ORIGINAL RESEARCH

Imaging in diagnosis, monitoring and outcome prediction of large vessel vasculitis: a systematic literature review and meta-analysis informing the 2023 update of the EULAR recommendations

Philipp Bosch ,[1] Milena Bond ,[2] Christian Dejaco ,[1,2] Cristina Ponte ,[3,4] Sarah Louise Mackie ,[5,6] Louise Falzon ,[7] Wolfgang A Schmidt ,[8] Sofia Ramiro [9,10]

ABSTRACT

Objectives To update the evidence on imaging for diagnosis, monitoring and outcome prediction in large vessel vasculitis (LVV) to inform the 2023 update of the European Alliance of Associations for Rheumatology recommendations on imaging in LVV.

Methods Systematic literature review (SLR) (2017–2022) including prospective cohort and cross-sectional studies (>20 participants) on diagnostic, monitoring, outcome prediction and technical aspects of LVV imaging. Diagnostic accuracy data were meta-analysed in combination with data from an earlier (2017) SLR.

Results The update retrieved 38 studies, giving a total of 81 studies when combined with the 2017 SLR. For giant cell arteritis (GCA), and taking clinical diagnosis as a reference standard, low risk of bias (RoB) studies yielded pooled sensitivities and specificities (95% CI) of 88% (82% to 92%) and 96% (95% CI 86% to 99%) for ultrasound (n=8 studies), 81% (95% CI 71% to 89%) and 98% (95% CI 89% to 100%) for MRI (n=3) and 76% (95% CI 67% to 83%) and 95% (95% CI 71% to 99%) for fluorodeoxyglucose positron emission tomography (FDG-PET, n=4), respectively. Compared with studies assessing cranial arteries only, low RoB studies with ultrasound assessing both cranial and extracranial arteries revealed a higher sensitivity (93% (95% CI 88% to 96%) vs 80% (95% CI 71% to 87%)) with comparable specificity (94% (95% CI 83% to 98%) vs 97% (95% CI 71% to 100%)). No new studies on diagnostic imaging for Takayasu arteritis (TAK) were found. Some monitoring studies in GCA or TAK reported associations of imaging with clinical signs of inflammation. No evidence was found to determine whether imaging severity might predict worse clinical outcomes.

Conclusion Ultrasound, MRI and FDG-PET revealed a good performance for the diagnosis of GCA. Cranial and extracranial vascular ultrasound had a higher pooled sensitivity with similar specificity compared with limited cranial ultrasound.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Ultrasound and MRI have good diagnostic accuracies for cranial giant cell arteritis (GCA).
⇒ Ultrasound and fluorodeoxyglucose positron emission tomography (FDG-PET) have been used to detect extracranial artery inflammation.

WHAT THIS STUDY ADDS

⇒ FDG-PET for cranial and extracranial artery inflammation has a good diagnostic performance in GCA with a pooled sensitivity of 76% (95% CI 67% to 83%) and pooled specificity of 95% (95% CI 71% to 99%).
⇒ Ultrasound of extracranial (mostly including axillary artery) as well as cranial arteries has higher sensitivity and comparable specificity to only scanning cranial arteries.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The results of this systematic literature review will inform an international task force formulating the 2023 update of the European Alliance of Associations for Rheumatology recommendations on the use of imaging in large vessel vasculitis (LVV), and can help clinicians in choosing the most adequate imaging technique for diagnostic purposes in LVV.
⇒ Vascular ultrasound for GCA will accurately detect more cases if the axillary arteries are included in the scan.

INTRODUCTION

Primary large vessel vasculitis (LVV) includes giant cell arteritis (GCA), Takayasu arteritis (TAK) and idiopathic aortitis. Previously, our group published a systematic literature review (SLR) and European Alliance of Associations for Rheumatology (EULAR) endorsed recommendations for the use of imaging techniques

in LVV, including ultrasound, MRI, fluorodeoxyglucose positron emission tomography (FDG-PET) and CT.[1 2]

That SLR, containing studies until 2017, concluded that ultrasound and MRI had a good diagnostic accuracy for cranial GCA. FDG-PET studies were scarce and only assessed extracranial arteries, preventing conclusions on its diagnostic accuracy on cranial vessels.[2] Furthermore, there was little data on the role of imaging for monitoring disease activity and outcome prediction.

Since the publication of the original EULAR recommendations, the use of large vessel imaging has become more widespread, leading to a noticeable increase in the number of publications reporting the use of imaging in diagnosis, monitoring disease activity and outcome prediction in LVV. Furthermore, new data on FDG-PET has been published,[3 4] as well as novel studies evaluating the potential of ocular imaging techniques for diagnosing GCA,[5 6] altogether adding to the existing literature. Lastly, more research has been done on the diagnostic value of extended ultrasound examinations in GCA, also including large vessels such as the axillary artery.[7 8] Whether these studies yield better higher sensitivities and specificities than studies only assessing cranial arteries warrants further exploration.

Our aim was to update the previous SLR and inform the EULAR task force on new evidence on the use of imaging for diagnosis, monitoring and outcome prediction in LVV.[9]

METHODS
Research questions
The protocol of this SLR has been published on PROSPERO (ID: CRD42022360545). The four research questions, depicted in the Population, Intervention, Control and Outcome format (PICO),[10] were identical to those of the 2017 SLR. They comprised the use of different imaging techniques for LVV in (1) diagnostics, (2) monitoring, (3) outcome prediction and (4) technical aspects that should be considered when using them (online supplemental tables S1a–d). Apart from the studies on imaging techniques included in the previous SLR (ultrasound, MRI, FDG-PET and CT),[2] studies on optical coherence tomography (OCT) and fluoresceine angiography (FA) were also included in this review, as studies reporting the diagnostic value of these imaging techniques had recently been published.

The population of interest was adults (≥18 years) with suspected primary LVV (for diagnostic accuracy studies) or with established primary LVV (for monitoring, prediction and technical aspects studies), particularly GCA, TAK and isolated aortitis. For diagnostic studies, the intervention was the imaging findings (=index test) and the comparator was either the clinical diagnosis or temporal artery biopsy (TAB) results (=reference standard). The outcome was test performance, including sensitivity, specificity, positive likelihood ratio (LR+) and negative LR (LR−). For monitoring studies, intervention

and comparator were 'performing imaging' and 'not performing imaging', respectively, and the outcome was disease activity. For prediction studies, intervention and comparator were 'positive imaging' and 'negative imaging', respectively, with a variety of outcomes such as disease complications, cumulative glucocorticoid dose and disease activity. Lastly, studies on technical aspects compared different imaging settings regarding 'optimal results' (not prespecified).

Only full research articles (excluding commentaries and letters) were eligible. These included prospective cohort studies involving >20 patients, with <50% lost to follow-up. Cross-sectional studies were only included for studies of diagnostic accuracy and technical aspects. Case–control and retrospective studies were excluded.

Literature search, data extraction, risk of bias assessment
The search strategy was developed and run by an experienced librarian (LF). For the 2017 SLR, Medline, Embase, the Cochrane database and Epistemonikos were searched from inception to 10 March 2017.[2] As some databases do not allow to perform a search from an exact date and manuscripts may have been published online ahead of print, or were accepted by a journal without being fully available at that time, the interval of the updated search was extended to 1 January 2016 to be sure not to miss any relevant papers. Consequently, for this SLR, the same databases as mentioned above were searched from 1 January 2016 to 16 November 2022 with the same search string used in the 2017 SLR. A new string concerning the new imaging techniques was additionally run from inception to 16 November 2022 (online supplemental table S2). Overlaps between the 2017 and updated SLR were checked manually and excluded.

The good sensitivity of the search strategy was confirmed by testing for key publications proposed by the steering committee. Two reviewers (PB and MB) independently screened all titles and abstracts using the online literature programme Rayyan.[11] Subsequently, 20% of full texts of potentially eligible studies were read by both reviewers. As the agreement to include studies was higher than the prespecified minimal cut-off value (kappa>0.7), the remaining articles were split between the two reviewers for inclusion and exclusion. A similar interrater reliability exercise was performed for data extraction and risk of bias (RoB) assessment and the remaining articles were again allocated between the reviewers. Disagreement during the reliability exercise was discussed for consensus and resolved by a methodologist (SR) if necessary.

