Validation of the Southend Giant Cell Arteritis Probability Score in a Scottish single centre fast-track pathway

Authors
Andrew R Melville 1, Karen Donaldson2, James Dale2, Anna Ciechomska2,3

Affiliations
1. Institute of Infection, Immunity & Inflammation, University of Glasgow, Glasgow, UK.
2. Rheumatology Department, University Hospital Wishaw, NHS Lanarkshire, UK.
3. School of Medicine, Dentistry and Nursing, University of Glasgow, Glasgow, UK.

Correspondence
Dr Anna Ciechomska,
University Hospital Wishaw, 50 Netherton Street, Wishaw ML2 0DP, United Kingdom.
Email: anna.ciechomska@glasgow.ac.uk
ORCID ID: 0000-0002-3136-438X
Abstract

**Objective:** To externally validate the Southend GCA Probability Score (GCAPS) in patients attending a GCA Fast-Track Pathway (GCA FTP) in NHS Lanarkshire.

**Methods:** Consecutive GCA FTP patients between November 2018 and December 2020 underwent GCAPS assessment as part of routine care. GCA diagnoses were supported by USS +/- TAB and confirmed at 6 months. Percentages of patients with GCA according to GCAPS risk group, performance of total GCAPS in distinguishing GCA / non-GCA final diagnoses, and test characteristics using different GCAPS binary cut-offs, were assessed. Associations between individual GCAPS components and GCA, and the value of USS and TAB in the diagnostic process, were also explored.

**Results:** 44/129 patients were diagnosed with GCA, including 0/41 GCAPS low risk patients (GCAPS <9), 3/40 medium risk (GCAPS 9-12), and 41/48 high risk (GCAPS >12). Overall performance of GCAPS in distinguishing GCA / non-GCA was excellent [ROC AUC 0.976 (95% CI 0.954 – 0.999)]. GCAPS cut-off ≥10 had 100.0% sensitivity and 67.1% specificity for GCA. GCAPS cut-off ≥13 had highest accuracy (91.5%), with 93.2% sensitivity and 90.6% specificity. Several individual GCAPS components were associated with GCA. Sensitivity of USS increased by ascending GCAPS risk group (nil, 33.3%, 90.2% respectively). TAB was diagnostically useful in cases where USS was inconclusive.

**Conclusion:** This is the first published study describing application of GCAPS outside the specialist centre where it was developed. Performance of GCAPS as a risk stratification tool was excellent. GCAPS may have additional value for screening GCA FTP referrals and guiding empirical glucocorticoid treatment.

**Key Messages**

1. External validation confirms the value of the Southend GCA Probability Score (GCAPS) as a risk assessment and also diagnostic tool.
2. GCAPS <10 may be sufficient to exclude GCA without fast-track review.
3. Patients with GCAPS ≥13 appear high risk for GCA and warrant empirical glucocorticoids.

**Keywords**

Giant cell arteritis, temporal arteritis, probability score, fast track pathway, ultrasound, diagnosis.
INTRODUCTION

International guidelines recommend temporal and axillary artery ultrasound as a first-line modality for confirmation of diagnoses of Giant Cell Arteritis (GCA) (1, 2). Fast track GCA pathways (GCA FTP) with application of ultrasound are growing in popularity, however the number and availability of trained sonographers is a potential limiting factor. Delayed clinical assessment carries the risk of complications, including loss of vision (3), as well as jeopardising the value of clinical findings and test results in patients started on glucocorticoids by their referring clinician.

Risk stratification scores could help prioritise referrals to GCA FTPs, facilitating timely assessment and treatment of patients with GCA, and avoidance of over-treatment with glucocorticoids in those with other diagnoses. The Southend GCA Probability Score (GCAPS), also known as the Southend GCA clinical pre-test probability score (4), developed in a specialist centre at Southend University Hospital, has shown promise in discriminating patients with low, medium, and high pre-test probability for GCA (5), and appears to augment the diagnostic performance of ultrasound. There is a suggestion that GCAPS alone might be sufficient to exclude GCA in low risk patients, without additional tests, which if confirmed could have major significance for resource allocation. These findings require validation, particularly given the potentially serious consequences of missed diagnoses. Evidence to support the use of GCAPS in external cohorts is currently limited, consisting of conference reports but no published studies (6, 7).

