreed in both rounds that the controls had no GCA. Agreement was also 100% in both rounds for the diagnosis of GCA in 3 patients with GCA. Disagreement occurred only when 5/12 and 3/12 sonographers missed the diagnosis of GCA in rounds 1 and 2, respectively, in 1 obese patient with bilateral axillary artery vasculitis, very small residual artery lumen, pronounced collateral flow, and normal temporal arteries (Figure 2).

In both exercises, reliabilities did not significantly differ whether halo sign or compression sign was evaluated. The detailed results are shown in Table 1, Table 2, and Table 3.

DISCUSSION

The inter- and intrarater reliabilities for performing US of temporal and axillary arteries in patients with GCA and controls were good to excellent for the diagnosis of GCA with experienced sonographers who were familiar with the US equipment.

Better reliabilities attained in the full exercise compared to the preliminary exercise could be explained by the following:

(1) Lack of sonographer training on the US equipment and its settings in the preliminary exercise. Only 42% of sonographers in the preliminary exercise and 50% in the full exercise were using similar equipment in their institutions. Even if a sonographer is familiar with a certain type of

Table 3. Intrarater reliability and agreement in the full exercise.

Variables	Mean Prevalence, %	Mean Agreement	Mean κ	Mean PABAK	Range of Agreement	Range, κ	Range, PABAK
US positive for GCA	61.8	0.96	0.91	0.92	0.83–1	0.57–1	0.7–1
Halo sign							
Temporal arteries, all segments	32.3	0.88	0.71	0.76	0.75–1	0.31–1	0.5–1
Common superficial temporal artery	37.9	0.87	0.71	0.73	0.67–1	0.25–1	0.3–1
Frontal branch	31.2	0.89	0.73	0.77	0.67–1	0.27–1	0.3–1
Parietal branch	27.8	0.89	0.71	0.78	0.67–1	0.25–1	0.3–1
Axillary arteries	52.4	0.90	0.80	0.81	0.73–1	0.46–1	0.5–1
Compression sign							
Temporal arteries, all segments	33.2	0.89	0.73	0.78	0.75–1	0.34–1	0.5–1
Common superficial temporal artery	36.8	0.89	0.75	0.77	0.67–1	0.25–1	0.3–1
Frontal branch	33.7	0.88	0.74	0.75	0.58–1	0.17–1	0.2–1
Parietal branch	29.2	0.89	0.72	0.77	0.67–1	0.25–1	0.3–1

GCA: giant cell arteritis; PABAK: prevalence-adjusted bias-adjusted kappa; US: ultrasound.

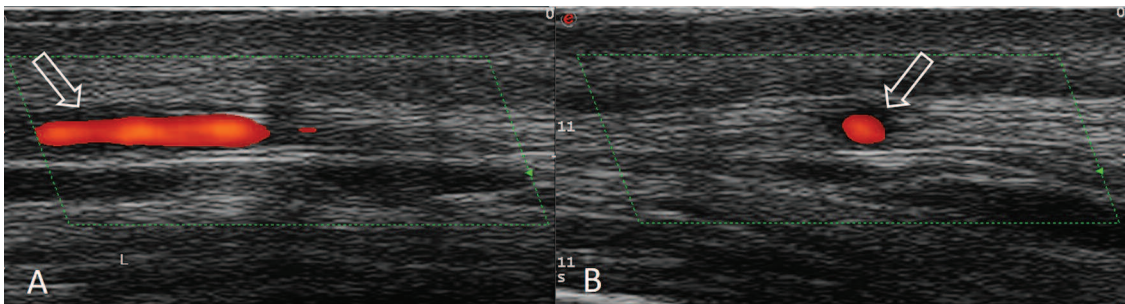


Figure 1. Small segmental, only slightly hypoechoic halo of a temporal artery branch in a patient of the full exercise with longstanding GCA. A. Longitudinal view. B. Transverse view. GCA: giant cell arteritis.

