Temporal Artery Biopsy: When Is It Worth the Headache?

Bronson T. Bomkamp, MD; Andrew J. Borgert, PhD; Alec J. Fitzsimmons, MPH; Travis J. Smith, MD; Irina Shakhnovich, MD; Jacob R. Peschman, MD

ABSTRACT

Introduction: Temporal artery biopsy is ordered when clinical symptoms and elevated C-reactive protein values and/or erythrocyte sedimentation rates suggest giant cell arteritis. The percentage of temporal artery biopsies positive for giant cell arteritis is low. The objectives of our study were to analyze the diagnostic yield of temporal artery biopsies at an independent academic medical center and to develop a risk stratification model for triaging patients for possible temporal artery biopsy.

Methods: We retrospectively reviewed the electronic health records of all patients who underwent temporal artery biopsy in our institution from January 2010 through February 2020. We compared clinical symptoms and inflammatory marker (C-reactive protein and erythrocyte sedimentation rate) values of patients whose specimens were positive for giant cell arteritis with those of patients with negative specimens. Statistical analysis included descriptive statistics, chi-square test, and multivariable logistic regression. A risk stratification tool, which included point assignments and measures of performance, was developed.

Results: Of 497 temporal artery biopsies for giant cell arteritis performed, 66 were positive and 431 were negative. Jaw/tongue claudication, elevated inflammatory marker values, and age were associated with a positive result. Using our risk stratification tool, 3.4% of low-risk patients, 14.5% of medium-risk patients, and 43.9% of high-risk patients were positive for giant cell arteritis.

Conclusions: Jaw/tongue claudication, age, and elevated inflammatory markers were associated with positive biopsy results. Our diagnostic yield was much lower when compared with a benchmark yield determined in a published systematic review. A risk stratification tool was developed based on age and the presence of independent risk factors.

• • •

Author Affiliations: Department of Medical Education, Gundersen Health System, La Crosse, Wisconsin (Bomkamp); Department of Medical Research, Gundersen Health System, La Crosse, Wis (Borgert, Fitzsimmons), Department of Surgery, Gundersen Health System, La Crosse, Wis (Smith, Shakhnovich); Department of Surgery, The Medical College of Wisconsin, Milwaukee, Wis (Peschman).

Corresponding Author: Bronson T. Bomkamp, MD, Mail Stop: C05-001, Department of Medical Education, Gundersen Health System, 1900 South Ave, La Crosse, WI 54601; phone 608.775.2331; email btbomkam@gundersenhealth.org.