Data were extracted using a prespecified data extraction sheet as previously described.[2] Briefly, this included general study characteristics, study design and details on the PICO of interest. For diagnostic accuracy studies, the number of true positives, true negative, false positives and false negatives (FN) were collected to calculate diagnostic test properties. A LR+ and LR− of >4 and <0.4, respectively, were prespecified as minimum thresholds for clinically useful imaging tests. Rules of thumb

for interpreting LR+ and LR− at the bedside have been described elsewhere.[12]

RoB assessment was performed using the Quality Assessment of Diagnostics Accuracy Studies-2 (QUADAS-2) tool for diagnostic accuracy studies and the Quality In Prognosis Studies (QUIPS) for monitoring and prediction studies.[13 14] No RoB assessment was performed for studies on technical aspects. For monitoring studies the domain 'prognostic factor measurement"'of the QUIPS was omitted due to lack of such a factor per design. Overall RoB is depicted as 'low', 'unclear' and 'high' for QUADAS-2 and 'low', 'moderate' and 'high' for QUIPS (online supplemental tables S3a,b).

Data analysis

Meta-analyses were performed for diagnostic accuracy studies using the imaging findings as index test, and either clinical diagnosis or TAB as two separate reference standards. Ultrasound findings of different aspects of arterial inflammation, including halo sign, compression sign, stenosis and occlusion were pooled, and reported here as 'halo sign', because they represent different aspects of the same pathological process and because the field has converged on halo sign as the central imaging feature of GCA. If studies reported halo sign and other ultrasound features of GCA separately, the data relating specifically to the halo sign were used in the meta-analysis.

Studies were only included in the meta-analysis if they reported all findings that would allow the calculation of sensitivity and specificity. For all other diagnostic accuracy studies, as well as studies on monitoring, prediction and technical aspects, individual results are reported separately.

To conduct meta-analyses, data from the 2017 SLR[2] and the current update were pooled. Random-effects bivariate generalised binomial mixed models were applied in case three or more studies were available for a certain index test and reference standard combination. These models are the adequate analytical method because they consider correlations between sensitivities and specificities and therefore estimate them together. If only three studies were available for a specific meta-analysis, univariate random-effects models were used instead. Meta-analysis was not performed if less than three studies were available. LR+/LR− was obtained using the delta method.[15]

We report the results of the meta-analysis both for all studies and separately for studies with low RoB. Further subanalyses were performed comparing diagnostic properties between studies assessing cranial arteries (ie, temporal, facial, occipital and maxillary arteries) only and studies assessing both cranial and extracranial arteries using imaging.

Sensitivity analyses were conducted considering studies, in which: (1) all patients were glucocorticoid naïve before imaging; (2) for ultrasound, ≥15 MHz probes were used; (3) for MRI, a 3 Tesla machine was used; (4) the clinical diagnosis was confirmed at a follow-up visit; (5) the index test and the reference standard were performed independently from one another (ie, without the assessors of one test being aware of the results of the other test). Sensitivity analyses were performed for all studies fulfilling the mentioned criteria, as well as restricted to the subset of these studies with low RoB whenever possible.

All analyses were performed using R V.4.2.1 using the 'lme4', 'msm' and 'forestploter' packages.

RESULTS

From a total of 6578 screened references, 38 eligible studies were found through the update. Together with the 43 studies identified by the 2017 SLR published, this gave 81 studies in total (online supplemental figure S1). Some of the 38 studies addressed multiple index tests, reference standards, research questions and mixed populations of GCA and TAK patients.

For GCA, 22 studies on diagnostic accuracy,[3–5 7 16–33] 8 on monitoring disease activity,[34–41] 5 on outcome prediction[4 34 42–44] and 5 on technical aspects[8 30 43 45 46] were found in the update. For TAK, there were no studies on diagnostic accuracy, four on monitoring[38 41 47 48] and one each on prediction[44] and technical aspects.[46] No studies on isolated aortitis satisfied the inclusion criteria.

Diagnostic accuracy studies

Regarding diagnostic accuracy studies in GCA, 17 were on ultrasound,[3 5 7 16–29] 3 on MRI,[5 30 31] 6 on FDG-PET[3 4 23 25 32 33] and 1 on FA in the update.[5] No studies were found for CT. Study characteristics and RoB assessment are depicted in tables 1–3 and online supplemental table S4a,b respectively.

Ultrasound

For the meta-analysis of ultrasound, with the clinical diagnosis as reference standard, and considering both the current update and the previous SLR, there were 12 (4 with low,[3 7 21 22] 7 with unclear[5 16 18–20 23 24] and 1 with high RoB[17]) and 11 studies (4 with low,[49–52] 2 with unclear[53 54] and 5 with high RoB),[55–59] respectively, summing up to a total of 1981 patients included. Pooled sensitivities and specificities were 75% (95% CI 66% to 83%) and 91% (95% CI 86% to 94%) when considering all studies (online supplemental figure S2) and 88% (95% CI 82% to 92%) and 96% (95% CI 86% to 99%) when focusing on low RoB studies (figure 1). For the latter group, pooled LR+ was 20.07 (95% CI 6.23 to 64.64) and pooled LR− was 0.13 (95% CI 0.09 to 0.18).

Pooling estimates from all studies assessing both cranial and extracranial arteries resulted in a higher sensitivity and similar specificity as compared with the analysis of only cranial arteries (see table 4). This result was also seen when filtering for low RoB studies (Sens: 93% (95% CI 88% to 96%) vs 80% (95% CI 71% to 87%) and Spec: 94% (95% CI 83% to 98%) vs 97% (95% CI 71% to 100%)) as shown in online supplemental table S5. The most frequently assessed extracranial arterial territories were the axillary (7/7 studies)[3 7 21 25 26 51 54] and the

Table 1 Characteristics of diagnostic accuracy studies on ultrasound in GCA

Author, year	Study design	N	Investigated arteries	Index test (elementary lesions)	Reference standard	RoB
Chrysidis, 2021[7]	Prospective cohort	106	Axillary, carotid, facial, temporal arteries	(Halo sign and compression sign in temporal and/or facial arteries) or (imt≥1 mm in axillary and/or imt>1.5 mm in carotid arteries)	Clinical diagnosis	Low
Conway, 2019[16]	Prospective cohort	162	Temporal arteries	Halo sign	Clinical diagnosis	Unclear
Czihal, 2021[17]	Cross-sectional	114	Temporal arteries	Compression sign with IMT≥0.7 mm	Clinical diagnosis	High
Estrada Alarcón, 2019[18]	Prospective cohort	52	Temporal arteries	Halo sign	Clinical diagnosis /TAB	Unclear
González Porto, 2020[19]	Prospective cohort	56	Temporal arteries	Halo sign	Clinical diagnosis	Unclear
Hansen, 2022[20]	Cross-sectional	78	Temporal arteries	Halo sign	Clinical diagnosis	Unclear
Imfeld, 2020[25]	Prospective cohort	102	Axillary, carotid, femoral, popliteal, subclavian, temporal, vertebral arteries	Halo sign with/without stenosis	Clinical diagnosis	High
Lecler, 2022[5]	Prospective cohort	45	Temporal arteries	Halo sign and/or thickened arterial wall	Clinical diagnosis	Unclear
Mukhtyar, 2020[21]	Prospective cohort	25	Axillary, temporal arteries	Halo sign and/or stenosis and/or occlusion	Clinical diagnosis	Low
Nielsen, 2020[3]	Prospective cohort	90	Axillary, carotid, temporal arteries	Halo sign with/without compression sign	Clinical diagnosis	Low
Noumegni, 2022[22]	Cross-sectional	80	Temporal arteries	Halo sign	Clinical diagnosis	Low
Rottenburger, 2021[23]	Cross-sectional	34	Temporal arteries	Halo sign with/without stenosis	Clinical diagnosis	Unclear
Sundholm, 2019[24]	Cross-sectional	75	Temporal arteries	Halo sign	Clinical diagnosis /TAB	Unclear
van der Geest, 2020[26]	Prospective cohort	89	Axillary, temporal arteries	Halo score and Halo count	Clinical diagnosis /TAB	Low
Bilyk, 2017[27]	Prospective cohort	71	Temporal arteries	Halo sign	TAB	Unclear
Gielis, 2021[28]	Prospective cohort	80	Temporal arteries	Halo sign with IMT>0.5 mm	TAB	Unclear
Moudrous, 2022[29]	Cross-sectional	93	Axillary, carotid, temporal arteries	Halo sign	TAB	Low

GCA, giant cell arteritis; IMT, intima media thickness; N, number of participants; RoB, risk of bias; TAB, temporal artery biopsy.

common carotid arteries (4/7 studies).[3 7 25 51] Sensitivity analyses revealed similar results, either when all studies were considered or when only low RoB studies were analysed (table 4 and online supplemental table S5).