NHS Lanarkshire is a health board in Scotland (UK) serving a population of 650,000 people across urban and rural communities. A GCA FTP was established in 2018, based on the earlier experiences of FTPs in other parts of the UK (7, 8).

AIMS

The aim of this study was to assess the performance of GCAPS in our cohort of patients from NHS Lanarkshire, including i) ability of GCAPS to categorise patients at the pre-assessment stage to groups of low, medium and high risk, ii) ability of GCAPS to discriminate GCA from non-GCA final diagnoses, and iii) whether GCAPS alone might be sufficient to exclude GCA in patients deemed low risk (i.e. below a specified binary cut-off).

Additional aims were to explore the predictive values of individual GCAPS components in this dataset, to report the diagnostic yield of ultrasonography of cranial and axillary arteries (USS) across GCAPS risk groups, and to describe the role of temporal artery biopsy (TAB) in the diagnostic process.

METHODS

Study design

We conducted a retrospective cohort study of consecutive patients with suspected GCA assessed on the NHS Lanarkshire GCA FTP over a 2-year period (November 2018 to
December 2020). According to NHS Research Ethics Committee guidelines, this work was classed as service evaluation and formal ethical approval was not required (9).

Clinical practice

The NHS Lanarkshire GCA FTP was supported by two consultant rheumatologists. Referral sources included primary and secondary care. Approximately 30% of referrals were rejected following telephone consultation on grounds of clinical implausibility (e.g. combination of young age, normal inflammatory indicators, clear alternative diagnosis). The remaining patients were assessed in person.

GCAPS evaluation was incorporated into routine clinical practice as an assessment aid and promising pre-test tool to prioritise referrals. All patients underwent USS, with additional tests (e.g. TAB, cross-sectional imaging) arranged at the discretion of the treating consultant. Diagnostic and therapeutic decisions were made on a case-by-case basis by the treating consultant.

US scans were performed by trained sonographers (AC, KD) at time of initial assessment, with GE Logiq S8 (linear probe ML6-15) and GE Logiq e (linear probe L4-12t), using settings recommended by EULAR (1). Non-compressible halo sign and intima media thickness were recorded for the superficial temporal artery (TA), frontal and parietal TA branches, and facial artery (FA), in both longitudinal and transverse planes. Intima media thickness (IMT) was recorded at the carotid, subclavian and axillary arteries.

Definitions

For this study, GCA diagnosis was defined pragmatically by clinician decision to treat as GCA following FTP review, with confirmation 6 months after initial visit. Most cases were supported by additional tests showing objective evidence of cranial and/or large vessel vasculitis, as per BSR and EULAR recommendations (1, 2). In some cases, additional test results were negative or inconclusive, but mitigating circumstances existed to explain non-positive results. Non-GCA diagnoses were also confirmed after 6 months by review of clinical notes.

Positive USS was defined by the presence of non-compressible halo sign (according to OMERACT definitions (10)), associated with decreased echogenicity of cranial arterial walls, and maximal IMT measurement ≥0.5mm for TA, ≥0.4mm for frontal (FB) and parietal (PB) TA branches and FA, and >1mm for subclavian and axillary arteries (11, 12). US scans demonstrating short sections of modestly thickened and hyperechoic IMT, irregular arterial walls (suggestive of atheromatic plaques), or borderline measurements were considered inconclusive (10).

GCAPS is a pre-test probability scoring system comprised of positive and negative integers based on patient demographics, symptoms, signs, laboratory findings, and competing differentials encountered during assessment of suspected GCA. Previous publications have considered scores of <9 to be low risk for GCA, 9 - 12 medium risk, and >12 high risk (4, 5).
Data collection

Patient demographics, time from referral to assessment, duration of steroid treatment, fulfillment of ACR classification criteria for GCA (excluding TAB information) (13), clinical features at presentation including GCAPS components [symptom duration, CRP, cranial pain, constitutional symptoms, polymyalgia rheumatica (PMR), ischaemic symptoms, visual abnormalities, temporal artery abnormalities, extra-cranial vascular abnormalities, cranial nerve palsies, alternative diagnosis more likely], total score, results of USS +/- TAB, and final diagnosis were collected routinely at first visit and follow-up visits and recorded in a local clinical database, anonymised prior to analyses.