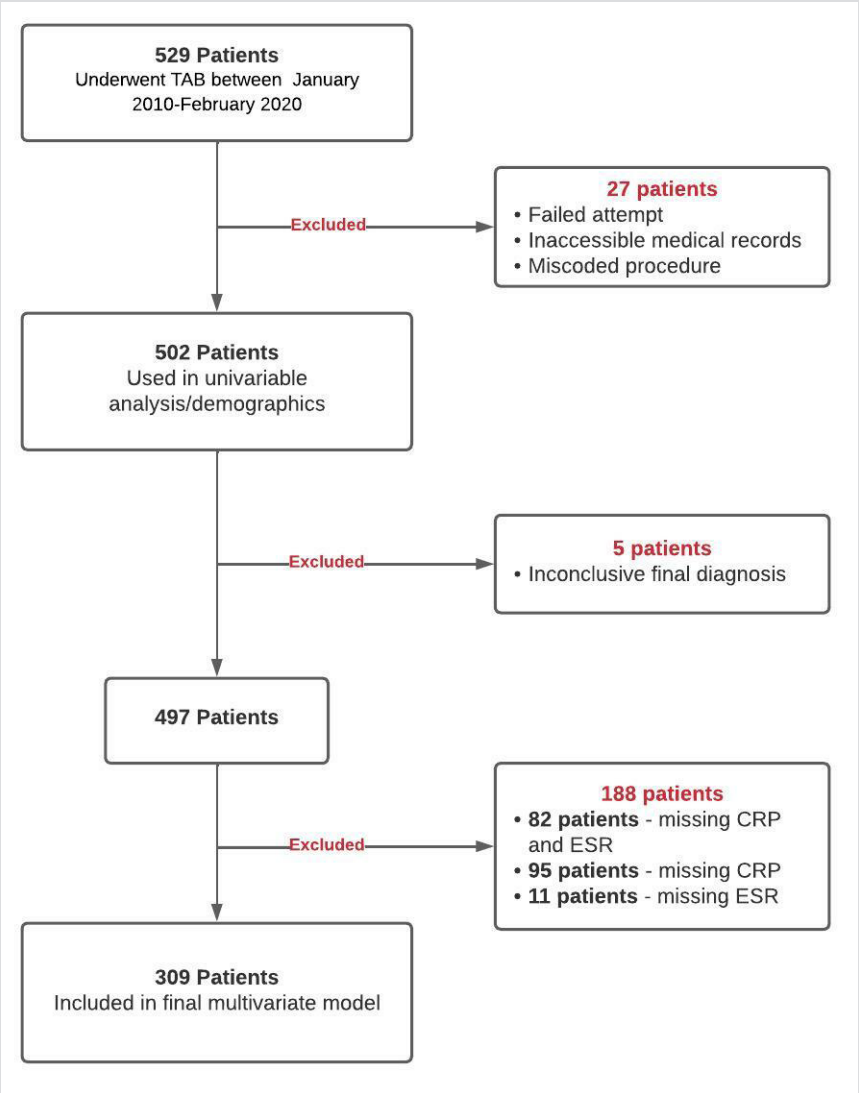
INTRODUCTION

Giant cell arteritis (GCA), also known as temporal arteritis, is an inflammatory vasculopathy that affects large- to medium-sized vessels, most often affecting patients of advanced age.[1] The pathophysiology involves the infiltration of giant cells into the vessel walls. These giant cells are created due to granulomatous changes of CD4+ T lymphocytes and macrophages.[2] A variety of symptoms have been reported in patients with temporal arteritis, including constitutional symptoms (fever, fatigue, and weight loss), headaches, jaw or tongue claudication, visual symptoms (transient vs permanent visual loss, diplopia, hallucinations), and musculoskeletal symptoms. If left untreated, GCA may progress to permanent blindness; therefore, early identification and treatment are paramount to evaluating patients with the potential of having the disease. Yet GCA diagnosis is challenging. Temporal artery biopsy (TAB) is the best confirmatory test, but the false-negative rate of biopsy is estimated to vary between 10% and 61%.[3]

GCA can manifest as extracranial disease, also known as large-vessel GCA. Studies have shown that large-vessel involvement is present in more than 80% of patients with GCA,[4] and patients with GCA are 17 times more likely than age/sex-matched patients without GCA to develop aneurysms of the thoracic aorta.[5] Evaluation for large-vessel GCA is different from that for temporal arteritis and was beyond the scope of this project, which focused on TAB for GCA.

The diagnosis of GCA has several considerations. In 1990, in an effort to standardize its diagnosis, the American College of Rheumatology (ACR) developed 5 criteria to distinguish GCA from other forms of vasculitis, criteria that have subsequently been used to assist with GCA diagnosis: over 50 years of age, new-onset headache, temporal artery tenderness or decreased temporal artery pulse, erythrocyte sedimentation rate (ESR) greater than 50 mm/hour, and arterial biopsy showing necrotizing arteritis characterized by the predominance of mononuclear cell infiltrates or a granulomatous process with multinucleated giant cells. When 3 of 5 criteria were met, their sensitivity and specificity were 93.5% and 91.2%, respectively.[6] ACR criteria do not require TAB to make the diagnosis. A Kaiser Permanente study published in 2013 developed a new protocol for temporal arteritis evaluation that included an enhanced clinical evaluation and the incorporation of color duplex ultrasonography and biopsy when appropriate. Using this protocol, their TAB yield increased from 8.5% to 24%.[7]

Figure 1. Patient Inclusion and Exclusion Flow Diagram



Abbreviations: TAB, temporal artery biopsy; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate.

The presence of "skip lesions" and inadequate specimen length may contribute to the false-negative rate of TAB. A recent 10-year retrospective review with over 1000 specimens published in 2020 suggested an optimal prefixation biopsy specimen length of 1.5 to 2 cm.[8] The studies included in a systematic review of 113 articles about GCA published in 2018 had a high degree of heterogeneity, rendering a meta-analysis unsuitable.[9] However, the median yield of TAB was 25% (95% CI, 0.21-0.27), which the authors concluded would provide a benchmark to determine whether TAB is under- or overutilized.