Two studies (one with low[26] and one with high RoB[25]) were not included in the meta-analysis due to missing data. Imfeld *et al* reported a sensitivity of 57% and a specificity of 97%,[25] while van der Geest *et al* reported

Table 2 Characteristics of diagnostic accuracy studies on MRI in GCA

Author, year	Study design	N	Arteries	Index test	Reference standard	RoB
Lecler, 2022[5]	Prospective cohort	45	Temporal, occipital arteries	Wall thickening and/or contrast enhancement	Clinical diagnosis	Unclear
Poillon, 2020[30]	Cross-sectional	79	Extracranial: temporal, occipital arteries intracranial: basilary, carotid, cerebral, vertebral arteries	Wall thickening and/or contrast enhancement	Clinical diagnosis	Low
Rodriguez-Régent, 2020[31]	Cross-sectional	32	Temporal, occipital arteries	Contrast enhancement	Clinical diagnosis	Low

GCA, giant cell arteritis; N, number of participants; RoB, risk of bias.

diagnostic test properties for the halo count (Sens: 78%, Spec: 55%) and the halo score (Sens: 78% and Spec: 61%) separately. Halo count and halo score include bilaterally the common temporal artery, the parietal and frontal branch as well as the axillary artery and score each segment with 0/1 or 0–3, respectively.[26]

Similar results for ultrasound were also found when using TAB as the reference standard. Five studies from the update (1 low,[29] 4 moderate RoB[18 24 27 28] and 10 studies) from the previous SLR (1 low,[60] 1 moderate,[54] 8 high RoB),[55–59 61–63] including 1126 participants, were analysed. These revealed pooled sensitivities of 73%

Table 3 Characteristics of diagnostic accuracy studies on FDG-PET in GCA

Author, year	Study design	N	Arteries	Index test	Reference standard	RoB
Imfeld, 2020[25]	Prospective cohort	102	Aorta, carotid, femoral, iliac, proximal arm arteries	Artery-SUVmax liver-SUVmean ratio >1 in the supra-aortic region or >1.3 in the aorto-iliofemoral region	Clinical diagnosis	High
Imfeld, 2022[33]	Prospective cohort	50	Abd/thor aorta, axillary, carotid, femoral, iliac, subclavian, vertebral arteries	Artery-SUVmax liver-SUVmean ratio >1 in the supra-aortic region or >1.3 in the aorto-iliofemoral region	Clinical diagnosis	High
Nielsen, 2020[3]	Prospective cohort	90	Maxillary, temporal, vertebral arteries	Vessel uptake above surrounding tissue uptake	Clinical diagnosis	Low
Rottenburger, 2021[23]	Cross-sectional	34	Maxillary, temporal, occipital arteries	Artery-SUVmax liver-SUVmean ratio >1	Clinical diagnosis	Unclear
Sammel, 2019[4]	Prospective cohort	64	Aortic arch, asc/desc aorta, axillary, brachiocephalic, carotid, maxillary, occipital, subclavian, temporal, vertebral arteries	Subjective global assessment for GCA based on intensity and distribution of vascular uptake	Clinical diagnosis/TAB	Low
Thibault, 2022[32]	Prospective cohort	51	Extracranial: abd/thor aorta, axillary, carotid, femoral, iliac, subclavian intracranial: maxillary, occipital, temporal, vertebral arteries	Similar or higher vessel uptake in respect to liver uptake	Clinical diagnosis/TAB	Low

FDG-PET, fluorodeoxyglucose positron emission tomography; GCA, giant cell arteritis; N, number of participants; RoB, risk of bias; SUV, standardised uptake value; TAB, temporal artery biopsy.
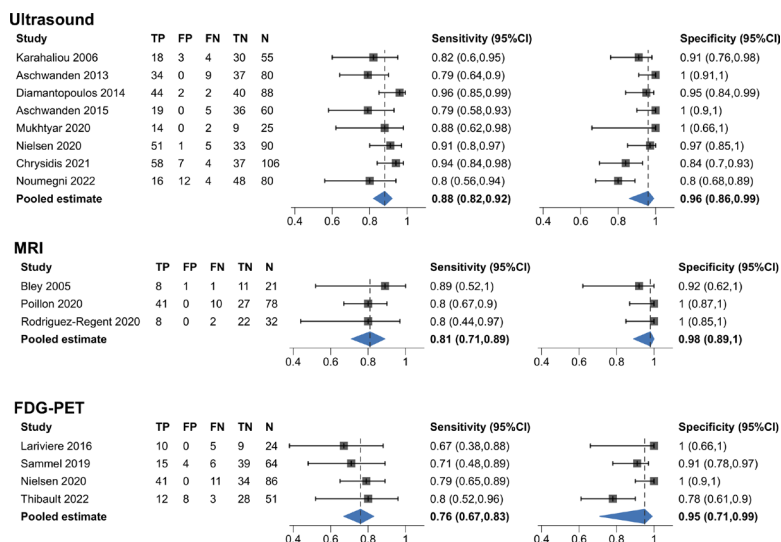
Figure 1 Diagnostic performance of ultrasound, MRI and FDG-PET in comparison with clinical diagnosis as reference standard for the diagnosis of GCA according to low RoB studies Diagnostic performance according to all studies is depicted in online supplemental figure S3–S5. This plot only contains studies with low risk of bias. FN, false negative; FP, false positive; FDG-PET, fluorodeoxyglucose positron emission tomography; GCA, giant cell arteritis; N, number of participants; RoB, Risk of bias; TN, true negative; TP, true positive.

(95% CI 61% to 83%) and specificities of 84% (95% CI 74% to 90%).

MRI

For MRI, with the reference standard defined as clinical diagnosis, there were three new studies (two low,[30 31] one moderate RoB[5]) and five studies from the previous SLR (one low,[64] three moderate[65–67] and one high RoB[68]) included in the meta-analysis (492 participants). All these studies assessed exclusively cranial vessels. Pooled sensitivity was 82% (95% CI 76% to 86%) and specificity was 92% (95% CI 84% to 97%) for all studies and 81% (95% CI 87% to 89%) and 98% (95% CI 89% to 100%) for studies with low RoB, respectively (figure 1 and online supplemental figure S3). For the group of studies with low RoB, pooled LR+ was 49.67 (95% CI 7.09 to 348.1) and pooled LR− was 0.19 (95% CI 0.12 to 0.31). Sensitivity analyses assessing only studies using 3 Tesla MRI, or studies in which the index test and reference standard were performed independently, confirmed the primary result (table 4, online supplemental table S5). No new studies on diagnostic properties of MRI with the reference standard TAB were found.

Fluorodeoxyglucose positron emission tomography

Four new studies (three with low,[3 4 32] one with unclear[23] RoB) and one study with low RoB[69] from the previous SLR on the value of FDG-PET in regard to the clinical diagnosis were included in the meta-analysis (n=259 participants). Pooled sensitivities and specificities were 80% (95% CI 70% o 97%) and 91% (95% CI 67% to 98%) for all studies, respectively (online supplemental figure S4), as well as 76% (95% CI 67% to 83%) and 95% (95% CI 71% to 99%) for studies with low RoB, respectively (figure 1), with pooled LR+ of 14.5 (95% CI

2.21 to 94.96) and pooled LR− of 0.26 (95% CI 0.18 to 0.37). Sensitivity analyses revealed similar results (table 4, online supplemental table S5). Two studies with high RoB, that were not included in the meta-analysis due to missing data, reported sensitivities of 72%[25] and 89%,[33] as well as specificities of 85% and 73%.

Only two studies with low RoB reported diagnostic properties for FDG-PET with TAB as reference standard (sensitivity: 92% (95% CI 62% to 100%) and specificity: 85% (95% CI 71% to 94%),[4] sensitivity: 70% (95% CI 42% to 98%) and specificity: 86% (95% CI 72% to 100%).[32]

FA and indocyanine green angiography

One study with high RoB evaluated the value of FA and indocyanine green angiography (ICG).[5] The delay in choroidal vessel filling (FA) or the presence of non-vascularised choroidal areas (ICG) were considered as positive imaging tests. Using the clinical diagnosis as reference standard, the authors found a sensitivity of 88% (95% CI 69% to 97%) and a specificity of 74% (95% CI 49% to 91%).

