Learning and reliability of colour Doppler ultrasound in giant cell arteritis

E. De Miguel¹, C. Castillo¹, A. Rodríguez², J.J. De Agustín³
and the Working Group Ultrasound Giant Cell Arteritis

ABSTRACT

Objective. To evaluate a standardized training program and a reliability exercise in colour Doppler ultrasound (CDUS) for giant cell arteritis (GCA).

Methods. Two workshops were conducted in 2007 and 2008 to train rheumatologists in CDUS for GCA ultrasound diagnosis. Twenty-nine and forty-three participants without previous experience in GCA ultrasound were admitted in 2007 and 2008, respectively. First, some theoretical knowledge about GCA ultrasound signs was provided; second, a reader evaluation session of temporal artery video recording examinations of 27 and 30 patients were projected in the 2007 meeting and the 2008 workshop, respectively (50% were cases and 50% were controls). Twenty-four cases were common to both reader sessions. A mean of six videos had to be assessed as normal or pathologic. Finally, hands-on scanning training was performed.

To assess the efficacy of the workshop: 1) a structured satisfaction questionnaire was graded (1-5 Likert scale), and 2) the reliability, specificity, and percentage of correctly classified cases by each participant were calculated.

Results. The kappa coefficient of inter-reader agreement for the 29 and 43 participants was excellent (κ: 0.846) in 2007 and (κ: 0.848) in 2008. The intra-reader kappa result was also excellent (κ: 0.950). The satisfaction, sensitivity, specificity, and percentage of correctly classified patients and controls were very high.

Conclusions. The proposed learning method seemed to be effective and well accepted by the target audience. The inter-reader reliability of GCA ultrasound was excellent. These encouraging results support the need for planned standardized training programs.

Introduction

Giant cell arteritis (GCA) is the most common form of systemic inflammatory vasculitis in adults (1). Prompt diagnosis and treatment are important to prevent serious vascular complications, such as blindness and other vascular events (2). GCA is sometimes diagnosed clinically, but a temporal-artery biopsy is generally recommended to confirm the diagnosis. Using American College of Rheumatology (ACR) criteria, fulfillment of any three or more criteria indicates a diagnosis of temporal arteritis with 93.5% sensitivity and 91.2% specificity (3). However, the ACR criteria was designed for research purposes and not for clinical diagnosis. In fact, poor positive predictive value has been found when these criteria were used for clinical diagnosis (4). For this reason, most authorities recommend pathological confirmation in all patients due to the potentially serious side effects of long-term corticosteroid therapy. However, although a temporal artery biopsy is the diagnostic gold standard for this disease, biopsy results may be negative in 9% to 44% of patients with a clinical diagnosis of GCA (5-7). This has prompted the use of imaging techniques, such as colour Doppler ultrasonography (CDUS), and magnetic resonance imaging or positron emission tomography (8-13). There is growing evidence that CDUS of the temporal arteries delineates a characteristic hypoechoic circumferential wall thickening (halo) around the lumen of inflamed temporal arteries, stenoses, and occlusions (14-16). Outcome measures in clinical practice should be valid and reliable. Many studies and a meta-analysis (17) have demonstrated the validity of CDUS in GCA, but few papers have shown the reliability of this imaging technique, and the studies have always involved a limited number of readers. Today, few rheumatology units perform CDUS in the diagnosis of
The workshop was divided into three different parts: a) a 20’ theoretical session; followed by b) an 80’ reading session of temporal artery examination video recordings; and finally c), a 50’ training session dedicated to hands-on scanning to become familiar with normal temporal artery ultrasound examination.

Materials and methods
Two ultrasonography workshops were conducted at the Spanish Society of Rheumatology Annual Meeting in 2007 and 2008 to train rheumatologists in CDUS for GCA ultrasound diagnosis. The workshop was divided into three different parts: a) a 20’ theoretical session; followed by b) an 80’ reading session of temporal artery examination video recordings; and finally c), a 50’ training session dedicated to hands-on scanning to become familiar with normal temporal artery ultrasound examination.

