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

Raashid Luqmani,[1]* Ellen Lee,[2] Surjeet Singh,[1] Mike Gillett,[2] Wolfgang A Schmidt,[3] Mike Bradburn,[2] Bhaskar Dasgupta,[4] Andreas P Diamantopoulos,[5] Wulf Forrester-Barker,[1] William Hamilton,[6] Shauna Masters,[1] Brendan McDonald,[7] Eugene McNally,[1] Colin Pease,[8] Jennifer Piper,[1] John Salmon,[9] Allan Wailoo,[2] Konrad Wolfe[10] and Andrew Hutchings[11]

[1]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[2]School of Health and Related Research, University of Sheffield, Sheffield, UK
[3]Immanuel Krankenhaus Berlin, Medical Centre for Rheumatology Berlin-Buch, Berlin, Germany
[4]Department of Rheumatology, Southend University Hospital NHS Foundation Trust, Southend, UK
[5]Department of Rheumatology, Hospital of Southern Norway, Kristiansand, Norway
[6]Primary Care Diagnostics, University of Exeter Medical School, Exeter, UK
[7]Department of Neuropathology and Ocular Pathology, Oxford University Hospitals NHS Trust, Oxford, UK
[8]Department of Rheumatology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[9]Oxford Eye Hospital, Oxford University Hospitals NHS Trust, Oxford, UK
[10]Department of Pathology, Southend University Hospital NHS Foundation Trust, Southend, UK
[11]Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, UK

*Corresponding author

Declared competing interests of authors: Raashid Luqmani received honoraria from GlaxoSmithKline (GSK), Nordic and Chemocentryx for training in the use of the Birmingham Vasculitis Activity Score and Vasculitis Damage Index, and personal fees from Roche outside the submitted work. Raashid Luqmani received grants from Fundação para a Ciência e Tecnologia (Portugal), Canadian Institute of Health Research, Arthritis Research UK, Patient Centered Outcomes Research Institute, Oxford University Hospitals NHS Trust Innovation Challenge Competition and Vasculitis UK. Raashid Luqmani has patents pending for a mechanical arm to automate acquisition of ultrasound images and analysis for reviewing ultrasound images. Bhaskar Dasgupta received personal fees from GSK, Servier, Roche, Merck, and Mundipharma and grants from Napp outside the submitted work. Andrew Hutchings was funded by a Medical Research Council special training fellowship in health services research during the development of the study. Jennifer Piper has a patent pending for an ultrasound arm.

Published November 2016
DOI: 10.3310/hta20900

Scientific summary

The Role of Ultrasound Compared to Biopsy of Temporal Arteries

Health Technology Assessment 2016; Vol. 20: No. 90
DOI: 10.3310/hta20900

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Giant cell arteritis (GCA) is a relatively common form of blood vessel inflammation, which usually affects people over the age of 50 years. GCA typically causes headaches and systemic upset, but can be associated with sudden and irreversible sight loss. For this reason, if a general practitioner (GP) sees a patient with suspected disease, the patient will usually be commenced on high doses of glucocorticoids, often before the diagnosis has been confirmed by further testing. Therefore, it is important to make the diagnosis correctly in order to decide on the need to continue high-dose glucocorticoid therapy to improve the condition and reduce the risk of visual loss. However, it is also important to avoid treating those without the condition, because there is a very high incidence of side effects associated with long-term glucocorticoid therapy. Temporal artery biopsy is the current gold standard test for establishing the diagnosis, with a high specificity but low sensitivity. It can be misleading in a significant number of cases. Up to 44% of patients with clinical features of GCA have a negative biopsy. There are many reasons for this, including the adequacy of the specimen obtained, the duration of glucocorticoid treatment prior to biopsy and the presence of skip lesions (intermittent, dispersed areas of abnormality in the artery that might be missed because not all areas of the artery will be sectioned for examination). Ultrasound and other imaging techniques are emerging as alternative tests to biopsy but have not been taken up widely. Ultrasound imaging can be used to assess both temporal arteries as well as both axillary arteries, which has been shown to increase the diagnostic yield.

Objectives

We aimed to test the clinical effectiveness and cost-effectiveness of ultrasound as an alternative to biopsy in the diagnosis of patients with a new presentation of possible GCA. The primary objectives of the study were (1) to evaluate the diagnostic performance (sensitivity and specificity) of ultrasound as an alternative to biopsy for diagnosing GCA in patients who are referred with suspected GCA and in whom a biopsy was going to be carried out; and (2) to perform a cost-effectiveness analysis to compare different potential investigation strategies for diagnosing GCA, incorporating either or both ultrasound and biopsy.