Direct comparisons of imaging techniques

Four studies directly compared ≥2 imaging techniques for GCA diagnosis (online supplemental table S6a,b).

Three studies (one with low,[3] two with high[23 25] RoB) used both ultrasound and FDG-PET as index test. Imfeld et al[25] and Rottenburger et al[23] reported a lower sensitivity (57% vs 94% and 53% vs 94%) and higher specificity (97% vs 59% and 94% vs 59%) for ultrasound compared with FDG-PET, respectively. While Imfeld et al assessed only cranial arteries with ultrasound and FDG-PET, Rottenburger et al investigated only extracranial arteries using both imaging modalities. Nielsen et al[3] found a higher sensitivity (91% vs 79%) but a lower specificity

Table 4 Subgroup and sensitivity analyses of all studies on diagnostic properties for giant cell arteritis using imaging as index test and clinical diagnosis as reference standard

	No of patients (no of studies)	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR− (95% CI)
Ultrasound					
Main analysis (all studies)*	1981 (23)[3 5 7 16–24 49–59]	0.75 (0.66 to 0.83)	0.91 (0.86 to 0.94)	8.28 (4.98 to 13.74)	0.27 (0.19 to 0.39)
Only cranial arteries†	1291 (18)[5 16–20 22–24 49 50 52 53 55–59]	0.70 (0.59 to 0.79)	0.91 (0.84 to 0.95)	7.60 (4.15 to 13.91)	0.33 (0.23 to 0.47)
Cranial and extracranial arteries	690 (5)[3 7 21 51 54]	0.89 (0.73 to 0.96)	0.91 (0.83 to 0.95)	9.34 (4.79 to 18.23)	0.13 (0.05 to 0.33)
100% GC naive‡	332 (5)[3 52 53 58]	0.78 (0.58 to 0.90)	0.87 (0.77 to 0.92)	5.83 (2.97 to 11.42)	0.25 (0.12 to 0.56)
≥15MHz probe	889 (13)[3 5 7 17 18 20 22–24 49 50 52 55]	0.80 (0.68 to 0.89)	0.90 (0.84 to 0.94)	8.19 (4.88 to 13.75)	0.22 (0.13 to 0.37)
Diagnosis at follow-up§	710 (9)[3 7 16 18 19 51–53 55]	0.81 (0.67 to 0.90)	0.86 (0.78 to 0.91)	5.80 (3.32 to 10.15)	0.22 (0.11 to 0.43)
MRI					
Main analysis (all studies)*	492 (8)[5 30 31 64–68]	0.82 (0.76 to 0.86)	0.92 (0.84 to 0.97)	10.62 (4.77 to 23.65)	0.20 (0.15 to 0.26)
3 Tesla	243 (6)[5 30 31 64 65 68]	0.84 (0.71 to 0.91)	0.93 (0.78 to 0.98)	11.59 (3.53 to 38.04)	0.18 (0.09 to 0.32)
Independent assessment§	295 (3)[30 31 67]	0.81 (0.74 to 0.86)	0.99 (0.26 to 1.00)	68.12 (0.30 to 15477.46)	0.19 (0.14 to 0.27)
FDG-PET					
Main analysis (all studies)*	259 (5)[3 4 23 32 69]	0.80 (0.70 to 0.87)	0.91 (0.67 to 0.98)	9.13 (2.17 to 38.4)	0.22 (0.15 to 0.33)
Diagnosis at follow-up§	225 (4)[3 4 32 69]	0.76 (0.67 to 0.83)	0.95 (0.71 to 0.99)	14.50 (2.21 to 94.96)	0.26 (0.18 to 0.37)
Independent assessment¶	184 (3)[3 4 23]	0.80 (0.70 to 0.87)	0.92 (0.52 to 0.99)	10.31 (1.14 to 93.45)	0.22 (0.14 to 0.34)

*Including all available studies from the 2017 SLR and the update.
†Including temporal, maxillary, occipital and facial arteries.
‡No glucocorticoids before index test was performed.
§Diagnosis was confirmed at follow-up visit.
¶Including studies in which the personnel evaluated the index and reference test independently.
FDG-PET, fluorodeoxyglucose positron emission tomography; GC, glucocorticoids; LR+/LR−, positive/negative likelihood ratio.

(79% vs 100%) for ultrasound compared with FDG-PET, assessing both cranial and extracranial arteries.

Lecler *et al*[5] (high RoB) compared ultrasound (Sens: 86%, Spec: 94%), MRI (100%, 84%) and FA/ICG (88%, 74%) at cranial arteries, reporting the highest sensitivity for MRI and the highest specificity for ultrasound.

Monitoring disease activity

Ten prospective studies assessing ultrasound or contrast enhanced ultrasound (CEUS),[34–37 47 48] FDG-PET,[38–40] MR-angiography (MRA)[41] and CT-angiography (CTA)[41] were retrieved. Overall, all imaging techniques were found to be capable of detecting vascular changes during follow-up, indicating their potential usefulness as tools for monitoring disease activity. A summary of the evidence for each imaging modality is provided below, while detailed information of each study is reported in online supplemental table S7.

Ultrasound

Six studies (three with low,[35–37] two with moderate[34 48] and one with high RoB[47]) investigated the role of ultrasound in GCA[34–37] and of ultrasound with or without contrast enhancement (CE) in TAK.[47 48]

With regard to GCA, findings of the previous SLR confirmed that, after treatment initiation, the improvement[34 36 37] and normalisation[34] of wall thickening tended to occur earlier and more consistently in temporal arteries as compared with large vessels. Moreover, some new evidence emerged: throughout follow-up, a weak to moderate correlation was observed between the halo sign (number of halos and intima–media thickness (IMT) of vessels affected by the halo) of temporal and axillary arteries and disease activity markers, assessed through the Birmingham vasculitis activity score (BVAS), or acute phase reactants (erythrocyte sedimentation rate, C reactive protein (CRP)).[35 36] Furthermore, a new ultrasound score to monitor disease activity (the OMERACT GCA Ultrasonography Score) was presented and tested prospectively, showing a large to very large magnitude of change (standardised mean difference between −1.19 and −2.16) over a 24-week period of follow-up after treatment initiation.[35] This score is calculated as the sum of IMT measured in every segment (bilateral common temporal arteries, frontal and parietal branches as well as axillary arteries) divided by the rounded cut-off values of IMTs in each segment and consecutively divided by the number of segments available.

For TAK, two prospective studies showed the association of ultrasound and CEUS[47 48] findings with disease activity; specifically, a reduction of the mean IMT was observed from baseline to 3 months follow-up (2.24 mm vs 1.85 mm, p=0.041). Moreover, CE in the carotid wall decreased more frequently in patients who were clinically active at baseline as compared with those who were inactive (overall, patients having a grade of CE>2 decreased from 53% to 34%, p=0.036).[47 48]

Positron emission tomography

Two studies with low RoB in GCA[39 40] and one study with moderate RoB in both GCA and TAK[38] evaluated the role of FDG-PET for monitoring disease activity.

Sammel *et al*[40] reported a decrement of the Total Vascular Score (TVS, considering 18 different vascular segments; scale: 0–54) in both cranial and extracranial arteries 6-month after treatment initiation in newly diagnosed GCA patients compared with baseline. Moreover, Quinn *et al*[39] demonstrated that in tocilizumab-treated patients, the PET Vascular Activity Score (PETVAS, evaluating 4 segments of the aorta and 11 branch arteries; scale: 0–27) decreased over time. Six months after treatment discontinuation, PETVAS increased in a small proportion of patients. The third study, which included both GCA and TAK patients, reported a decrease in arterial standardised FDG uptake values (SUV) among patients with an increase in immunosuppressive treatment whereas those without a change in medication showed stable target to background ratio (obtained by the normalisation of the vascular FDG uptake to liver SUV) values.[38]

MRA and CTA

One study with moderate RoB, including both GCA and TAK patients, presented and validated a new quantitative composite score for clinical trials using CTA or MRA. This score evaluated the presence of dilation and stenosis in 17 arterial territories.[41] The authors reported a larger change of the score in patients who were clinically active at baseline as compared with those who were inactive according to the Indian Takayasu Clinical Activity Score (Δ between active and inactive patients in the imaging change score: 1.56 (0.66), p=0.020), the change of patient global assessment (3.09 (1.02), p=0.003), but not the change of National Institute of Health score (1.19 (1.0) p=0.26). Moreover, using the radiologist's assessment as reference standard, the score showed excellent discrimination (area under the receiver operating characteristic curve=0.996, sensitivity 77%, specificity 96%) for detecting disease progression.