Most biopsies are performed in the clinic setting within 72 hours of receiving the biopsy request. Timing is crucial to improve diagnostic yield because significant histological findings can resolve after 2 weeks of steroid use.[10]

The primary objective of this study was to evaluate the yield of TAB at our institution and to compare it with the reported benchmark of 25% biopsies positive for GCA. Our secondary objective was to create a risk stratification tool with management recommendations in conjunction with our Department of Rheumatology. Our working group agreed that our tool would be designed to favor a negative predictive value, given the risk of untreated GCA. The intent of this tool and the associated management recommendations were focused on improving care for patients in 2 ways: (1) to identify those patients for whom biopsy results—whether positive or negative—would be unlikely to alter management, thereby obviating the need for the procedure and avoiding its associated risks and costs, and (2) to help limit the prescribing of unnecessary steroids, given their potential for long-term complications.

METHODS

Following Institutional Review Board approval, we reviewed the electronic health records (EHR) of patients who underwent TAB from January 2010 through February 2020 in our medical center. TABs were performed either in the clinic or the operating room by general or vascular surgeons working with general surgery residents. The operative technique employed is consistent within the

Table 1. Preoperative Patient Characteristics and Postoperative Diagnoses of Positive and Negative Temporal Artery Biopsy Cases

Demographic and Clinical Characteristics	Positive (n = 66)	Negative (n = 431)	*P* value
Demographic			
Age, mean (SD), y	74.6 (8.5)	71.8 (11.5)	0.1
Sex, no. (%)			0.6
Men	26 (39.4)	154 (35.7)	
Women	40 (60.6)	277 (64.3)	
BMI, mean (SD), kg/m^2	28.0 (6.1)	31.0 (8.2)	< 0.01
Laboratory values			
CRP (n = 320), mean (SD), mg/L	8.0 (7.7)	4.3 (6.0)	< .001
ESR (n = 404), mean (SD), mm/h	72.9 (29.7)	55.7 (33.8)	< 0.01
TAB setting, no. (%)			0.8
Clinic	63 (95.5)	402 (93.3)	
Operating room	3 (4.6)	29 (6.7)	
Surgical specimen length, mean (SD), mm	1.5 (0.7)	1.5 (0.7)	0.5
Symptoms, no. (%)			
Localized headache	51 (77.3)	330 (76.6)	0.9
Scalp tenderness	3 (4.5)	43 (10.0)	0.2
Jaw or tongue claudication	25 (37.9)	82 (19.0)	< 0.01
Visual abnormalities	30 (45.5)	178 (41.3)	0.5
Biopsies, no. (%)			0.1
Unilateral	55 (83.3)	391 (90.7)	
Bilateral	11 (16.7)	40 (9.3)	

Abbreviations: BMI, body mass index; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; TAB, temporal artery biopsy.
Note: All percentages may not total 100% due to rounding; symptom percentages will not equal 100% because some patients had more than 1 symptom.

practice and is based on the steps outlined in "Bedside Procedures for General Surgeons: Part 2."[11]

Commonly reported symptoms that have been identified as concerning for temporal arteritis in previous studies, including localized headache, scalp tenderness, jaw/tongue claudication, and visual abnormalities, were captured from the EHRs of study patients.[4,7] Symptoms that were not specifically documented in the clinician note were considered absent and coded as not present. Laboratory values obtained included ESR and C-reactive protein (CRP) performed within 2 weeks before or 3 days after the biopsy. Procedural information, including the anatomical location of the biopsy (laterality), the setting where the biopsy was performed (clinic vs operating room), and specimen length, also were included for analysis, as were complications within 30 days, including bleeding and surgical site infection.

Descriptive statistics were used to report biopsy positivity and complication rates. Clinical, demographic, and outcome data were compared using chi-square and Fisher exact tests for categorical data, trend tests for ordinal categorical data, and Wilcoxon rank-sum tests for continuous data. After converting age and laboratory values into ordinal categories, factors showing significant univariate associations with the final clinical diagnosis were included in a multivariate logistic regression model as potential covariates. An iterative cross-validation process was employed to construct the final multivariate model. Candidate models were created by adding potential covariates to a logistic regression model in descending order of the strength of association between the covariate and biopsy status via univariable analysis, starting with the covariate with the strongest association. In each iteration, an additional candidate variable was added to the logistic regression model, and a 4-fold, 25-repetition cross-validation procedure was executed whereby the dataset was split into 4 groups; a model was trained on 3 of the 4 groups and then evaluated against the single remaining group via receiver operating characteristic area under the curve (AUC) analysis. This was repeated 25 times for all 4 possible training group combinations (100 total steps), and the average AUC was calculated for this iteration. The covariates included in the iteration with the highest average AUC were selected as the final model.