Data analysis

Descriptive characteristics were expressed as number (n) (%), mean (SD), or median (IQR), depending on data type and distribution. The performance of GCAPS in predicting final diagnosis of GCA was assessed by Receiver Operating Characteristic (ROC) analyses, and the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of different GCAPS binary cut-off values were calculated. Multivariable logistic regression was used to test the associations between individual GCAPS components (positive integers only) and confirmed GCA. The sensitivity, specificity, PPV, NPV, and accuracy of USS for GCA diagnosis, in the whole group and according to GCAPS risk group, were also calculated.

Patients with ineligible or insufficient data were excluded from analyses. Minimal missing data was anticipated, and no imputation of missing data planned. Analyses were carried out using IBM SPSS Statistics v.21.0 (IBM Corp) and R v.4.1.2 (R Core Team).

RESULTS

Clinical characteristics

133 patients were assessed on the GCA FTP during the study period, of which 129 had evaluable data and were included in analyses (of 4 excluded, 1 was referred with existing GCA diagnosis, 1 was referred following incidental finding of vasculitis on PET-CT, and 2 were lost to follow-up prior to confirmation of diagnosis).

83/129 (64.3%) were female, and mean (SD) age was 69.6 (9.5) years. Median (IQR) time from referral to assessment was 3 (2,5) days. 93 (72.1%) patients had received glucocorticoids for ≥24 hours at time of assessment, with median (IQR) time on steroids of 4 (1,7) days.

In the whole group of 129 patients, GCAPS ranged from 2 – 24 with a median (IQR) score of 10 (8, 15). 41 (31.8%) patients were classed as low risk (GCAPS <9), 40 (31.0%) as medium risk (GCAPS 9-12), and 48 (37.2%) as high risk (GCAPS >12).
Diagnostic pathway

A confirmed diagnosis of GCA was made in 44/129 patients (34.1%). Of these, 33 (75.0%) fulfilled ACR classification criteria for GCA (excluding TAB information). Figure 1 depicts the route to diagnosis for all patients, organised by GCAPS risk group. All 129 patients were assessed by USS, and 27 (20.9%) had TAB. In 40/44 cases, GCA diagnosis was supported by positive USS, positive TAB, or both. Of the remaining 4, 2 were diagnosed clinically, 1 developed scalp necrosis with skin biopsy result suggestive of GCA, and 1 developed relapse after initial assessment following rapid steroid taper, with subsequent USS positive (see Supplementary Table S1, available at Rheumatology Advances in Practice online).

Figure 2 shows the final number of patients with GCA according to risk group. There were no patients with GCA in the low risk group, 3 in medium risk group (7.5%), and 41 in the high-risk group (85.4%).

GCAPS components - overview

Table 1 shows GCAPS components and total scores for the whole group, as well as by GCAPS risk group, and by final GCA diagnosis. Most patients presented with cranial pain (120/129 [93.0%]), with similar percentages across risk groups, and between GCA / non-GCA groups. Only 3/129 patients (2.3%) were <50 years of age, none of whom had GCA. No cases of GCA had CRP <10. Ophthalmological abnormalities (i.e. anterior ischaemic optic neuropathy, central retinal artery occlusion, visual field defect, relative afferent pupillary defect) were rare, occurring in 7/129 patients (5.4%; 2/7 medium risk, 5/7 high risk; 5/7 GCA, 2/7 non-GCA).

Supplementary Table S2, available at Rheumatology Advances in Practice online, shows additional GCAPS data for the whole group and subgroups as per table 1. Extracranial vascular signs (i.e. bruits or loss of pulse) and cranial nerve palsies were rare, affecting 5/129 (3.9%) and 2/129 (1.6%) respectively. Alternative diagnosis more likely than GCA occurred in 57/129 patients (44.2%).

GCAPS components - regression analyses

In multivariable logistic regression analyses, the following GCAPS components were associated with final diagnosis of GCA: increasing age group [OR 4.59 (95% CI 1.17 – 17.95); p=0.03], increasing CRP [OR 7.45 (95% CI 1.80 – 30.88); p=0.006], presence of combined constitutional symptoms (compared to absence of constitutional symptoms) [OR 34.72 (95% CI 2.45 – 491.55); p=0.009], presence of ischaemic symptoms [OR 29.96 (95% CI 3.43 – 262.07); p=0.002], presence of TA tenderness [OR 16.19 (95% CI 1.63 – 160.51; p=0.02], and presence of TA thickening [OR 92.98 (95% CI 4.59 – 1871.49; p=0.003] (both compared to absence of TA changes). Female sex was associated with reduced likelihood of GCA [OR 0.08 (95% CI 0.01 – 0.66); p=0.02] (see Supplementary Table S3, available at Rheumatology Advances in Practice online).
GCAPS performance – ROC analyses

Figure 3 shows the ROC curve for overall performance of GCAPS in discriminating final diagnoses of GCA from non-GCA, with an AUC of 0.976 (95% CI 0.954 – 0.999), indicating excellent diagnostic ability (14). Maximum Youden index (a measure of optimal compromise between sensitivity and specificity) (15) was 0.873, corresponding to a GCAPS binary cut-off value of ≥13.