Outcome prediction

Five prospective studies providing information on the role of ultrasound in GCA patients[34 42 43] and FDG-PET in GCA[40] and in both GCA and TAK patients[44] for outcome prediction were included. No data were retrieved for the remaining imaging techniques. Overall, imaging did not predict response to treatment, occurrence of relapses or ischaemic complications. The presence of FDG uptake at baseline, however, was linked with the subsequent development of angiographic changes (stenosis, occlusion or aneurysm) in both GCA and TAK (online supplemental table S8).

Ultrasound

Three studies (two with low,[42 43] one with moderate RoB[34] including GCA patients only, provided information on this aspect. Over a 2-year period, no differences were

found between patients with and without a reduction of IMT during follow-up concerning subsequent relapses, cumulative glucocorticoid doses and need for glucocorticoid sparing agents.[34] Moreover, no differences in terms of new ischaemic complications were reported between patients with or without a halo sign at baseline.[43] Lastly, a composite model, taking into account the IMT of temporal and axillary arteries as well as the bilateral presence of the halo in the same vascular territories, showed no predictive value for any of these ultrasound parameters concerning the clinical outcomes evaluated (ie, visual loss, vascular damage index, glucocorticoid doses >10 mg/day of prednisone equivalent at 6 months and/or the need for adjunctive immunosuppressants).[42]

Positron emission tomography

One study with low RoB, including only GCA patients, reported that a baseline TVS>10 could not predict relapses, glucocorticoid cumulative doses and CRP levels over a 12-month follow-up.[40] Another study with low RoB and a mixed population of both GCA and TAK patients revealed that territories with abnormal FDG uptake at baseline were at a higher risk of developing subsequent angiographic changes (OR 19.5, 95% CI 2.44 to 156).[44]

Technical aspects of imaging techniques

Five studies providing information on technical aspects of ultrasound,[8 43] FDG-PET[45 46] and MR-A[30] were included. The time points after treatment initiation and examination of vascular territories were found to be important factors determining the sensitivity of ultrasound and FDG-PET. Details of each study are shown in online supplemental table S9. No information was retrieved for the remaining imaging techniques.

Ultrasound

Two cross-sectional studies were eligible.[8 43] A subanalysis of the Temporal Artery Biopsy versus ULtrasound study evaluated the temporal artery IMT during the first week after glucocorticoid treatment initiation.[43] The results showed that the IMT was greater in patients undergoing ultrasound evaluation at the day of treatment start compared with those having the ultrasound exam 1 week after glucocorticoids were initiated (r=0.30, p=0.001). The second study compared the ultrasound examination of temporal arteries and axillary arteries (visualised from the axilla) with an extended protocol additionally including the proximal axillaries and brachial, the subclavian and the common carotid arteries.[8] The extended approach was more sensitive in detecting large vessel inflammation than the limited one (p<0.001).

Positron emission tomography

Two prospective studies were retrieved.[45 46] In the article by Nielsen *et al*, 24 GCA glucocorticoid-naïve patients with a positive FDG-PET exam at baseline were included and divided into two groups undergoing a second scan either at day 3 or day 10 after starting glucocorticoids.[45] While in the first group, 100% of patients still had a positive scan after 3 days, only 36% of patients in the second group showed tracer uptake after ten days. The second study included both GCA and TAK patients. More scans were considered positive with an acquisition time of 2 hours (77%) as compared with 1 hour (56%, p<0.01).[46]

MRI

A single cross-sectional study on GCA patients, comparing different reformatting techniques and using TAB or the clinical diagnosis as reference standard, was included.[30] The multiplanar reformatting 3D contrast enhanced vessel wall (CEVW) MR had a specificity of 100% to show inflammatory changes of extracranial arteries and was more sensitive than the axial only 3D (80% vs 73%, p=0.046) and 2D CEVW MRI (80% vs 70%, p=0.03).

DISCUSSION

This SLR confirms the good diagnostic accuracy of ultrasound and MRI, and furthermore shows a comparable performance of FDG-PET with these techniques for diagnosing GCA, taking clinical diagnosis as the reference standard. Ultrasound and FDG-PET studies also included extracranial arteries in their assessment for GCA, while for MRI, data are only available on cranial arteries. While studies reporting direct comparisons of imaging techniques were few and in part contradictory, indirect comparisons suggested that ultrasound might have a higher sensitivity than MRI and FDG-PET with a comparable specificity.

Since our first SLR, containing studies until 2017,[2] several meta-analyses have reported diagnostic properties of ultrasound in GCA patients. While Nakajima *et al*[70] found similar pooled estimates (Sens: 86%, Spec: 95%) than we did (Sens: 88%, Spec: 96%), Sebastian *et al*[71] reported a lower sensitivity (Sens: 67%, Spec: 95%) and Rinagel *et al*[72] a lower sensitivity and specificity (Sens: 68%, Spec: 81%).

One possible explanation is that the other reviews included retrospective and prospective studies, generally leading to a more heterogenous sample with higher RoB. Rinagel *et al* only included studies assessing temporal arteries, which may have reduced the sensitivity to detect pathology in patients with LV-GCA compared with studies also assessing extracranial arteries.

An important result of our SLR was the observation that the ultrasound assessment of cranial plus extracranial arteries had a higher sensitivity than the investigation of cranial arteries alone, while specificity was comparable. The increased sensitivity for a GCA diagnosis can be explained by the higher number of scanned arteries and therefore higher chance of finding imaging signs of vasculitis. To improve sensitivity, it may be sufficient to assess axillary arteries in addition to temporal arteries rather than applying a more extended scanning protocol. The axillary arteries were the most frequently scanned extracranial vessels in these studies and the data suggest that there is no incremental benefit of including

additional extracranial arteries.[7 51 73] A separate analysis of the diagnostic accuracy of FDG-PET and MRI studies for cranial and extracranial arteries was not possible because the number of studies reporting sensitivity and specificity for cranial or extracranial arteries separately was insufficient.

We only included prospective and cross-sectional studies in our SLR to increase precision, but also to reduce the RoB, which is often high in retrospective studies. Selection bias may occur when imaging is only performed in certain patient groups (eg, dubious cases), rather than in all patients with suspected disease. Expectation bias may lead to an overestimation of diagnostic properties when the imaging assessor is also aware of the clinical symptoms of a patient (which is common in retrospective studies). Lastly, the selection bias inherent in case–control studies can lead to an overestimation of the value of the imaging technique, as controls are usually not patients with suspected GCA, but rather healthy controls or patients with other diseases, leading to an unrealistically large contrast between cases and controls. We have, therefore, opted for the stricter approach by excluding both retrospective and case–control studies. We believe that further meta-analyses in the field should have a similar approach, which is now possible, due to the large number of high-quality studies.

For ultrasound studies we pooled the results of different imaging abnormalities (ie, halo sign, compression sign, occlusion and stenosis) as they not only describe the same underlying inflammatory process, that is, inflammatory thickening of the arterial wall, but also have similar diagnostic properties.[2 49 50] This reflects new insights in the field and clarity that was obtained throughout the last years of GCA research.

SLRs including meta-analysis on FDG-PET have also been published before, reporting similar results to ours. These reviews, however, used the 1990 ACR criteria and TAB rather than the clinical diagnosis as reference standard.[74 75] TAB has a considerable risk of FN results, especially in GCA patients with large vessel involvement, while the 1990 ACR criteria, also relying on TAB, largely focus on cranial symptoms that are not necessarily present in LV-GCA.[76–79] Moreover, when assessing the diagnostic performance of an imaging test, we are interested in its performance compared with the clinical diagnosis and not to classification criteria that are intended to select a homogeneous population for clinical trials. Classification criteria should not be used for clinical diagnosis.

Using the clinical diagnosis as reference standard has the disadvantage of potential circular reasoning as imaging data are often considered to support the diagnosis. Nevertheless, when performing sensitivity analyses focusing on studies with independent index tests and reference standards, we did not observe different results concerning diagnostic properties, which confirms the robustness of our findings.