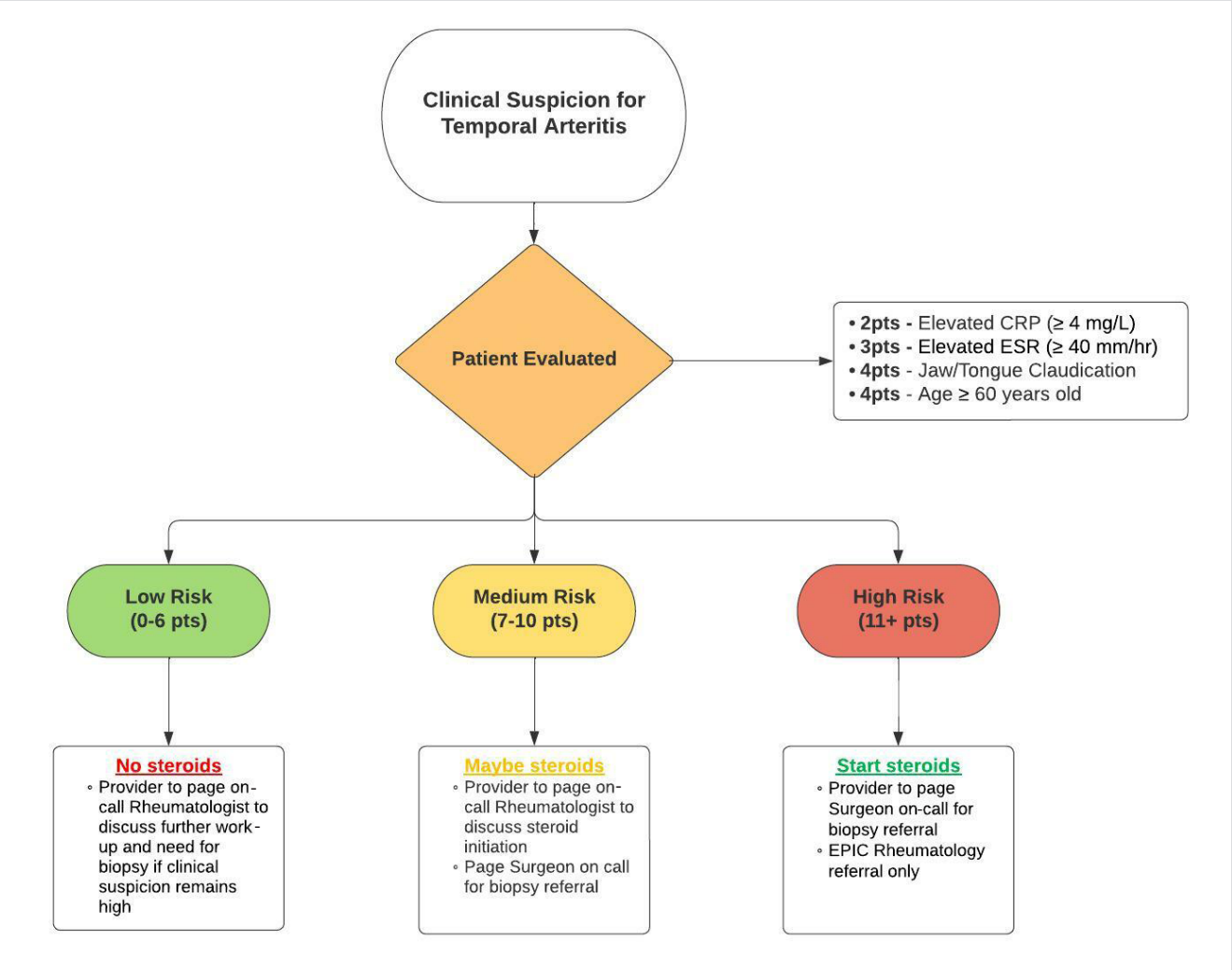
A points-based risk score was created from the final multivariate model by assigning points to a risk factor based on the resulting odds ratio for that factor, rounded to the nearest whole number. Point totals were then categorized as low-, medium-, or high-risk based on the proportion of positive clinical diagnoses associated with those point totals. Final model performance was assessed via receiver operating characteristic analysis and calculation of standard binary classification performance metrics using a simplified low- versus medium/high-risk classification system. All analyses were performed using the SAS 9.4 software suite (Cary, North Carolina).

