Received: 6 July 2020 | Revised: 24 August 2020 | Accepted: 20 September 2020

DOI: 10.1111/coa.13652

ORIGINAL ARTICLE

WILEY

A prospective multicentre external validation study of the Liverpool Peritonsillar abscess Score (LPS) with a no-examination COVID-19 modification

David Selwyn[1] | Ding Yang[2] | Elliot Heward[3] | Ashwin Kerai[4] | Elinor Thompson[4] | Abulgasem Shommakhi[4] | Scott Faulkner[4] | Richard Siau[4] | Hussein Walijee[4] | Tom Hampton[5] | Dorota Chudek[6] | Supriya Singhera[7] | Waqas Din[1] | Andrew S. Lau[4]

[1]Bradford Royal Infirmary, Bradford, UK

[2]Royal Stoke University Hospital, Stoke-on-Trent, UK

[3]Stepping Hill Hospital, Stockport, UK

[4]Arrowe Park Hospital, Wirral, UK

[5]Mersey ENT Research Collaborative (MERC), Merseyside, UK

[6]Royal Liverpool University Hospital, Liverpool, UK

[7]Wythenshawe Hospital, Manchester, UK

Correspondence
David Selwyn, Department of Otolaryngology, Hull Royal Infirmary, HU3 2JZ, UK.
Email: d.selwyn@nhs.net

Abstract

Objectives: Our primary aim was to validate the Liverpool Peritonsillar abscess Score (LPS) externally in a new patient cohort. Our secondary aim was to modify the LPS in the light of the COVID-19 pandemic to produce a no-examination variant for use in this instance.

Design: Prospective multicentre external validation study.

Setting: Six different secondary care institutions across the United Kingdom.

Participants: Patients over 16 years old who were referred to ENT with any uncomplicated sore throat such a tonsillitis or peritonsillar abscess (PTA).

Main outcome measures: Sensitivity, specificity, positive predictive value and negative predictive value for both the original LPS model and the modified model for COVID-19.

Results: The LPS model had sensitivity and specificity calculated at 98% and 79%, respectively. The LPS has a high negative predictive value (NPV) of 99%. The positive predictive value (PPV) was slightly lower at 63%. Receiver operating characteristic (ROC) curve, including the area under the curve (AUROC), was 0.888 which indicates very good accuracy.

Conclusions: External validation of the LPS against an independent geographically diverse population yields high NPV. This may support non-specialist colleagues who may have concerns about mis-diagnosing a PTA. The COVID-19 modification of the LPS has a similar NPV, which may be of use where routine oral examination is to be avoided during the COVID-19 pandemic.

1 | INTRODUCTION

1.1 | Background

Peritonsillar abscesses (PTA, also known as quinsies) are a frequently occurring presentation to emergency departments and represent a common reason for referral to an ENT service. Management options include drainage of the abscess and antibiotic therapy. Without intervention, PTA can develop into deep neck space infections (DNSI) which carry significant morbidity and mortality.[1] Patients who undergo drainage of their PTA have been shown to experience quicker resolution of symptoms.[2]

Oral presentation, British Otolaryngology and Allied Research Sciences (BOARS) Meeting, 15th November 2019

Reports suggest that there are increasing numbers of inpatient admissions to secondary care with PTA and DNSI.[3] It is therefore important for both non-specialists and specialists to arrive at the correct diagnosis in order to facilitate early and appropriate treatment. We previously reported the development of a predictive score for peritonsillar abscess through a prospective multicentre observational study (Liverpool Peritonsillar abscess Score, LPS).[4]

The LPS is an additive threshold score consisting of five variables with the aim of predicting the likelihood of PTA in adult patients. Internal validation of the LPS from the initial study sample produced a high sensitivity and negative predictive value (NPV; 96% for both). The LPS functions very much like the Wells score (which estimates the probability of venous thromboembolism), and its strength lies in its NPV.[5] We theorise that the PPV is lower because a proportion of patients with tonsillitis will have trismus, and because the LPS was developed using aspirate-proven PTA (rather than peritonsillar cellulitis).

Our aim in developing the LPS was to support non-specialists in preventing a missed diagnosis of PTA and to facilitate remote triage by secondary care services and to avoid any need for significant extra resources or funding in the process.

1.2 | Covid-19 modification

During the current COVID-19 pandemic, ENT UK has issued updated guidance, reserving oral examination only for severe cases of sore throat.[6] A validated predictive score not reliant on oral examination may therefore be useful to the non-specialist when triaging or referring a suspected PTA.

1.3 | Objectives

Our primary aim was to validate the Liverpool Peritonsillar abscess Score (LPS) externally against a separate cohort of patients through a multicentre prospective observational study. Our secondary aim was to modify the LPS in the light of the COVID-19 pandemic to produce a no-examination variant for use in this instance.