Only a single study on ocular imaging techniques was found by the SLR, reporting lower diagnostic yields than ultrasound and MRI.[5] The role of this technique, as well as that of other ocular imaging modalities such as OCT, ultra-wide-field colour fundus photography or transocular sonography in LVV still remains unclear.[6 80 81]

Evidence on the role of imaging for monitoring disease activity or outcome prediction in LVV is still limited but, compared with the 2017 SLR,[2] some new data emerged for ultrasound and FDG-PET reporting a change of imaging signs of inflammation after treatment initiation.

Previous studies on FDG-PET seemed not to discriminate between clinically active and inactive disease, while newer studies revealed an association between FDG uptake and change in treatment in both GCA and TAK.

There is a growing evidence that the reduction of ultrasound measured IMT occurs earlier in temporal than in axillary arteries. We can only speculate whether the different ultrasonographic behaviour of temporal arteries and large vessels is due to distinct histopathological features of the arteries or to technical limitations of ultrasound assessment (ie, some persistent thickening may be present also in the temporal arteries but remaining below the technical detection limit). The latter hypothesis is supported by histology studies that report a persisting T helper 1-driven tissue inflammation in TABs despite clinical remission.[82 83]

Weak to moderate associations between ultrasound signs of activity and markers of disease activity were found. New data on the sensitivity to change of different imaging techniques have emerged and imaging scores for ultrasound and MRI reported good results for construct validity and reliability.[35 41] However, studies using imaging as a monitoring tool, especially to assess its value in guiding treatment decisions compared with a clinical approach alone, are still missing.

The number of studies on the predictive value of imaging was low. In line with previous results,[2] no association was found between imaging and relapses, cumulative glucocorticoid doses or new ischaemic complications in LVV.[34 40 43] Retrospective studies, partially reporting contradictory results for the predictive value of ultrasound in regard to cumulative glucocorticoid doses and relapses, need further investigation in prospective cohorts.[76 79 84]

Some studies have reported on the technical aspects of LVV imaging, highlighting the importance of timely imaging after treatment initiation. They have shown that the interpretation of activity based on FDG-PET and MR-A also depends on acquisition times and reformatting methods, respectively. Timing remains a crucial factor in performing imaging: as the sooner it is performed, the more sensitive the test is. However, it is unclear what the time limit is for obtaining positive results, and whether this differs across imaging modalities.

A strength of our SLR is the relatively high number of studies with a low RoB on diagnostic properties for ultrasound, MRI and FDG-PET, allowing us to present sensitivities and specificities, together with LR+/LR− of these studies as the main results of this SLR. The results of

sensitivity analyses were in line with the primary results, confirming the robustness of our data. Having excluded retrospective and case–control studies increased the quality and precision of our data. The main limitations are the lack of new literature on the diagnostic properties of TAK and the heterogeneity and low number of prospective monitoring and prediction studies.

In summary, we report good diagnostic performances of ultrasound, MRI and FDG-PET for the clinical diagnosis of GCA and a better diagnostic sensitivity when adding the axillary arteries to the ultrasound examination. Prospective cohort studies on monitoring disease activity and on the prediction of relevant outcomes are needed, as well as studies on the value of imaging for the diagnosis of TAK. This SLR will inform the task force updating the EULAR recommendations on the use of imaging in LVV.[9]

Author affiliations
[1]Department of Rheumatology and Immunology, Medical University of Graz, Graz, Austria
[2]Department of Rheumatology, Hospital of Bruneck (ASAA-SABES), Teaching Hospital of the Paracelsius Medical University, Brunico, Italy
[3]Rheumatology Research Unit, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
[4]Rheumatology Department, Hospital de Santa Maria, Centro Hospitalar de Lisboa Norte, EPE, Lisbon, Portugal
[5]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK
[6]Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[7]Health Economics and Decision Science, The University of Sheffield, Sheffield, UK
[8]Department of Rheumatology, Immanuel Krankenhaus Berlin, Medical Centre for Rheumatology Berlin-Buch, Berlin, Germany
[9]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[10]Department of Rheumatology, Zuyderland Medical Center, Heerlen, The Netherlands

Twitter Philipp Bosch @philipp_j_bosch, Milena Bond @milena_bond, Cristina Ponte @cristinadbponte and Sofia Ramiro @sofiaramiro82

Contributors All authors helped designing the study. LF performed the literature search. PB and MB performed literature screening, data extraction, risk of bias assessment and meta-analysis. All authors participated in the interpretation of the data. PB and MB prepared the manuscript. PB and SR are responsible for the overall content as the guarantors. All authors critically appraised the manuscript for important intellectual content and approved the final manuscript.

Disclaimer The views expressed in this article are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Competing interests PB has received speaker fees by Janssen and project grants by Pfizer. MB has received research grants from AbbVie. CD has received consulting/speaker's fees from Abbvie, Eli Lilly, Janssen, Novartis, Pfizer, Roche, Galapagos and Sanofi, all unrelated to this manuscript. He is an editorial board member of ARD. WAS has received speaker honoraria from Abbvie, Amgen, Bristol Myers Squibb, Chugai, Lilly, Johnson & Johnson, Medac, Novartis, Pfizer, Roche, Sanofi, and UCB; consultancy fees from Abbvie, Amgen, Bristol Myers Squibb, Chugai, GlaxoSmithKline, Novartis, Roche, and Sanofi. He is principal investigator of phase 2 and phase 3 trials sponsored by Abbvie, GlaxoSmithKline, Novartis, and Sanofi. CP is or has been the principal investigator of studies by AbbVie, Sanofi and Novartis and has received consulting/speaker's fees from CSL Vifor, AbbVie, AstraZeneca, GlaxoSmithKline and Roche, all unrelated to this manuscript. SLM reports: Consultancy on behalf of her institution for Roche/Chugai, Sanofi, AbbVie, AstraZeneca, Pfizer; Investigator on clinical trials for Sanofi, GSK, Sparrow; speaking/lecturing on behalf of her institution for Roche/Chugai, Vifor, Pfizer, UCB, Novartis and AbbVie; chief investigator on STERLING-PMR trial, funded by NIHR;

patron of the charity PMRGCAuk. No personal remuneration was received for any of the above activities. Support from Roche/Chugai to attend EULAR2019 in person and from Pfizer to attend ACR Convergence 2021 virtually. SLM is supported in part by the NIHR Leeds Biomedical Research Centre. The views expressed in this article are those of the authors and not necessarily those of the NIHR, the NIHR Leeds Biomedical Research Centre, the National Health Service or the UK Department of Health and Social Care. SR has received research grants and/or consultancy fees from AbbVie, Eli Lilly, Galapagos, MSD, Novartis, Pfizer, UCB.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Philipp Bosch http://orcid.org/0000-0002-6783-6422
Milena Bond http://orcid.org/0000-0002-5400-2955
Christian Dejaco http://orcid.org/0000-0002-0173-0668
Cristina Ponte http://orcid.org/0000-0002-3989-1192
Sarah Louise Mackie http://orcid.org/0000-0003-2483-5873
Louise Falzon http://orcid.org/0000-0002-9449-6056
Wolfgang A Schmidt http://orcid.org/0000-0001-7831-8738
Sofia Ramiro http://orcid.org/0000-0002-8899-9087

REFERENCES
1 Dejaco C, Ramiro S, Duftner C, *et al*. EULAR recommendations for the use of imaging in large vessel vasculitis in clinical practice. *Ann Rheum Dis* 2018;77:636–43.
2 Duftner C, Dejaco C, Sepriano A, *et al*. 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.
3 Nielsen BD, Hansen IT, Keller KK, *et al*. Diagnostic accuracy of ultrasound for detecting large-vessel giant cell arteritis using FDG PET/CT as the reference. *Rheumatology (Oxford)* 2020;59:2062–73.
4 Sammel AM, Hsiao E, Schembri G, *et al*. Diagnostic accuracy of positron emission tomography/computed tomography of the head, neck, and chest for giant cell arteritis: a prospective, double-blind, cross-sectional study. *Arthritis Rheumatol* 2019;71:1319–28.
5 Lecler A, Hage R, Charbonneau F, *et al*. Validation of a multimodal algorithm for diagnosing giant cell arteritis with imaging. *Diagn Interv Imaging* 2022;103:103–10.
6 Maldiney T, Greigert H, Martin L, *et al*. Full-field optical coherence tomography for the diagnosis of giant cell arteritis. *PLoS One* 2020;15:e0234165.
7 Chrysidis S, Døhn UM, Terslev L, *et al*. Diagnostic accuracy of vascular ultrasound in patients with suspected giant cell arteritis (Eureka): a prospective, multicentre, non-Interventional, cohort study. *Lancet Rheumatol* 2021;3:e865–73.
8 Bull Haaversen AC, Brekke LK, Kermani TA, *et al*. Extended ultrasound examination identifies more large vessel involvement in patients with giant cell arteritis. *Rheumatology (Oxford)* 2023;62:1887–94.
9 Dejaco C, Ramiro S, Bond M, *et al*. EULAR recommendations for the use of imaging in large vessel vasculitis in clinical practice: 2023 update. *Ann Rheum Dis* 2023:ard-2023-224543.
10 Sackett D, Richardson W, Rosenberg W, *et al*. Evidence-based medicine. how to practice and teach EBM. In: *Evidence-based medicine*. Churchill Livingston, 2005: 2.