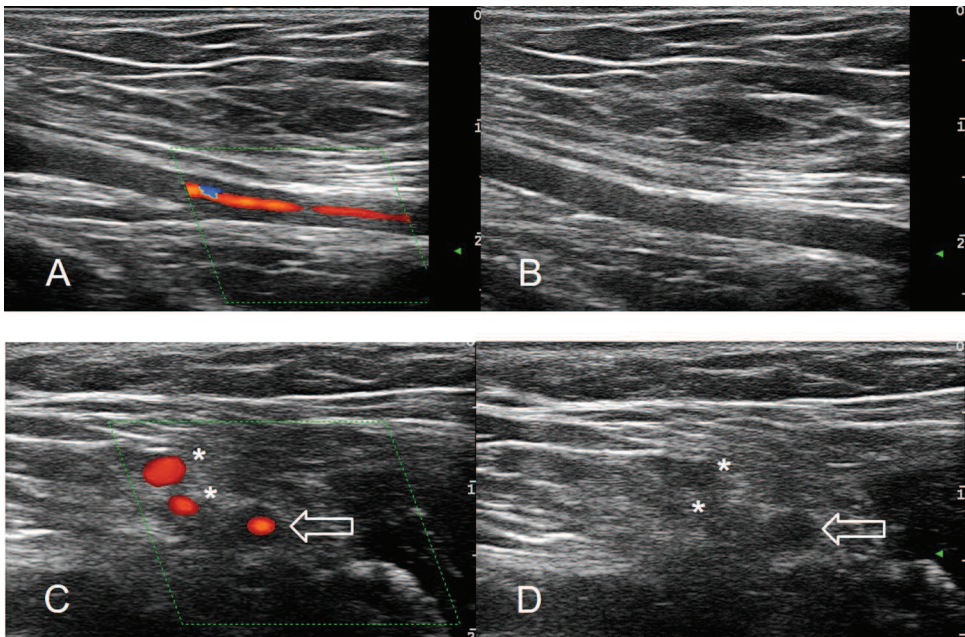


Figure 2. Halo sign of an axillary artery of a patient with GCA in the full exercise in longitudinal (A, B) and transverse (C, D) views. It is only visible when applying color Doppler (A, C). In the transverse views, collateral vessels without halo sign (*) appear larger than the lumen of the affected axillary (arrows). GCA: giant cell arteritis.

machine, experience with the settings is important as these may considerably influence the appearance of the US images. (2) Only 58% of sonographers in the preliminary exercise had performed > 300 examinations in suspected GCA compared to 92% in the full exercise. The European Federation of Societies for Ultrasound in Medicine and Biology minimum training requirements for rheumatologists performing musculoskeletal US demand a minimum of 300 US examinations for achieving level I competency²⁵. Our current study suggests that this requirement may also apply for temporal and axillary artery US in suspected GCA.

(3) More time was provided for each examination in the full exercise because 67% of sonographers of the preliminary exercise said they felt that time restrictions had hampered the results. An examination time of 15–20 min appears to be optimal for examining temporal and axillary arteries in suspected GCA.

(4) The time frame when performing US is important for image interpretation. In patients with untreated GCA, the pathology is much more pronounced than in patients with longstanding, treated disease. The real-time patient-based reliability exercises, according to an OMERACT algorithm, are faced with this shortcoming, because it is impossible to obtain patients with untreated GCA for these exercises. The disease was more longstanding and pathologies were subtler in the preliminary exercise, with 52% of examined anatomical segments showing ambivalent findings compared to 29% in the full exercise. The sensitivity of temporal artery US decreases rapidly with glucocorticoid treatment. In 1 study, the sensitivity compared to the final clinical diagnosis dropped from 88% in patients who had been untreated or who had received glucocorticoids for not longer than 1 day, to 50% in patients who had been treated for 2 days or longer²⁶. Another study, however, found that a residual halo sign may persist for 8 weeks in half of the patients²⁷. In axillary arteries, US pathology may remain longer, for months and years, but it also decreases over time⁷. Nevertheless, as halo size decreases and halo echogenicity increases with treatment, it is more difficult to differentiate normal from abnormal findings in treated established GCA. This is probably also the case for histology because giant cells do not persist longer than 6 months²⁸. Arteriosclerosis may be a potential confounder in the mainly elderly GCA population. It is, however, far less common in the temporal and axillary arteries than in the carotid and femoral arteries.

Few studies have yet assessed real-time patient-based reliabilities for US in suspected GCA. As for other indications and other imaging methods, reliability was higher when investigated for only 2 sonographers from the same institution. Agreement of 2 sonographers examining temporal arteries for halo sign, stenosis, and occlusions was 95% for the diagnosis of GCA in 1 study². In another study, 2 sonographers evaluating the compression sign of temporal arteries disagreed only in 1 of 60 patients²⁹. A single study with multiple sonogra-

phers from Spain found excellent reliability with a κ value of 0.85 for interreader reliability and of 0.95 for intrareader reliability after a training workshop³⁰. The reliability in our study may be lower probably because of a tighter protocol.