GCAPS performance – binary cut-offs

Using a GCAPS binary cut-off of ≥9, 44/88 patients were correctly identified as GCA, and 41/41 as non-GCA, giving a sensitivity of 100.0% and specificity of 48.2%. Using a cut-off of ≥10, sensitivity remained 100.0%, but specificity was higher at 67.1%. Maximum accuracy (91.5%) was seen with a cut-off of ≥13 (also the maximum Youden index in ROC analyses). These results, along with additional cut-off values, are detailed in table 2.

Ultrasound performance

In the whole group, 38/39 (97.4%) patients with a positive USS had a final diagnosis of GCA, while 84/90 patients (93.3%) with a negative or inconclusive US had a non-GCA final diagnosis. Comparing positive scans to scans that were negative or inconclusive, USS had an overall sensitivity of 86.4% and specificity of 98.8%.

Sensitivity increased by risk group (incalculable for low risk group as only 1 positive USS and no final diagnoses of GCA, 33.3% for medium risk, 90.2% for high risk). There were 2/41 (4.9%) inconclusive scans in the low risk group, 7/40 (17.5%) in medium risk, and 3/48 (6.3%) in high risk. Diagnostic performance of USS is summarised in Supplementary Table S4, available at Rheumatology Advances in Practice online.

Of patients with a final diagnosis of GCA and positive USS, 31/38 were positive for TA or TA branches, 1/38 was positive for large vessels only, and 6/38 were positive for both TA (or branches) and large vessels. No cases were identified where a GCA diagnosis was based on facial, carotid or subclavian artery measurements alone (data available in 21/38).

Only one patient with a negative USS had a final diagnosis of GCA; this was a clinical diagnosis (TAB inconclusive) (see Supplementary Table S2, available at Rheumatology Advances in Practice online). A further 4 patients with inconclusive USS had final GCA diagnoses. Conversely, only one patient with a positive USS had a final non-GCA diagnosis. This patient had pre-existing rheumatoid arthritis and developed symptoms suggestive of PMR, but only had partial response to moderate dose glucocorticoids. GCAPS was 8. USS assessment of TAs was positive unilaterally, with borderline axillary artery IMT changes. TAB (performed after 3 weeks on glucocorticoids) and PET-CT were negative. The patient was ultimately treated with tocilizumab for active rheumatoid arthritis, with resolution of symptoms, and repeat US was not undertaken.
Temporal artery biopsy results

Of 27 TABs performed, 7 were positive, 18 negative, and 2 inconclusive (see figure 1 for TAB results with corresponding GCAPS risk group, USS result, and final diagnosis). TAB result was consistent with USS result and final diagnosis in 12 cases; 5 had positive USS / positive TAB (all GCA); 7 had negative USS / negative TAB (all non-GCA). Conflicting USS and TAB results were seen in 3 cases; all had positive USS / negative TAB; 2 had a final diagnosis of GCA, 1 non-GCA (= patient described in Ultrasound performance section). No patient had negative USS followed by positive TAB (of 9 performed). In 10 cases, TAB was performed following inconclusive USS; of these, final diagnosis reflected TAB result in 7 cases (2 diagnosed with GCA following positive TAB, 5 non-GCA following negative TAB), but in 3 cases a diagnosis of GCA was reached despite negative TAB (see Supplementary Table S2, available at Rheumatology Advances in Practice online). TAB was inconclusive in 2 cases, both following negative USS in GCAPS high risk patients; 1 had a final diagnosis of GCA (see Supplementary Table S2), 1 non-GCA.