Participants
Twenty-nine participants, all of whom were rheumatologists without previous experience in GCA ultrasound, were admitted to the 2007 workshop, and 43 were admitted to the 2008 workshop. Admission was voluntary, and the selection process based on the order of registration. The participants were asked for their collaboration to analyze the inter-reader reliability of an individual and blinded lecture and the future use of these data in a report. In the 2007 workshop, we used a sheet of paper indicating the number of every case and the answer, while in 2008, an electronic real-time wireless voting system was used after every case video collection.

Theoretical session
Prior to the reading session, a theoretical seminar about 20 minutes in length was carried out. In this part of the workshop, the typical ultrasound signs of GCA – halo, stenosis, or occlusion – were shown by digital video recordings of CDUS temporal artery examinations, including normal pathological temporal artery ultrasound examinations. After the theoretical knowledge instruction was finished, the reading session began.

Reading session
Temporal artery frontal and parietal branches video recording examinations from 27 and 30 patients were examined during the reading session in the 2007 and 2008 workshops, respectively. An average of six videos was shown for each patient; 13 patients that had been diagnosed with GCA and 14 controls were shown in 2007, while 15 diagnosed patients and 15 controls were included in 2008. Diagnosis was obtained based on the clinical histories of the patients at least three months after the ultrasound exploration.

The video recording evaluation process was independently performed, and each of the participants was blinded to the diagnoses of the patients. Every case was answered as normal or pathologic. As the workshop was educational, after every case was answered by the participants, the diagnosis was elucidated and was discussed in an interactive manner to resolve disputes and improve knowledge. Halo was the ultrasound sign assessed (Fig. 1). To avoid recognition, the video recordings used for the reading session were different from those shown at the theoretical session. Twenty-four cases were common to both reader sessions, and seven readers participated in both workshops; this was used in the intra-reader reliability.

The videos shown were captured by one experienced ultrasonographer with Acuson-Antares Siemens equipment using a linear scanner (5-13 MHz). Only three cases were captured with a Logic 5 (5-12 MHz) in the 2007 reader exercise, and another three were captured with a Logic 9 General Electric system (9-14 MHz probe) in 2008. Most of the videos were longitudinal scans of both the frontal and parietal sides of the superficial temporal arteries, although some videos were transverse scans.

Acquisition training session
Finally, hands-on scanning training was performed to familiarize participants with normal temporal artery ultrasound examinations. Experienced ultrasonographers conducted the participants’ examinations and taught them how to examine the temporal, frontal, and parietal rams, including common superficial artery localization, in longitudinal and transverse planes. The duration of hands-on scanning for each participant was around seven minutes.

Outcome measures
We used a two-level quantitative technique to assess the workshops.
1. Subjective learning perception with regard to levels of satisfaction and efficiency of the educational process
achieved by the participants: This was measured with an anonymous questionnaire structured with graded responses (1-5 Likert scale) that was administered at the end of the workshop; 22 of the 29 participants answered the questionnaire in 2007, and 35 of the 43 in 2008 (Table I).

2. Objective measures: a) Kappa inter-and intra-reader reliability; and b) validity, sensitivity, specificity, and percentage of correctly classified cases and controls against the gold standard diagnosis. For the validity exercise, the clinical diagnosis of GCA was established by experienced rheumatologists. The gold standard used was the final clinical diagnosis confirmed at the end of a follow-up period of at least 6 months by the rheumatologist responsible for the patients. The ACR 1990 GCA criteria were used for classification, and all patients with GCA final clinical diagnoses fulfilled 3 or more ACR criteria. Only five patients with GCA did not have temporal artery biopsy performed.

**Statistical analyses**
The kappa coefficient for categorical variables was used for the inter-reader and intra-reader statistical analysis. For every participant, the sensitivity, specificity, and percentage of correctly classified patients were analysed against the gold standard diagnosis using the STATA programme version 10.0 (StataCorp, Collage Station, TX).

**Results**
The results obtained from the questionnaire showed an excellent overall satisfaction rating of the 12 items scored, with a mean Likert score of 4.9 (range of 4.8 to 5 points) in the 2007 workshop, and a mean of 4.7 (range 4.5-4.9) in the 2008 meeting. The lesser satisfaction was due to insufficient time in the hands-on scanning exercise and the increased time spent on theoretical knowledge, participation, and reading exercise (see Table I).