RESULTS

During the study period, 529 patients underwent biopsy procedures. Twenty-seven were aborted before obtaining a specimen, 3 were inconclusive, 1 had reported evidence of prior arteritis, and 1 vein specimen. Of the 497 remaining biopsy specimens, 66 were positive for GCA and 431 were negative; 188 were excluded from the final multivariate analysis for lack of an ESR and/or CRP result (Figure 1).

Demographic and clinical data are provided in Table 1. Sixty-four percent of study patients were women. The study population's mean age was 72.2 ± 11.2 years. The majority of biopsies were performed in the clinic (465/497; 94.4 %), and the remaining 32 (6.6%) were performed in the operating room. Pathologic positive diagnosis was identified in 66 (13.2%) specimens. Jaw/tongue claudication was significantly associated with a TAB positive for GCA, as were elevated CRP (≥ 4 mg/L) and ESR (≥ 40 mm/hr.).

Figure 2. Current Assessment and Recommendation Workflow Utilizing Risk Stratification Tool



Abbreviations: CRP, C-reactive protein; ESR, ethryocyte sedimentation rate; pts, points.

Ninety-seven percent of biopsies positive for GCA were acquired from patients 60 years of age or older. Subgroup analysis by period (2010-2013, 2014-2016, and 2017-2020) to exclude potential biases of practice pattern change over time revealed no significant differences in positivity rate or exclusion rate from risk stratification development. Although lower body mass index was associated with a positive result, it was no longer significant after our multivariate analysis.

Using a cross-validation model selection procedure, a final multivariate logistic regression model was constructed that included the covariates age, jaw/tongue claudication, ESR, and CRP. From the cross-validation procedure, the average

Table 2. Risk Stratification Tool Development and Performance Assessment

Risk Factor Point Assignment			
Risk factor	**Odds Ratio**	***P* value**	**Points Assigned**
CRP ≥ 4 mg/L	1.98 (0.91-4.28)	0.0837	+2
ESR ≥ 40 mm/h	3.10 (1.14-8.46)	0.0272	+3
Jaw or tongue claudication	3.63 (1.78-7.38)	0.0004	+4
Age ≥ 60 years	3.94 (0.87-17.81)	0.0748	+4
Model Risk Status vs Biopsy Outcome			
	Negative, n	**Positive, n**	**Total, n**
Low risk (≤ 6 points)	112	4	116
Elevated risk (7+ points)	153	40	193
Total	265	44	309
Measures of Model Performance, %			
Negative predictive value	96.6		
Positive predictive value	20.7		
Sensitivity	90.9		
Specificity	42.3		
Accuracy	49.2		

Abbreviations: CRP, C-reactive protein; ESR, erythrocyte sedimentation rate.

AUC for this model was 0.75 ± 0.08. Using these elements, we created our risk stratification tool (Figure 2) using odds ratios, as described above, after excluding patients with incomplete data points. Two points were given for an elevated CRP concentration and 3 points for an elevated ESR. Four points were given for jaw/tongue claudication and 4 points for age 60 years or older. Patients were then divided into low-risk (0-6 points), medium-risk (7-10 points), and high-risk (11+ points) risk groups (Table 2). We placed our patient data into our algorithm to validate our points-based risk scoring system. Of the original 502 patients, 309 were evaluated using our new algorithm. The model's negative predictive value was 96.6%, and its positive predictive value was 20.7% when comparing low-risk patients with any medium- or high-risk patients. When further assigning the 309 patients to low-, medium-, and high-risk categories based upon the points assigned in our algorithm, we found TAB positive for GCA in 3.4% (4/116) of low-risk, 14.5% (22/152) of medium-risk, and 43.9% (18/41) of high-risk patients.

DISCUSSION

GCA is a severe disease potentially leading to permanent blindness. There are no universal protocols available, and diagnosis remains a challenge. Although ACR classification criteria do not require TAB, most rheumatology clinicians favor pathological confirmation of the disease. Our primary goal was to identify our institution's TAB yield, which we determined to be approximately 13%. This study identified specific clinical symptoms, laboratory values, and age values based on data from our EHR system that help categorize patients as at low-, medium-, or high-risk of having a TAB positive for GCA and created a point-based clinical tool to guide patient care (Figure 2). In addition to helping avoid low-yield biopsies, it also helps referring clinicians identify specific laboratory tests and clinical symptoms to include in their workup, with the ultimate score guiding decision-making on steroid use and need for rheumatology collaboration and biopsy referral.