The secondary objectives in the study were to evaluate:

- sequential diagnostic performance and cost-effectiveness of biopsy following ultrasound in patients who have a negative ultrasound, compared with either ultrasound or biopsy alone
- the clinical effectiveness and cost-effectiveness of providing ultrasound results or biopsy results alone on treatment decisions proposed by participating clinicians
- the diagnostic performance of ultrasound in specific subgroups such as individuals at high or low risk of disease and the level of variation in ultrasound appearance in terms of halo size and degree of stenosis.

Methods

We conducted a multicentre, prospective study of new cases of suspected GCA. In order to ensure standardisation of the new technique of ultrasound scanning of temporal arteries, we needed to develop a training programme for sonographers to standardise the performance and interpretation of ultrasound assessment of temporal and axillary arteries. The training consisted of a Microsoft PowerPoint® version 97–2003 (Microsoft Corporation, Redmond, WA, USA) presentation, an online assessment in which sonographers were required to correctly identify video images of scans as showing or not showing features

consistent with GCA (75% pass mark required), the provision of video and still images from 10 control individuals and one patient with active GCA and evidence of ultrasound abnormalities consistent with the diagnosis. All images were reviewed by an expert panel. If scan techniques were suspected to be inadequate, the sonographers underwent retraining. We compared the standard of care in the investigation of GCA [clinical evaluation, measurement of acute phase response and temporal artery biopsy (TAB) from the most affected artery] with ultrasound. All patients underwent both tests in sequence (ultrasound first, followed by biopsy) within 7 days of commencing high-dose steroids for the suspected diagnosis of new-onset GCA. We did not provide any training in TAB or in the interpretation of the results by pathologists, because these are established techniques in NHS care. All patients received normal care as decided by their clinician. The clinician reviewed the patient at baseline, 2 weeks later (after the scan and biopsy had been performed) and 6 months later. We used the clinician diagnosis made at 2 weeks as the primary outcome measure for the study. We also established a reference standard diagnosis based on the clinician's submitted diagnosis, any revisions by 6 months, or any revisions by the expert panel who reviewed all case records. The scan result was kept blinded from the clinician treating the patient to avoid the results influencing the decision on diagnosis. The clinician was supplied with all other results including the biopsy because this would reflect standard of care. After the clinician submitted their diagnosis at 2 weeks, they were provided with the scan results on request if they were considering withdrawal of therapy if they concluded that the patient did not have GCA. We excluded patients with a previous diagnosis of GCA or a previous TAB; we also excluded patients who had been treated with high-dose steroids for any other reason for more than 7 days prior to the ultrasound and biopsy. We created 30 case vignettes from the data obtained from patients in the study who either did or did not have a reference standard diagnosis of GCA and in whom the biopsy or scan could have been either positive or negative. We undertook two interobserver rater exercises to assess variation in diagnosis based on interpretation of ultrasound and biopsy images.

Results

We recruited 35 centres for the study and provided ultrasound training for 49 sonographers. Only 26 sonographers from 22 sites completed the training. Two of these sites did not take any further part in the study, leaving 24 sonographers at 20 recruiting sites. Seven sonographers passed all three components of the training at first attempt; 13 required further attempts to pass; and four were exempted from some parts of the training because they had already demonstrated expertise.

We enrolled 430 patients with suspected GCA into the study; 44 withdrew or did not have their scan and biopsy within 10 days of starting glucocorticoid therapy and five withdrew before completing a follow-up assessment. The remaining 381 patients were included in the primary analysis. The median age [interquartile range (IQR)] was 71 years (64–78 years); 72% were female. The median time between first symptom onset and baseline was 31 days (IQR 10–93 days, $n = 377$); the median time between symptom onset and starting steroids was 33 days (IQR 13–99 days, $n = 379$). The reference diagnosis was based on the 2-week and 6-month clinical diagnosis, as well as the opinion of an expert review panel, which assessed all of the patient data apart from the ultrasound results. In total, 257 out of 381 patients were considered to have GCA on the basis of the reference standard diagnosis. Twenty-one patients had a change of diagnosis from the original clinician's final available diagnosis following an expert review: in eight patients, the diagnosis was changed from GCA to not GCA; in a further 13 patients, the diagnosis was switched from not GCA to GCA. We compared the relative performance of ultrasound and biopsy in 257 patients with a reference standard diagnosis of GCA and 124 patients diagnosed as not having GCA. The incidence of polymyalgia rheumatica was 28 cases at baseline, and ischaemic optic neuropathy was reported in 9.7% of cases with GCA at baseline (compared with 4.8% in the non-GCA group), 6.2% of cases at 2 weeks (compared with 0.8%) and 4% of cases at 6 months (compared with 2.8%). Baseline comorbidity included hypertension in 52.5% of patients, diabetes mellitus in 14.2% of patients, ischaemic heart disease in 7.3% of patients, heart failure in 5% of patients, malignancy in 2.4% of patients and fracture in 0.3% of patients. At 6 months after diagnosis, the incidence of hypertension had increased to

55.8% and the incidence of diabetes mellitus had increased to 18.2%. Four low-trauma fractures occurred during the follow-up period.