DISCUSSION

This is the first published study of application of GCAPS outside the specialist centre where it was developed. In our setting, a newly established GCA FTP where GCAPS was incorporated as part of standard clinical assessment, overall performance of GCAPS in predicting final GCA diagnoses was excellent. Prevalence of GCA increased across GCAPS risk groups, with none in the low risk group, 7.5% in the medium risk group, and 85.4% in the high risk group, indicating effective stratification at the pre-assessment stage. These results are comparable with previous publications from Southend (4, 5) and provide external validation to support the use of GCAPS as a tool for prioritisation of referrals in clinical practice.

Furthermore, the absence of GCA among low risk patients (i.e. GCAPS <9) in our cohort lends weight to the suggestion that GCAPS alone may be sufficient to exclude GCA. Of more than 350 patients included in the recent study by Sebastian et al, none with GCAPS <9 ultimately had GCA (5). These results suggest it may be feasible to adopt a GCAPS binary cut-off when accepting or rejecting referrals to the GCA FTP, thus avoiding imaging and/or specialist review of low risk patients. This would help focus resources on those truly at risk and facilitate rapid review. Indeed, our data suggest a score of <10 (not GCA) or ≥10 (possible GCA) would be optimal for this purpose, as all GCA cases were captured (i.e. 100% sensitivity), but with fewer “false positives” than GCAPS <9 (i.e. higher specificity). Early results from an ongoing prospective multi-centre study, presented at EULAR Congress 2021, also indicated an optimal cut-off point ≥10 (100% sensitivity, 69% specificity) (16), and authors involved in a GCA FTP in Luton have suggested applying this value when accepting referrals (7).

A GCAPS cut-off of ≥13 was “optimal” in our cohort in terms of maximum combined sensitivity and specificity (93.2% and 90.6% respectively), and maximum accuracy (91.5%), meaning greatest number of correct final GCA diagnoses. While less appropriate for screening referrals (where maximising sensitivity helps avoid missed diagnoses), GCAPS
≥13 might be a pragmatic threshold for initiation of empirical glucocorticoid treatment, with rapid GCA FTP review prior to glucocorticoid initiation for GCAPS 10-12. Such an approach would minimise over-treatment, increase confidence in interpretation of negative clinical and ultrasound findings, and carry a low complication risk.

Compared to the recent study by Sebastian et al, a higher proportion of our patients had a final diagnosis of GCA (34% versus 25%), and our distribution across GCAPS risk groups was shifted to the higher end (32% low risk, 31% medium risk, 37% high risk in our study, versus 43% low risk, 39% medium risk, 18% high risk in the previous study) (5). These discrepancies may reflect a more stringent approach to screening referrals to the GCA FTP in our centre, with more rejected based on clinician judgement at the pre-assessment stage.

We found several individual GCAPS components associated with final diagnosis of GCA, including age, CRP, constitutional symptoms (>1), ischaemic symptoms, TA tenderness, and TA thickening. Cranial pain was non-significant, as most GCA FTP patients exhibited this symptom, meaning similar representation in GCA and non-GCA groups. Female sex was associated with lower likelihood of GCA, despite a higher GCAPS value. While more females than males ultimately had GCA in our cohort, a higher proportion of females had non-GCA final diagnoses. These findings are broadly consistent with a recent systematic review of diagnostic accuracy of clinical features in GCA, where ischaemic symptoms, abnormalities of TA, and high ESR were associated with GCA, while younger age (<70) and normal CRP were associated with absence of GCA; interestingly sex was not found to affect GCA likelihood in this study (17). In our cohort, rarer manifestations, such as a visual abnormalities, extracranial vascular abnormalities, loss of TA pulse, and cranial nerve palsies, were not associated with GCA, likely because of low event rates, and/or because these events provoked referral for exclusion of GCA, with some found to have alternative causes. Our regression analyses were limited by sample size (which also explains the broad confidence intervals), but are nonetheless of interest to those involved in clinical assessment.

The strengths of our study include that GCAPS assessment was contemporaneous and prospectively documented, reducing the risk of recording bias inherent in retrospective analyses. Additionally, our GCA / non-GCA final diagnoses were supported by objective evidence in almost all cases, were confirmed at 6 months, and were unaffected by missing data or loss to follow-up. We are therefore confident in the internal validity of our main exposure and outcome variables.