Using the published systematic review with a median yield of 25% as benchmark, we have an opportunity to improve our diagnostic approach, as well as the potential to decrease unnecessary biopsies and their associated comorbidities, reduce the use of steroids, and decrease the cost. Although the overperformance of TABs could be attributable to many factors, one factor we have identified in our health system is the lack of a consistent diagnostic approach. Prior studies have reported that biopsy yield is improved when a referral is made utilizing a multidisciplinary approach.[5] We were unable to identify yield rates by referring clinicians. However, reports from our surgeons suggest that rheumatologists are more likely than those in other specialties to have completed a workup that includes considering all the potential risk factors we have identified in our study. This is evidenced by nearly 40% of the patients who underwent biopsy being excluded from the risk stratification tool due to the lack of a CRP or ESR drawn during the diagnostic workup.

In addition to directing the next steps for the at-risk patient, our algorithm will ensure that patients have the appropriate laboratory workup before undergoing a biopsy. It also elicits direct communication with Rheumatology, which is actively involved in the entire process of diagnosis and treatment. They review every pathology report and clinical context and, ultimately, determine if a patient will come off steroids or if biologic therapy is necessary. Every therapy is individualized, and TAB is an additional clinical tool for them. A high-quality negative biopsy that is not degraded by steroids is the gold standard for a rheumatologist. They can then taper off steroids with clinical confidence while knowing that the patient is unlikely to lose vision. Neither TABs nor steroids are entirely benign. Although TAB is a minor procedure and complications are very low, cost and morbidities, such as discomfort at the biopsy site and bleeding, are associated with the procedure.

We acknowledge that low risk does not equate to no risk, and that the 3.4% of patients so categorized could have serious consequences. Thus, they will be followed closely by the rheumatologist. If their clinical concern remains high, they will be referred for a biopsy and treated at clinician discretion. The on-call rheumatologist is contacted in all cases suggestive of GCA and is involved in the entire course of diagnosis and treatment.

Our future direction will be to institute our risk stratification tool and prospectively track patients to further refine this protocol. This will include the addition of other adjunct diagnostic tools currently not employed by our institution. The first is color duplex ultrasonography.[12] For example, low-risk patients with normal ultrasonographic findings will not need a confirmatory test with TAB to rule out GCA, while high-risk patients with abnormal ultrasonographic findings will have a GCA diagnosis and avoid TAB. Patients with medium risk will undergo TAB. Ultrasonography can provide further guidance when clinical suspicion and biopsy results do not correlate—for example, when skip lesions or disease phenotype spares cranial branches but can be seen in axillary arteries. In addition, eye examinations are not currently a routine referral at our institution for patients with symptoms suggestive of GCA and, therefore, could not be included in our analysis but may be helpful in future models. Finally, we will also need to better follow all patients with a TAB negative for GCA result to understand which may continue treatment with steroids due to high clinical suspicion and which go on to potentially have worsening symptoms and need a repeat biopsy or steroids due to presumed false-negative initial biopsy results.

This study is not without limitations. First, we started with 529 patients, 27 of whom were excluded due to failed biopsy attempts, incomplete or inaccessible medical records, or miscoded procedures. Five patients were excluded owing to inconclusive diagnoses, and 188 patients were excluded from the final multivariate

analysis and risk stratification tool owing to incomplete laboratory workup. Year of diagnosis and TAB exclusion rate were not significantly associated. Some studies have found that bilateral TAB has an increased yield compared with unilateral.[13] In our study, the difference in yield of bilateral biopsies versus unilateral biopsies was approaching significance. As mentioned previously, we did not capture the percentage of patients who continued to undergo treatment despite a negative biopsy result.

Several patients included in this study likely were not evaluated for all the investigated symptoms. Therefore, our consistent approach was to count the absence of a reported symptom as negative. However, if a patient had a particular symptom—jaw claudication, for example—but was not questioned by their clinician, that symptom would have been incorrectly categorized as not present. Additionally, we evaluated biopsies only as positive or negative. In working with our rheumatologists, we found that they typically treat GCA in settings where there is loss of the internal elastic lamina, sometimes referenced as healed arteritis. This may mean that other arterial segments could have active disease. Loss of the internal elastic lamina was not consistently reported in our pathology results. Therefore, biopsies considered negative for GCA in our study may have existed in patients who had or were still treated for presumed GCA. These other features, which were present but not documented in the formal report, were relayed during any pathologist and clinician follow-up discussions.