Our study has limitations. The reliability may depend on the severity of the pathologic findings. Because all patients were receiving glucocorticoid treatment, reliability may have been impaired by ambivalent pathology. The repetition of the examination sequence on the same day may have led to overestimation of intrareader reliability. Although similar US equipment was used, even machines of the same type may exhibit different image features. Our study was performed with current high-quality modern 6–18 MHz probes. Probes for examining temporal arteries should provide frequencies of ≥ 15 MHz³¹. Probes with frequencies > 20 MHz will further increase resolution and allow reliable measurement of the IMC of temporal arteries³². Very few of the sonographers participating in our study are using these probes. Further, IMC measurement of axillary arteries could have a role in future US protocols in suspected GCA.

These exercises following the OMERACT Ultrasound Group guidelines show that the OMERACT-derived definitions of halo and compression signs of temporal and axillary arteries are applicable in recent-onset GCA with excellent inter- and intrareader reliabilities for the diagnosis of GCA if sonographers are experienced, are provided sufficient time for examination, and are familiar with the US equipment, high frequency probes > 15 MHz, and settings.

ACKNOWLEDGMENT

The authors are indebted to the patients for their participation in the study. We also thank Katerina Achilleos and Kenny Schlüter, who helped organize the workshops in Southend and Berlin, respectively.

REFERENCES

1. Muratore F, Pipitone N, Salvarani C, Schmidt WA. Imaging of vasculitis: state of the art. *Best Pract Res Clin Rheumatol* 2016;30:688-706.
2. Schmidt WA, Kraft HE, Vorpahl K, Völker L, Gromnica-Ihle EJ. Color duplex ultrasonography in the diagnosis of temporal arteritis. *N Engl J Med* 1997;337:1336-42.
3. Aschwanden M, Daikeler T, Kesten F, Baldi T, Benz D, Tyndall A, et al. Temporal artery compression sign—a novel ultrasound finding for the diagnosis of giant cell arteritis. *Ultraschall Med* 2013; 34:47-50.
4. Diamantopoulos AP, Haugeberg G, Lindland A, Myklebust G. The fast-track ultrasound clinic for early diagnosis of giant cell arteritis significantly reduces permanent visual impairment: towards a more effective strategy to improve clinical outcome in giant cell arteritis? *Rheumatology* 2016;55:66-70.
5. Patil P, Williams M, Maw WW, Achilleos K, Elsideeg S, Dejaco C, et al. Fast track pathway reduces sight loss in giant cell arteritis: results of a longitudinal observational cohort study. *Clin Exp Rheumatol* 2015;33:S-103-6.
6. Luqmani R, Lee E, Singh S, Gillett M, Schmidt WA, Bradburn M, et al. The role of ultrasound compared to biopsy of temporal arteries in the diagnosis and treatment of giant cell arteritis (TABUL): a diagnostic accuracy and cost-effectiveness study. *Health Technol Assess* 2016;20:1-238.