11 Ouzzani M, Hammady H, Fedorowicz Z, *et al*. Rayyan—a web and mobile app for systematic reviews. *Syst Rev* 2016;5:210.

12 McGee S. Simplifying likelihood ratios. *J Gen Intern Med* 2002;17:646–9.

13 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.

14 Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006;144:427–37.

15 Takwoingi Y, Guo B, Riley RD, *et al*. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Stat Methods Med Res* 2017;26:1896–911.

16 Conway R, O'Neill L, McCarthy GM, *et al*. Performance characteristics and predictors of temporal artery ultrasound for the diagnosis of giant cell arteritis in routine clinical practice in a prospective cohort. *Clin Exp Rheumatol* 2019;37 Suppl 117:72–8.

17 Czihal M, Köhler A, Lottspeich C, *et al*. Temporal artery compression sonography for the diagnosis of giant cell arteritis in elderly patients with acute ocular arterial occlusions. *Rheumatology (Oxford)* 2021;60:e213–4.

18 Estrada Alarcón P, Reina D, Navarro Ángeles V, *et al*. Doppler ultrasonography of superficial temporal artery in a cohort of patients with strong clinical suspicion of giant cell arteritis. *Med Clin (Barc)* 2019;153:151–3.

19 González Porto SA, Silva Díaz MT, Reguera Arias A, *et al*. A comparative study of doppler ultrasound against temporal artery biopsy in the diagnosis of giant cell arteritis. *Reumatol Clin (Engl Ed)* 2020;16:313–8.

20 Hansen MS, Terslev L, Jensen MR, *et al*. Comparison of temporal artery ultrasound versus biopsy in the diagnosis of giant cell arteritis. *Eye (Lond)* 2023;37:344–9.

21 Mukhtyar C, Myers H, Scott DGI, *et al*. Validating a diagnostic GCA ultrasonography service against temporal artery biopsy and long-term clinical outcomes. *Clin Rheumatol* 2020;39:1325–9.

22 Noumegni SR, Jousse-Joulin S, Hoffmann C, *et al*. Comparison of halo and compression signs assessed by a high frequency ultrasound probe for the diagnosis of giant cell arteritis. *J Ultrasound* 2022;25:837–45.

23 Rottenburger C, Mensch N, Imfeld S, *et al*. 18F-FDG PET/CT compared with ultrasound and biopsy for detection of vasculitis of the temporal artery branches. *Swiss Med Wkly* 2021;151:w20512.

24 Sundholm JKM, Pettersson T, Paetau A, *et al*. Diagnostic performance and utility of very high-resolution ultrasonography in diagnosing giant cell arteritis of the temporal artery. *Rheumatol Adv Pract* 2019;3:rkz018.

25 Imfeld S, Aschwanden M, Rottenburger C, *et al*. 18F]FDG positron emission tomography and ultrasound in the diagnosis of giant cell arteritis: congruent or complementary imaging methods *Rheumatology (Oxford)* 2020;59:772–8.

26 van der Geest KSM, Borg F, Kayani A, *et al*. Novel ultrasonographic halo score for giant cell arteritis: assessment of diagnostic accuracy and association with ocular ischaemia. *Ann Rheum Dis* 2020;79:393–9.

27 Bilyk JR, Murchison AP, Leiby BT, *et al*. The utility of color duplex Ultrasonography in the diagnosis of giant cell arteritis: A prospective, masked study. *Trans Am Ophthalmol Soc* 2017;115:T9.

28 Gielis JF, Geelhoed R, Yogeswaran SK, *et al*. Evaluation of temporal artery duplex ultrasound for diagnosis of temporal arteritis. *J Surg Res* 2021;261:320–5.

29 Moudrous W, Visser LH, Yilmaz T, *et al*. A new prediction model for giant cell arteritis in patients with new onset headache and/or visual loss. *Ann Med* 2022;54:2770–6.

30 Poillon G, Collin A, Benhamou Y, *et al*. Increased diagnostic accuracy of giant cell arteritis using three-dimensional fat-saturated contrast-enhanced vessel-wall magnetic resonance imaging at 3 T. *Eur Radiol* 2020;30:1866–75.

31 Rodriguez-Régent C, Ben Hassen W, Seners P, *et al*. 3D T1-weighted black-blood magnetic resonance imaging for the diagnosis of giant cell arteritis. *Clin Exp Rheumatol* 2020;38 Suppl 124:95–8.

32 Thibault T, Durand-Bailloud B, Soudry-Faure A, *et al*. PET/CT of cranial arteries for a sensitive diagnosis of giant cell arteritis. *Rheumatology (Oxford)* 2023;62:1568–75.

33 Imfeld S, Scherrer D, Mensch N, *et al*. A simplified PET/CT measurement routine with excellent diagnostic accuracy for the diagnosis of giant cell arteritis. *Diagnostics (Basel)* 2022;12:728.

34 Aschwanden M, Schegk E, Imfeld S, *et al*. Vessel wall plasticity in large vessel giant cell arteritis: an ultrasound follow-up study. *Rheumatology (Oxford)* 2019;58:792–7.

35 Dejaco C, Ponte C, Monti S, *et al*. The provisional OMERACT ultrasonography score for giant cell arteritis. *Ann Rheum Dis* 2022;82:556–64.

36 Ponte C, Monti S, Scirè CA, *et al*. Ultrasound halo sign as a potential monitoring tool for patients with giant cell arteritis: a prospective analysis. *Ann Rheum Dis* 2021;80:1475–82.

37 Sebastian A, Kayani A, Prieto-Pena D, *et al*. Efficacy and safety of tocilizumab in giant cell arteritis: a single centre NHS experience using imaging (ultrasound and PET-CT) as a diagnostic and monitoring tool. *RMD Open* 2020;6:e001417.

38 Dashora HR, Rosenblum JS, Quinn KA, *et al*. Comparing semiquantitative and qualitative methods of vascular 18F-FDG PET activity measurement in large-vessel vasculitis. *J Nucl Med* 2022;63:280–6.

39 Quinn KA, Dashora H, Novakovich E, *et al*. Use of 18F-Fluorodeoxyglucose positron emission tomography to monitor tocilizumab effect on vascular inflammation in giant cell arteritis. *Rheumatology (Oxford)* 2021;60:4384–9.

40 Sammel AM, Hsiao E, Schembri G, *et al*. Cranial and large vessel activity on positron emission tomography scan at diagnosis and 6 months in giant cell arteritis. *Int J Rheum Dis* 2020;23:582–8.

41 Tombetti E, Godi C, Ambrosi A, *et al*. Novel angiographic scores for evaluation of large vessel vasculitis. *Sci Rep* 2018;8:15979.

42 Monti S, Ponte C, Pereira C, *et al*. The impact of disease extent and severity detected by quantitative ultrasound analysis in the diagnosis and outcome of giant cell arteritis. *Rheumatology (Oxford)* 2020;59:2299–307.

43 Ponte C, Serafim AS, Monti S, *et al*. Early variation of ultrasound halo sign with treatment and relation with clinical features in patients with giant cell arteritis. *Rheumatology (Oxford)* 2020;59:3717–26.

44 Quinn KA, Ahlman MA, Alessi HD, *et al*. Association of 18F-Fluorodeoxyglucose positron emission tomography and angiographic progression of disease in large-vessel vasculitis. *Arthritis Rheumatol* 2023;75:98–107.

45 Nielsen BD, Gormsen LC, Hansen IT, *et al*. Three days of high-dose glucocorticoid treatment attenuates large-vessel 18F-FDG uptake in large-vessel giant cell arteritis but with a limited impact on diagnostic accuracy. *Eur J Nucl Med Mol Imaging* 2018;45:1119–28.