CONCLUSIONS

Jaw/tongue claudication, elevated inflammatory markers, and older age were associated with positive biopsy results. Therefore, we developed a risk stratification tool that might increase our positive TAB yield from 13% to 20% (nearer to the proposed 25% benchmark) by avoiding biopsies and the potential morbidity of steroid use in most low-risk patients. This change should reduce unnecessary procedures and reduce health care spending. After implementing this tool, we plan to prospectively evaluate our changes and their effect on patient care.

Funding/Support: None declared.

Financial Disclosures: None declared.

REFERENCES

1. Gonzalez-Gay MA, Vazquez-Rodriguez TR, Lopez-Diaz MJ, et al. Epidemiology of giant cell arteritis and polymyalgia rheumatica. *Arthritis Rheum.* 2009;61(10):1454-1461. doi:10.1002/art.24459

2. Cid MC, Campo E, Ercilla G, et al. Immunohistochemical analysis of lymphoid and macrophage cell subsets and their immunologic activation markers in temporal arteritis. Influence of corticosteroid treatment. *Arthritis Rheum*. 1989;32(7):884-893.

3. Ashton-Key MR, Gallagher PJ. False-negative temporal artery biopsy. *Am J Surg Pathol*. 1992;16(6):634-635. doi:10.1097/00000478-199206000-00014

4. Koster MJ, Matteson EL, Warrington KJ. Large-vessel giant cell arteritis: diagnosis, monitoring and management. *Rheumatology (Oxford).* 2018;57(suppl_2):ii32-ii42. doi:10.1093/rheumatology/kex424

5. Evans JM, O'Fallon WM, Hunder GG. Increased incidence of aortic aneurysm and dissection in giant cell (temporal) arteritis. A population-based study. *Ann Intern Med.* 1995;122(7):502-507. doi:10.7326/0003-4819-122-7-199504010-00004

6. Hunder GG, Bloch DA, Michel BA, et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis Rheum.* 1990;33(8):1122-1128. doi:10.1002/art.1780330810

7. Alberts M. Temporal arteritis: improving patient evaluation with a new protocol. *Perm J.* 2013;17(1):56-62. doi:10.7812/TPP/12-067

8. Chu R, Foster C, Ali M, et al. Optimal length and usefulness of temporal artery biopsies in the diagnosis of giant cell arteritis: a 10-year retrospective review of medical records. L*ancet Rheumatol*. 2020;2(12):e774-e778. doi:10.1016/S2665-9913(20)30222-8

9. Ing EB, Wang DN, Kirubarajan A, et al. Systematic review of the yield of temporal artery biopsy for suspected giant cell arteritis. *Neuroophthalmology*. 2018;43(1):18-25. doi:10.1080/01658107.2018.1474372

10. Ray-Chaudhuri N, Kiné DA, Tijani SO, et al. Effect of prior steroid treatment on temporal artery biopsy findings in giant cell arteritis. *Br J Ophthalmol.* 2002;86(5):530-532. doi:10.1136/bjo.86.5.530

11. Cogbill TH, Marcos BS. Bedside procedures for general surgeons: part 2. In: Ashley SW, ed. *Decker Surgery*. Decker Medicine; 2017.

12. Karahaliou M, Vaiopoulos G, Papaspyrou S, Kanakis MA, Revenas K, Sfikakis PP. 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(4):R116. doi:10.1186/ar2003

13. Chung SH, Morcos MB, Ng B. A longer or bilateral temporal artery biopsy (TAB) is more likely to yield a positive result for giant cell arteritis (GCA) in the Veterans Health Administration (VHA) Database cohort. *Arthritis Rheumatol.* 2018;70(suppl 10):3108.

advancing the art & science of medicine in the midwest

WMJ

WMJ (ISSN 1098-1861) is published through a collaboration between The Medical College of Wisconsin and The University of Wisconsin School of Medicine and Public Health. The mission of *WMJ* is to provide an opportunity to publish original research, case reports, review articles, and essays about current medical and public health issues.

Visit www.wmjonline.org to learn more.