In the 2007 workshop, the kappa correlation coefficient of inter-reader agreement for the 29 participants was excellent, with a kappa value of 0.846 (CI 0.668-0.937; p<0.0001). The kappa coefficient for each of the 29 readers is shown in Table II.

The maximum sensitivity was achieved by 21 of the 43 participants, and the lowest sensitivity reached was 86.7% in two cases. The sensitivity achieved by 24 of the 29 participants was 100%, so 82.7% of the readers attained maximum sensitivity. Furthermore, the sensitivity of the other five readers was 92.3%. The specificity results were also good, with a lowest value of 64.2%, and a perfect 100% attained by three readers. The percentage of correctly classified patients ranged from 100 to 77.8, with a mean value of 93.4 (Table II).

In the 2008 workshop, electronic wireless answers were used to check possible bias in the 2007 exercise, but the kappa correlation coefficient of inter-reader agreement for the 43 readers was also excellent, with a value of 0.848 (CI 0.681-0.936; p<0.0001), similar to the previous one. The sensitivity was slightly lower, but the specificity and percentage of cases and controls correctly classified was improved (see Table II).

In the case of sensitivity, 100% was achieved by 21 of the 43 participants, and the lowest sensitivity reached was 86.7% in two cases. The specificity was 100% in 19 cases, and the lowest specificity was 86.7% in seven cases. Finally, the percentage of cases correctly classified was between 100% (11 readers) and 90% (7 participants).

Seven readers participated in both workshops, and it was, thus, possible to perform an intra-reader reliability exercise with a kappa value of 0.950. These readers demonstrated a trend toward better results in comparison with the remaining participants, but without statistical difference. Sensitivity in the first lecture was 100% and in the second was 98.1%, the specificity in the first workshop was 88.8% while in the second it was 98.1%; and finally, the percentage of cases correctly classified was 94.2 in the first exercise, and 98.1 in the second.

**Discussion**
Within the last decade, an increasing number of rheumatologists have incorporated musculoskeletal ultrasound as a valuable imaging tool in their clinical practice.
practice. In the Cunnington et al. (18) study, the principal reason given for not performing musculoskeletal ultrasound was lack of training in this technique (75% of the cases assessed). An increasing number of papers propose how musculoskeletal ultrasound training for rheumatologists should be performed (18-26). This is the first study to provide information on an evaluated teaching method in GCA. A low teacher-to-learner ratio was employed in our method, with excellent results. If this type of low teacher-to-learner method is demonstrated to be efficient, it will be easier to perform US, because the high teacher-to-learner ratio actually employed in current US training, potentially imposes a limitation on the expansion of rheumatological US.

The validity of colour Doppler US for GCA diagnosis has been widely demonstrated (8, 15, 17), although it is not yet widely used. If the validity of this technique is accepted, why is it not used in practice? Perhaps this is due to lack of training in this technique, or due to the general idea that US is observer-dependent. The workshop was focused on two topics, as well as its on the reliability of US and the development of a teaching-learning procedure for GCA US diagnosis. Teaching CDUS in GCA was of additional interest to us because GCA is a low prevalence disease, meaning there are fewer chances for learning than with other more prevalent diseases. Measuring effectiveness in a short learning model is not easy. We selected the following as outcomes: subjective measurement of the learning perception achieved by the participants in different aspects of the workshop (education focused on who is learning), and objective measures, such as reliability and validity, via a reader exercise that was administered at the same time as the teaching method.

The questionnaire indicated excellent subjective perception with regard to the usefulness of the workshop (Table I). Some difficulties were demonstrated by the participants in acquisition of correct ultrasound temporal artery images due to the limited time available for hands-on scanning. Approximately seven minutes was allocated per participant, which was clearly insufficient, although participants could see and learn from the exploration of other participants. This was also perceived by the teaching team. W. Schmidt et al. reported that a sonographer experienced in vascular ultrasound must examine at least 30 people without temporal arteries to accurately assess temporal arteries (27). Therefore, acquiring the ability to detect GCA requires an investment of time. However, with theoretical and short practical training, we thought the participants would learn to recognize normal and pathological findings and understand elemental exploration, which would begin with the process of self-learning.