46 Quinn KA, Rosenblum JS, Rimland CA, *et al*. Imaging acquisition technique influences interpretation of positron emission tomography vascular activity in large-vessel vasculitis. *Semin Arthritis Rheum* 2020;50:71–6.

47 Dong Y, Wang Y, Wang Y, *et al*. Ultrasonography and contrast-enhanced ultrasound for activity assessment in 115 patients with carotid involvement of Takayasu arteritis. *Mod Rheumatol* 2022:roac107.

48 Ma L-Y, Li C-L, Ma L-L, *et al*. Value of contrast-enhanced ultrasonography of the carotid artery for evaluating disease activity in Takayasu arteritis. *Arthritis Res Ther* 2019;21:24.

49 Aschwanden M, Daikeler T, Kesten F, *et al*. Temporal artery compression sign--a novel ultrasound finding for the diagnosis of giant cell arteritis. *Ultraschall Med* 2013;34:47–50.

50 Aschwanden M, Imfeld S, Staub D, *et al*. The ultrasound compression sign to diagnose temporal giant cell arteritis shows an excellent Interobserver agreement. *Clin Exp Rheumatol* 2015;33:S–113.

51 Diamantopoulos AP, Haugeberg G, Hetland H, *et al*. Diagnostic value of color doppler ultrasonography of temporal arteries and large vessels in giant cell arteritis: a consecutive case series. *Arthritis Care Res (Hoboken)* 2014;66:113–9.

52 Karahaliou M, Vaiopoulos G, Papaspyrou S, *et al*. Colour duplex sonography of temporal arteries before decision for biopsy: a prospective study in 55 patients with suspected giant cell arteritis. *Arthritis Res Ther* 2006;8:R116.

53 Habib HM, Essa AA, Hassan AA. Color duplex ultrasonography of temporal arteries: role in diagnosis and follow-up of suspected cases of temporal arteritis. *Clin Rheumatol* 2012;31:231–7.

54 Luqmani R, Lee E, Singh S, *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.

55 Nesher G, Shemesh D, Mates M, *et al*. The predictive value of the halo sign in color doppler ultrasonography of the temporal arteries for diagnosing giant cell arteritis. *J Rheumatol* 2002;29:1224–6.

56 Pfadenhauer K, Weber H. Duplex sonography of the temporal and occipital artery in the diagnosis of temporal arteritis. A prospective study. *J Rheumatol* 2003;30:2177–81.

57 Reinhard M, Schmidt D, Hetzel A. Color-coded sonography in suspected temporal arteritis-experiences after 83 cases. *Rheumatol Int* 2004;24:340–6.

58 Salvarani C, Silingardi M, Ghirarduzzi A, *et al*. Is duplex ultrasonography useful for the diagnosis of giant-cell arteritis? *Ann Intern Med* 2002;137:232–8.

59 Schmidt WA, Kraft HE, Vorpahl K, *et al*. Color duplex ultrasonography in the diagnosis of temporal arteritis. *N Engl J Med* 1997;337:1336–42.

60 Romera-Villegas A, Vila-Coll R, Poca-Dias V, *et al*. The role of color duplex sonography in the diagnosis of giant cell arteritis. *J Ultrasound Med* 2004;23:1493–8.

61 LeSar CJ, Meier GH, DeMasi RJ, *et al*. The utility of color duplex ultrasonography in the diagnosis of temporal arteritis. *Journal of Vascular Surgery* 2002;36:1154–60.

62 Murgatroyd H, Nimmo M, Evans A, *et al*. The use of ultrasound as an aid in the diagnosis of giant cell arteritis: a pilot study comparing histological features with ultrasound findings. *Eye (Lond)* 2003;17:415–9.

63 Zaragozá García JM, Plaza Martínez A, Briones Estébanez JL, *et al*. Value of the doppler-ultrasonography for the diagnosis of temporal arteritis. *Med Clin (Barc)* 2007;129:451–3.

64 Bley TA, Weiben O, Uhl M, *et al*. Assessment of the cranial involvement pattern of giant cell arteritis with 3T magnetic resonance imaging. *Arthritis Rheum* 2005;52:2470–7.

65 Siemonsen S, Brekenfeld C, Holst B, *et al*. 3T MRI reveals extra- and intracranial involvement in giant cell arteritis. *American Journal of Neuroradiology* 2015;36:91–7.

66 Bley TA, Uhl M, Carew J, *et al*. Diagnostic value of high-resolution MR imaging in giant cell arteritis. *AJNR Am J Neuroradiol* 2007;28:1722–7.

67 Klink T, Geiger J, Both M, *et al*. Giant cell arteritis: diagnostic accuracy of MR imaging of superficial cranial arteries in initial diagnosis-results from a multicenter trial. *Radiology* 2014;273:844–52.

68 Geiger J, Bley T, Uhl M, *et al*. Diagnostic value of T2-weighted imaging for the detection of superficial cranial artery inflammation in giant cell arteritis. *J Magn Reson Imaging* 2010;31:470–4.

69 Lariviere D, Benali K, Coustet B, *et al*. Positron emission tomography and computed tomography angiography for the diagnosis of giant cell arteritis: a real-life prospective study. *Medicine (Baltimore)* 2016;95:e4146.

70 Nakajima E, Moon FH, Junior NC, *et al*. Accuracy of Doppler ultrasound in the diagnosis of giant cell arteritis: a systematic review and meta-analysis. *Adv Rheumatol* 2023;63:5.

71 Sebastian A, Coath F, Innes S, *et al*. Role of the halo sign in the assessment of giant cell arteritis: a systematic review and meta-analysis. *Rheumatol Adv Pract* 2021;5:rkab059.

72 Rinagel M, Chatelus E, Jousse-Joulin S, *et al*. Diagnostic performance of temporal artery ultrasound for the diagnosis of giant cell arteritis: a systematic review and meta-analysis of the literature. *Autoimmun Rev* 2019;18:56–61.

73 Oshinsky C, Bays AM, Sacksen I, *et al*. The usefullness of Subclavian artery ultrasound assessment in giant cell arteritis evaluation. *J Clin Rheumatol* 2023;29:43–6.

74 Besson FL, Parienti J-J, Bienvenu B, *et al*. Diagnostic performance of ^{18}F-Fluorodeoxyglucose positron emission tomography in giant cell arteritis: a systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging* 2011;38:1764–72.

75 Soussan M, Nicolas P, Schramm C, *et al*. Management of large-vessel vasculitis with FDG-PET: a systematic literature review and meta-analysis. *Medicine (Baltimore)* 2015;94:e622.

76 Muratore F, Kermani TA, Crowson CS, *et al*. Large-vessel giant cell arteritis: a cohort study. *Rheumatology (Oxford)* 2015;54:463–70.

77 Schmidt WA, Seifert A, Gromnica-Ihle E, *et al*. Ultrasound of proximal upper extremity arteries to increase the diagnostic yield in large-vessel giant cell arteritis. *Rheumatology* 2008;47:96–101.

78 Brack A, Martinez-Taboada V, Stanson A, *et al*. Disease pattern in cranial and large-vessel giant cell arteritis. *Arthritis Rheum* 1999;42:311–7.

79 Bosch P, Dejaco C, Schmidt WA, *et al*. Association of ultrasound-confirmed axillary artery vasculitis and clinical outcomes in giant cell arteritis. *Semin Arthritis Rheum* 2022;56:152051.

80 Czihal M, Lottspeich C, Köhler A, *et al*. Transocular sonography in acute arterial Occlusions of the eye in elderly patients: diagnostic value of the spot sign. *PLoS One* 2021;16:e0247072.

81 Poignet B, Bonnin P, Gaudric J, *et al*. Correlation between ultra-wide-field retinal imaging findings and vascular supra-aortic changes in Takayasu arteritis. *J Clin Med* 2021;10:4916.

82 Deng J, Younge BR, Olshen RA, *et al*. TH17 and TH1 T-cell responses in giant cell arteritis. *Circulation* 2010;121:906–15.

83 Maleszewski JJ, Younge BR, Fritzlen JT, *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.

84 Tomelleri A, Campochiaro C, Sartorelli S, *et al*. Presenting features and outcomes of cranial-limited and large-vessel giant cell arteritis: a retrospective cohort study. *Scand J Rheumatol* 2022;51:59–66.