All patients underwent USS in our study, which is a further strength in permitting a full assessment of USS performance in relation to GCAPS. Overall, USS performed excellently as a diagnostic tool for GCA, consistent with its established role in clinical practice (12, 18). We included facial, carotid and subclavian artery scanning as routine, but did not identify any cases dependent on these measurements alone, suggesting results should be applicable to centres conducting standard temporal and axillary scans (per (1)). Specificity appeared high in all risk groups, suggesting USS is useful for excluding GCA regardless of pre-test probability. Specificity values were subject to overestimation, however, as we grouped negative and inconclusive scans together in these analyses (19). This also contributed to a lower sensitivity estimate for the medium risk group (33%), which had the most inconclusive scans. Nonetheless, sensitivity increased steeply across risk groups; indeed, the positive diagnostic yield in the low risk group was nil, as the only positive USS was a false-positive

https://mc.manuscriptcentral.com/rheumap
that led to further tests including TAB and PET-CT (albeit this was an unusual case with possible overlapping final diagnoses). This strengthens our argument against routine USS examination for low risk patients.

TAB was useful in selected cases to confirm or exclude GCA, particularly where USS was inconclusive. Several TABs performed following clear positive or negative USS appeared diagnostically superfluous. This likely reflects the natural learning curve associated with a newly established FTP (including use of TAB as a reference test in the early stages) and increasing confidence in USS over time.

Limitations largely reflect the real-world setting. Temporal arteries were assessed with a 13MHz probe in a proportion of patients, which is below the recently recommended frequency of $\geq 15MHz$, and associated with reduced resolution of ultrasound image. IMT cut-offs used routinely in our centre are adapted from published data (11, 12), rounded up for smaller vessels to account for probe availability. While this could increase false negative rates, diagnoses were not dependent on USS alone, and no missed diagnoses were identified at follow-up. Median time to assessment was 3 days and the majority of patients were already treated with glucocorticoids, which is known to reduce the diagnostic value of USS (20), and may also explain our inconclusive scan rate. USS assessors were not blinded to GCAPS, meaning prior knowledge could have influenced USS interpretation and GCA diagnosis. However, conferring pre-test probability is part of the function of GCAPS, and USS interpretation was guided by international standards. A small number of GCA diagnoses were not supported by positive USS or TAB, but we felt a pragmatic definition that incorporated convincing clinical cases was appropriate for the study type. Our results came from a single centre, but enhance the applicability of GCAPS overall by demonstrating effective performance in a newly established GCA FTP, distinct from the original specialist unit.

Introduction of GCAPS to GCA FTP referral pathways may help enhance the quality and appropriateness of referrals from primary care (21). This would likely require training for referrers and capacity for dialogue with specialists, to mitigate the possibility of inadequate assessment and underestimation of risk. In NHS Lanarkshire we have conducted an educational meeting for general practitioners to introduce the concept of GCAPS. While some GCAPS components are unambiguous, others (e.g. jaw claudication, TA assessment, alternative diagnoses) may be subject to misinterpretation. In our experience, uncertainty tends towards over-scoring and unnecessary referral and treatment, rather than missed diagnoses. Additional validation of GCAPS performance when used in this setting would be desirable.

In summary, GCAPS performed excellently in stratifying patients by risk, and discriminating GCA from non-GCA final diagnoses. Our results provide external validation of existing data from Southend. GCAPS appears sufficient to exclude GCA in low risk cases, and so may have clinical relevance beyond pre-test probability, i.e. as a diagnostic tool. Moving forward, use of a GCAPS binary cut-off as an entry requirement to the GCA FTP might reduce the need for imaging +/- specialist review in a proportion of patients.
Funding

No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this article.

Disclosure Statement

K.D. has received speaker fees from Menarini Pharma UK. A.C. has received speaker, travel and registration fee sponsorship from Novartis, Abbvie, Chugai Pharma, Colgene, Roche and Lilly. The remaining authors have declared no conflicts of interest.

Data availability statement

Data underlying this article will be shared on reasonable request to the corresponding author.
REFERENCES


Table 1. GCAPS total score and selected components for all patients, by risk group, and by final GCA diagnosis.