7. Schmidt WA, Seifert A, Gromnica-Ihle E, Krause A, Natusch A. Ultrasound of proximal upper extremity arteries to increase the diagnostic yield in large-vessel giant cell arteritis. *Rheumatology* 2008;47:96-101.
8. 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.
9. Ball EL, Walsh SR, Tang TY, Gohil R, Clarke JM. Role of ultrasonography in the diagnosis of temporal arteritis. *Br J Surg* 2010;97:1765-71.
10. Arida A, Kyprianou M, Kanakis M, Sfrikakis PP. The diagnostic value of ultrasonography-derived edema of the temporal artery wall in giant cell arteritis: a second meta-analysis. *BMC Musculoskelet Disord* 2010;11:44.
11. Duftner C, Dejaco C, Sepriano A, Falzon L, Schmidt WA, Ramiro S. Imaging in diagnosis, outcome prediction and monitoring of large vessel vasculitis: A systematic literature review and meta-analysis informing the EULAR recommendations. *RMD Open* 2018;4:e000612.
12. Tuckwell K, Collinson N, Dimonaco S, Klearman M, Blockmans D, Brouwer E, et al; GIACTA Investigators. Newly diagnosed vs. relapsing giant cell arteritis: Baseline data from the GIACTA trial. *Semin Arthritis Rheum* 2017;46:657-64.
13. Dejaco C, Ramiro S, Duftner C, Besson FL, Bley TA, Blockmans D, et al. EULAR recommendations for the use of imaging in large vessel vasculitis in clinical practice. *Ann Rheum Dis* 2018; 77:636-43.
14. Chrysidis S, Duftner C, Dejaco C, Schäfer VS, Ramiro S, Carrera G, et al. Definitions and reliability assessment of elementary ultrasound lesions in giant cell arteritis: a study from the OMERACT Large Vessel Vasculitis Ultrasound Working Group. *RMD Open* 2018;4: e000598.
15. Terslev L, Gutierrez M, Christensen R, Balint PV, Bruyn GA, Delle Sedie A, et al; OMERACT US Gout Task Force. Assessing elementary lesions in gout by ultrasound: results of an OMERACT patient-based agreement and reliability exercise. *J Rheumatol* 2015;42:2149-54.
16. Terslev L, Naredo E, Aegerter P, Wakefield RJ, Backhaus M, Balint P, et al. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce-Part 2: reliability and application to multiple joints of a standardised consensus-based scoring system. *RMD Open* 2017;3:e000427.
17. Filippou G, Scire CA, Damjanov N, Adinolfi A, Carrara G, Picerno V, et al. Definition and reliability assessment of elementary ultrasonographic findings in calcium pyrophosphate deposition disease: a study by the OMERACT Calcium Pyrophosphate Deposition Disease Ultrasound Subtask Force. *J Rheumatol* 2017;44:1744-9.
18. Christensen R, Bliddal H, Henriksen M. Enhancing the reporting and transparency of rheumatology research: a guide to reporting guidelines. *Arthritis Res Ther* 2013;15:109.
19. Kottner J, Gajewski BJ, Streiner DL. Guidelines for Reporting Reliability and Agreement Studies (GRRAS). *Int J Nurs Stud* 2011;48:659-60.
20. Schmidt WA. Ultrasound in the diagnosis and management of giant cell arteritis. *Rheumatology* 2018;57:ii22-31.
21. Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol Bull* 1971;76:365-77.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
23. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993;46:423-9.
24. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res Social Adm Pharm* 2013;9:330-8.
25. Terslev L, Hammer HB, Torp-Pedersen S, Szkudlarek M, Iagnocco A, D'Agostino MA, et al. EFSUMB minimum training requirements for rheumatologists performing musculoskeletal ultrasound. *Ultraschall Med* 2013;34:475-7.
26. Hauenstein C, Reinhard M, Geiger J, Markl M, Hetzel A, Treszl A, et al. Effects of early corticosteroid treatment on magnetic resonance imaging and ultrasonography findings in giant cell arteritis. *Rheumatology* 2012;51:1999-2003.
27. De Miguel E, Roxo A, Castillo C, Peiteado D, Villalba A, Martin-Mola E. The utility and sensitivity of colour Doppler ultrasound in monitoring changes in giant cell arteritis. *Clin Exp Rheumatol* 2012;30:S34-8.
28. Maleszewski JJ, Younge BR, Fritzlen JT, Hunder GG, Goronzy JJ, Warrington KJ, et al. Clinical and pathological evolution of giant cell arteritis: a prospective study of follow-up temporal artery biopsies in 40 treated patients. *Mod Pathol* 2017;30:788-96.
29. Aschwanden M, Imfeld S, Staub D, Baldi T, Walker UA, Berger CT, et al. The ultrasound compression sign to diagnose temporal giant cell arteritis shows an excellent interobserver agreement. *Clin Exp Rheumatol* 2015;33:S113-5.
30. De Miguel E, Castillo C, Rodriguez A, De Agustin JJ; Working Group Ultrasound Giant Cell Arteritis. Learning and reliability of colour Doppler ultrasound in giant cell arteritis. *Clin Exp Rheumatol* 2012;27:S53-8.
31. Monti S, Floris A, Ponte C, Schmidt WA, Diamantopoulos AP, Pereira C, et al. The use of ultrasound to assess giant cell arteritis: review of the current evidence and practical guide for the rheumatologist. *Rheumatology* 2017;57:227-35.
32. Schäfer VS, Juche A, Ramiro S, Krause A, Schmidt WA. Ultrasound cut-off values for intima-media thickness of temporal, facial and axillary arteries in giant cell arteritis. *Rheumatology* 2017; 56:1479-83.