The results demonstrate the ease with which even non-expert individuals can identify normal and pathological ultrasound images in temporal artery examinations. The good reliability of US is probably, in part, a consequence of the high quality of the video images, but the records were not previously selected based on quality. Another possible bias in the high reliability of the study is whether participants can learn from their own mistakes, as they submitted comments after every case. If the lecture had been continuous, the results may have been less promising. Our principal interest was teaching, and therefore, we used reliability as a learning outcome. However, to explore whether this method produced a bias, we performed another blinded teaching exercise using the classical blind inter-reader methodology, without knowing the results of the patients at any time. The readers were three foreign rheumatology residents with theoretical knowledge of colour Doppler ultrasonography of the temporal artery. We used the same videos utilized in the 2007 workshop and the results did not demonstrate significant statistical differences, but had a slightly lower inter-reader kappa value of 0.796. The sensitivity, specificity, and percentage of well-classified subjects were 90.3, 95.4, and 92.6 respectively (28).

To the best of our knowledge, there are thus far no good reliability data for temporal artery ultrasound, and this is the biggest reliability study in GCA US. Previous studies have included only two participants, which supposes a higher possibility of agreement (7, 15, 17). In the study performed by Schmidt et al., the rates of agreement between the two ultrasonographers were 100% for the initial halo sign, 98% for stenoses, 96% for occlusions, and 95% for all three features (15). Similar results were obtained by Salvarani et al. (7), with two ultrasonographers and 100% agreement for halo diagnosis. The inter-reader agreement was 95% with two different sonographers performing temporal artery ultrasound with regard to positive or negative findings (15).

In our reliability exercise, the kappa correlation coefficient for the halo sign was 0.84, but with 29 and 43 non-expert readers. We believe this result shows that the reliability of US in clinical practice could be excellent. However, comparing videos in ultrasound studies is superior to comparing images, and our study cannot provide information on the combination of acquiring and reading images, which usually provide less agreement. Further, reliability in this study is a secondary objective and an outcome used to evaluate a standardized training, not a real exercise of reliability. However, we are of the opinion that publication of our results will contribute greatly to the field, as a similar exercise has never been done with this many readers. While not a perfect reliability exercise, ours is a good approximation.

Finally, the diagnostic gold standard used was the final clinical diagnosis confirmed at the end of a 6 month follow-up period by the patients’ rheumatologists’ all GCA patients included fulfilled 3 or more ACR criteria. Only five of the GCA patients did not have biopsies taken, but we did not want introduce a debate over the use of biopsies versus ultrasound in this paper. Sensitivity and specificity (Table II) was only an outcome measure of learning with a pre-test probability of 50%; they were not a primary objective of the study.

In conclusion, CDUS is reliable, independent of the observer, and can be learned with a short training method. The teaching strategies appeared to be efficient and well-received by the target audience, providing good evidence for a potential model for future training.
References