<table>
<thead>
<tr>
<th></th>
<th>All (n=129)</th>
<th>Low risk (n=41)</th>
<th>Med risk (n=40)</th>
<th>High risk (n=48)</th>
<th>GCA (n=44)</th>
<th>Not GCA (n=85)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCAPS (median, IQR)</td>
<td>10 (8, 15)</td>
<td>6 (5.5, 7.5)</td>
<td>10 (9, 11)</td>
<td>16 (14, 18)</td>
<td>16 (14, 18)</td>
<td>9 (6, 11)</td>
</tr>
<tr>
<td>Age:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- &lt;50 (n, %)</td>
<td>3 (2.3)</td>
<td>2 (4.9)</td>
<td>1 (2.5)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>3 (3.5)</td>
</tr>
<tr>
<td>- 50-60 (n, %)</td>
<td>19 (14.7)</td>
<td>5 (12.2)</td>
<td>9 (22.5)</td>
<td>5 (10.4)</td>
<td>3 (6.8)</td>
<td>16 (18.8)</td>
</tr>
<tr>
<td>- 61-65 (n, %)</td>
<td>21 (16.3)</td>
<td>13 (31.7)</td>
<td>1 (2.5)</td>
<td>7 (14.6)</td>
<td>6 (13.6)</td>
<td>15 (17.6)</td>
</tr>
<tr>
<td>- &gt;65 (n, %)</td>
<td>86 (66.7)</td>
<td>21 (51.2)</td>
<td>29 (72.5)</td>
<td>36 (75.0)</td>
<td>35 (79.5)</td>
<td>51 (60.0)</td>
</tr>
<tr>
<td>Female (n, %)</td>
<td>83 (64.3)</td>
<td>22 (53.7)</td>
<td>31 (77.5)</td>
<td>30 (62.5)</td>
<td>24 (54.5)</td>
<td>59 (69.4)</td>
</tr>
<tr>
<td>CRP:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- 0-5 (n, %)</td>
<td>30 (23.4)</td>
<td>19 (46.3)</td>
<td>11 (27.5)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>30 (35.3)</td>
</tr>
<tr>
<td>- 6-10 (n, %)</td>
<td>8 (6.3)</td>
<td>4 (9.8)</td>
<td>3 (7.5)</td>
<td>1 (2.1)</td>
<td>0 (0)</td>
<td>8 (9.4)</td>
</tr>
<tr>
<td>- 11-25 (n, %)</td>
<td>22 (17.2)</td>
<td>11 (26.8)</td>
<td>10 (25.0)</td>
<td>1 (2.1)</td>
<td>2 (4.7)</td>
<td>20 (23.5)</td>
</tr>
<tr>
<td>- &gt;25 (n, %)</td>
<td>68 (53.1)</td>
<td>7 (17.1)</td>
<td>16 (40.0)</td>
<td>45 (95.7)</td>
<td>41 (95.3)</td>
<td>27 (31.8)</td>
</tr>
<tr>
<td>Cranial pain (n, %)</td>
<td>120 (93.0)</td>
<td>38 (92.7)</td>
<td>36 (90.0)</td>
<td>46 (95.8)</td>
<td>42 (95.5)</td>
<td>78 (91.8)</td>
</tr>
<tr>
<td>ConstitutionalP:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- single (n, %)</td>
<td>35 (27.1)</td>
<td>5 (12.2)</td>
<td>11 (27.5)</td>
<td>19 (39.6)</td>
<td>16 (36.4)</td>
<td>19 (22.4)</td>
</tr>
<tr>
<td>- combined (n, %)</td>
<td>19 (14.7)</td>
<td>1 (2.4)</td>
<td>4 (10.0)</td>
<td>14 (29.2)</td>
<td>14 (31.8)</td>
<td>5 (5.9)</td>
</tr>
<tr>
<td>PMR (n, %)</td>
<td>27 (20.9)</td>
<td>1 (2.4)</td>
<td>10 (25.0)</td>
<td>16 (33.3)</td>
<td>16 (36.4)</td>
<td>11 (12.9)</td>
</tr>
<tr>
<td>Ischaemicc (n, %)</td>
<td>45 (34.9)</td>
<td>3 (7.3)</td>
<td>9 (22.5)</td>
<td>33 (68.8)</td>
<td>30 (68.2)</td>
<td>15 (17.6)</td>
</tr>
<tr>
<td>Visuald (n, %)</td>
<td>7 (5.4)</td>
<td>0 (0)</td>
<td>2 (5.0)</td>
<td>5 (10.4)</td>
<td>5 (11.4)</td>
<td>2 (2.4)</td>
</tr>
<tr>
<td>Temporal artery:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- tender (n, %)</td>
<td>24 (18.6)</td>
<td>7 (17.1)</td>
<td>4 (10.0)</td>
<td>13 (27.1)</td>
<td>12 (27.3)</td>
<td>12 (14.1)</td>
</tr>
<tr>
<td>- thickened (n, %)</td>
<td>10 (7.8)</td>
<td>0 (0)</td>
<td>1 (2.5)</td>
<td>9 (18.8)</td>
<td>9 (20.5)</td>
<td>1 (1.2)</td>
</tr>
<tr>
<td>- pulseless (n, %)</td>
<td>6 (4.7)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>6 (12.5)</td>
<td>6 (13.6)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