In total, 101 cases had a biopsy consistent with the diagnosis and 162 cases had an ultrasound consistent with the diagnosis. In 70% of patients, the results of biopsy and ultrasound were concordant (74 positive and 192 negative), giving a kappa statistic of 0.35. In 27 patients the results were positive for biopsy but negative for ultrasound and in 88 patients the results were negative for biopsy but positive for ultrasound. The sensitivity of biopsy versus ultrasound was 39% [95% confidence interval (CI) 33% to 46%] compared with 54% (95% CI 48% to 60%) and the specificity was 100% (95% CI 97% to 100%) compared with 81% (95% CI 73% to 88%). We analysed change in the degree of ultrasound abnormality (based on halo size) during the 7-day assessment period. We found that the halo was likely to be much smaller after at least 4 days of high-dose glucocorticoid therapy than after fewer than 4 days of steroid therapy. The biopsy positivity rate also diminished significantly within 3 days of starting high doses of glucocorticoids. We also evaluated alternative strategies for combining the two tests and incorporating clinical judgement (the doctor's assessment of GCA based on the patient's characteristics and available test results). The most cost-effective strategy was to perform an ultrasound examination of all patients with suspected GCA. This strategy was more sensitive (93% vs. 91%), less specific (77% vs. 81%) and more cost-effective (incremental net monetary benefit of £485 per patient) than current standard practice, that is, a strategy involving biopsy and clinical judgement alone. The cost-effectiveness analysis accounted for the cost of the testing and the consequences of correct or incorrect diagnosis resulting in drug toxicity (e.g. fracture), as well as irreversible, potentially preventable, sight loss from anterior ischaemic optic neuropathy.

We measured inter-rater agreement for the two tests in a series of 30 cases selected from among the cohort. Among sonographers the intraclass correlation coefficient for agreement was 0.61 (95% CI 0.48 to 0.75) and for biopsy the intraclass correlation coefficient was 0.62 (95% CI 0.49 to 0.76). Agreement was strongest where the pathological findings included the presence of giant cells, but was much weaker for cases in which only minimal change was found. This suggests that the current approach of classifying test results as either positive or negative may be too simplistic because the tests are not always enough by themselves to make the diagnosis, as there can be some degree of uncertainty in interpreting the results.

Conclusions

We conclude that ultrasound, in comparison with TAB, is a more sensitive and cost-effective investigation in suspected cases of GCA. However, over one-third of the patients eventually diagnosed with GCA had neither a positive scan nor a positive biopsy, highlighting the importance of assessing the patient for clinical indicators to support the diagnosis rather than relying on test results alone. Temporal artery biopsy has a much lower sensitivity in the current study (39%) than in previously published figures (> 70%). It is tempting to speculate that this may reflect better awareness of the diagnosis of GCA and a willingness of GPs to commence treatment early in the disease course. This is supported by the relatively short time from symptom onset to diagnosis in this cohort (31 days) compared with other cohorts. The findings have potential implications for improving the management of GCA through the more effective use of available techniques to provide a clinically effective and cost-effective strategy for the diagnosis of GCA. Further research should address the issue of an independent reference diagnosis, standards for interpreting and reporting test results and the evaluation of ultrasound training and should explore the acceptability of these new diagnostic strategies in GCA. Some clinicians and patients may be uncomfortable with a strategy that does not involve a biopsy and may prefer to perform biopsies in all cases that are ultrasound negative, or in all cases of patients at medium or high risk of GCA in terms of clinical features but who have a negative ultrasound scan, to provide further evidence to rule in the disease as well as to support withdrawing therapy if both tests are negative. Although these combined strategies would be more expensive than our proposal (because a proportion of patients would require a biopsy), they remain more cost-effective than current practice (performing a biopsy in all suspected cases) and may be more acceptable to patients and clinicians.

Future work

Further research should address the issue of an independent reference diagnosis, standards for interpreting and reporting test results and the evaluation of ultrasound training and should explore the acceptability of these new diagnostic strategies in GCA.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 08/64/01. The contractual start date was in January 2010. The draft report began editorial review in June 2015 and was accepted for publication in July 2016. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).