Appendix

Working Group Ultrasound Giant Cell Arteritis

Working Group Ultrasound Giant Cell Arteritis: Alvarez F (Rheumatology Unit, Hospital Nuestra Señora de la Candelaria, Tenerife, Spain), Babio J (Rheumatology Unit, Hospital de Cabueñes, Gijón, Spain), Bendaña JM (Rheumatology Unit, Hospital de Ourense, Ourense, Spain), Betancourt JJ (Rheumatology Unit, Hospital Universitario de Canarias, La Laguna, Tenerife, Spain), Cancio J (Rheumatology Unit, Hospital Comarcal Del Alto Deba, Arrasate/Modragón, Spain), Castro P (Rheumatology Unit, Hospital Universitario Ramón y Cajal, Madrid, Spain), Díaz-Cordovés G (Hospital Regional Universitario Carlos Haya, Málaga, Spain), Espiño P (Rheumatology Unit, Hospital Regional Universitario Carlos Haya, Málaga, Spain), Falcón MJ (Rheumatology Unit, Hospital Nuestra Señora de la Candelaria, Tenerife, Spain), Fernández O (Rheumatology Unit, Hospital Civil De Basurto, Bilbao, Spain), Gallego AM (Rheumatology Unit, Hospital General de Ciudad Real, Ciudad Real, Spain), García ME (Rheumatology Unit, Hospital de Cabueñes, Gijón, Spain), García M (Rheumatology Unit, La Paz University Hospital, Madrid, Spain), García C (Hospital de Belvitge, Rheumatology Unit, Barcelona, Spain), García E (Hospital Universitario German Trias y Pujol, Rheumatology Unit, Badalona, Spain), García PV (Rheumatology Unit, Hospital General de Ciudad Real, Ciudad Real, Spain), Guillen CA (Rheumatology Unit, Hospital Universitario Ramón y Cajal, Madrid, Spain), Hegedus I (Rheumatology Unit, National Institute of Rheumatology and Physiotherapy, Budapest, Hungary), Holgado S (Hospital Universitario German Trias y Pujol, Rheumatology Unit, Badalona, Spain), Jiménez FG (Hospital Regional Universitario Carlos Haya, Málaga, Spain), Jiménez M (Rheumatology Unit, Hospital Universitario Puerta De Hierro, Madrid, Spain), Llerín C (Rheumatology Unit, Hospital Nuestra Señora de la Candelaria, Tenerife, Spain), Lluch P (Rheumatology Unit, Hospital Mateu Orfila, Menorca, Spain), Luna C (Rheumatology Unit, Hospital Universitario de Canarias, La Laguna, Tenerife, Spain), Martín JA (Rheumatology Unit, Hospital Nuestra Señora de la Candelaria, Tenerife, Spain), Martínez E (Rheumatology Unit, Hospital Universitario Ramón y Cajal, Madrid, Spain), Medeiros D (Instituto Portugués de Reumatología, Lisboa, Portugal), Mediavilla MJ (Instituto Portugués de Reumatología, Lisboa, Portugal), Mera A (Hospital Clínico Universitario de Santiago de Compostela, Santiago de Compostela, Spain), Míguez S (Rheumatology Unit, Hospital Universitario German Trias y Pujol, Badalona, Spain), Miquel X (Rheumatology Unit, Hospital General D'Igualada, Igualada, Spain), Moreno J (Rheumatology Unit, Hospital Virgen de La Arrixaca, Murcia, Spain), Morlà R (Rheumatology Unit, Hospital de Santa Tecla, Tarragona, Spain), Muñoz J (Rheumatology Unit, Iamas. Hospital De L'Esperança y Hospital Del Mar, Barcelona, Spain), Ordóñez MC (Rheumatology Unit, La Paz University Hospital, Madrid, Spain), Pérez E (Rheumatology Unit, Hospital Universitario Ramón y Cajal, Madrid, Spain), Polo MA (Rheumatology Unit, Fundación Hospital de Calahorra. Calahorra. La Rioja, Spain), Quirosa S (Rheumatology Unit, Hospital Clínico San Cecilio, Granada, Spain), Ramírez J (Rheumatology Unit, Hospital Universitario Virgen Del Rocío, Sevilla, Spain), Ribeiro A (Rheumatology Unit, Hospital Ponte de Lima, Portugal), Richi P (Rheumatology Unit, Hospital Infantia Sofia, Madrid, Spain), Rodríguez C (Rheumatology Unit, Hospital Doctor Negrin, Las Palmas de Gran Canaria, Spain), Roxo AS (Rheumatology Unit, HospitalPonte de Lima, Portugal), Salles M (Hospital Universitario German Trias y Pujol, Rheumatology Unit, Badalona), Sánchez J (Rheumatology Unit, Hospital General Carlos Haya, Málaga, Spain), Sánchez O (Rheumatology Unit, Hospital General de Elda. Elda. Alicante, Spain), Soley JM (Rheumatology, Lleida, Spain), Soria D (Rheumatology Unit, H.U. Puerta del Mar. Cadiz, Spain), Soro S (Rheumatology Unit, Hospital General de Ciudad Real, Ciudad Real, Spain), Tuneu R (Rheumatology, Centre Hospitalari Manresa, Spain), Varas B (Rheumatology, Hospital Universitario Ramón y Cajal, Madrid, Spain), Weruaga A (Rheumatology Unit, Hospital de Cabueñes, Gijón, Spain).