GCAPS = Giant cell arteritis probability score. PMR = symptoms suggestive of polymyalgia rheumatica.

a) One missing CRP value (high risk / GCA groups). b) Constitutional symptoms = night sweats, weight loss, fever. c) Ischaemic symptoms = jaw / tongue claudication, unioocular blurring, diplia, amaurosis fugax. d) Ophthalmological abnormality = anterior ischaemic optic neuropathy, central retinal artery occlusion, visual field defect, relative afferent pupillary defect.
Table 2. Diagnostic performance of different GCAPS binary cut-off values.

<table>
<thead>
<tr>
<th>GCAPS cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥8*</td>
<td>100.0%</td>
<td>36.5%</td>
<td>44.9%</td>
<td>100.0%</td>
<td>58.1%</td>
</tr>
<tr>
<td>≥9</td>
<td>100.0%</td>
<td>48.2%</td>
<td>50.0%</td>
<td>100.0%</td>
<td>65.9%</td>
</tr>
<tr>
<td>≥10</td>
<td>100.0%</td>
<td>67.1%</td>
<td>61.1%</td>
<td>100.0%</td>
<td>78.3%</td>
</tr>
<tr>
<td>≥11</td>
<td>95.5%</td>
<td>74.1%</td>
<td>65.6%</td>
<td>96.9%</td>
<td>81.4%</td>
</tr>
<tr>
<td>≥12</td>
<td>93.2%</td>
<td>83.5%</td>
<td>74.5%</td>
<td>95.9%</td>
<td>86.8%</td>
</tr>
<tr>
<td>≥13</td>
<td>93.2%</td>
<td>90.6%</td>
<td>83.7%</td>
<td>96.3%</td>
<td>91.5%</td>
</tr>
<tr>
<td>≥14</td>
<td>79.5%</td>
<td>97.6%</td>
<td>94.6%</td>
<td>90.2%</td>
<td>91.5%</td>
</tr>
<tr>
<td>≥15</td>
<td>72.7%</td>
<td>98.8%</td>
<td>97.0%</td>
<td>87.5%</td>
<td>89.9%</td>
</tr>
<tr>
<td>≥16</td>
<td>61.4%</td>
<td>100%</td>
<td>100%</td>
<td>83.3%</td>
<td>86.8%</td>
</tr>
</tbody>
</table>

GCAPS = Giant cell arteritis probability score. PPV = Positive predictive value. NPV = Negative predictive value. a) GCAPS cut-off of ≥8 means a score of ≥8 is considered “test positive” and a score of <8 is considered “test negative”.
Figure legends.

Figure 1. Flow chart showing route to diagnosis for all patients, according to GCAPS risk group. GCAPS = Giant cell arteritis probability score.

Figure 2. Bar chart displaying final GCA diagnosis by GCAPS risk group. GCAPS = Giant cell arteritis probability score.

Figure 3. Receiver operating characteristic curve for total GCAPS in predicting final diagnosis of GCA. AUC = 0.976 (95% CI 0.954 – 0.999). Maximum Youden index (sensitivity + specificity – 1) = 0.873, corresponding to a GCAPS binary cut-off value of \( \geq 13 \). GCAPS = Giant cell arteritis probability score. US = ultrasound. TAB = temporal artery biopsy. ND = Not done.
Figure 1. Flow chart showing route to diagnosis for all patients, according to GCAPS risk group.

247x129mm (330 x 330 DPI)
Figure 2. Bar chart displaying final GCA diagnosis by GCAPS risk group.

224x131mm (300 x 300 DPI)
Caption: Figure 3. Receiver operating characteristic curve for total GCAPS in predicting final diagnosis of GCA. AUC = 0.976 (95% CI 0.954 – 0.999). Maximum Youden index (sensitivity + specificity – 1) = 0.873, corresponding to a GCAPS binary cut-off value of ≥13.