2 | MATERIALS AND METHODS

2.1 | Reporting standards

This study and manuscript adhere to the STROBE reporting guidelines for observational studies.[7]

2.2 | Ethical considerations and institutional review

This project was registered with the clinical governance departments at six secondary care sites: Bradford Royal Infirmary (coordinating), Royal Liverpool University Hospital, Arrowe Park Hospital, Stepping Hill Hospital, Wythenshawe Hospital and Royal Stoke Hospital, all in the United Kingdom. After collection at each site, truly pseudo-anonymised data (unique study number) were shared with the coordinating site while complying with information governance and data protection procedures.

Key Points

- There are increasing numbers of inpatient admissions with peritonsillar abscesses (PTA) and deep neck space infections
- The Liverpool Peritonsillar abscess Score (LPS) is a predictive quantitative score of five variables to be used as an adjunct in assessing patients with sore throats
- The LPS model has a high sensitivity and negative predictive value (98% and 99%)
- The LPS may be used as a tool to assist non-specialist colleagues in recognising PTA
- A modified LPS which removes the need for oral examination still has high sensitivity (94%) and a negative predictive value (98%) which may be of particular use during the COVID-19 pandemic

2.3 | Sample size estimation

When developing the LPS, we based our previous sample size estimate on the need for a certain number of events per variable (EPV). However, in the external validation of a multivariable predictive model, statistical modelling suggested the need for "at least 100 events and ideally 200 (or more) events".[8] We therefore set our sample size at 200.

2.4 | Definitions

In line with our previous work, we defined a PTA as clinical suspicion based on typical symptoms and signs with positive aspiration of pus from the peritonsillar space. Non-PTA sore throat, including tonsillitis and peritonsillar cellulitis, was defined either on senior review (ENT specialty registrar ST3 or more senior) or when there had been suspicion of PTA but a negative aspirate.

2.5 | Inclusion and exclusion criteria

Data were collected prospectively from August 2018 onwards until we met our sample size requirement. We included consecutive patients over 16 years old who were referred to the ENT service for assessment of sore throat such as tonsillitis or PTA. Patients were assessed in the local emergency department or surgical assessment unit according to standard local protocol. We did not apply any length of stay criteria.

The LPS is validated for assessing the risk of peritonsillar abscess in patients ≥16 years. It does not replace clinical judgement and appropriate PPE should be used.

Unilateral sore throat (rated 80:20 or more by patient)	3
Trismus (inter-incisor distance <3cm)	2
Male gender	1
Pharyngeal voice chance (hot potato voice)	1
Uvular deviation (away from affected side)	1

Covid-19 Modification: clinicians may remove the uvular deviation component to avoid examining the mouth; there is no change to the cut-off value (see Results for classification function).

FIGURE 1 The Liverpool Peritonsillar abscess Score (LPS). This is an additive score with a cut-off value of 4. Scores of 0-3 predict no PTA whereas the likelihood of PTA increases with scores of 4-8. The LPS is validated for assessing the risk of peritonsillar abscess in patients ≥16 years. It does not replace clinical judgement and appropriate PPE should be used. COVID-19 modification: Clinicians may remove the uvular deviation component to avoid examining the mouth; there is no change to the cut-off value (see Results for classification function)

TABLE 1 Demographics of patients with and without PTA

	PTA n = 54	Non-PTA n = 151	*P*-value
Median age (years)	27 (range from 17 to 63)	32 (range from 16 to 85)	.029
Male gender (percentage)	65	37	<.001

We excluded any patients under 16 years old. We also excluded any patients referred with, or who were subsequently proved to have, supra- or epiglottitis or a DNSI.

2.6 | Data sources

Data quality and uniformity were ensured through the use of a single data collection proforma across all sites. Investigators were briefed on the use of the proforma. Data were collated locally before being pooled using unique study numbers. Investigators scored patients both at the point of referral from either General Practitioners or non-specialists in the local emergency department, and on assessment by an ENT doctor. The final diagnosis was confirmed both by the presence of pus on aspiration and senior review.

The LPS COVID-19 modification was developed to assist in risk stratification of PTA during the COVID-19 pandemic. Since uvular deviation was the only variable requiring oral examination with a tongue depressor, we removed it from the parent LPS and examined the classification function of the modified score.

2.7 | Statistical methods

Data from each collection site were collated prospectively and entered into Excel workbooks (Microsoft Office 365, Microsoft, Redmond, WA, USA). We analysed sample demographics and reported summary measures. We also calculated the classification function and area under the receiver operating characteristics curve (AUROC). Statistical associations were determined using the Mann-Whitney *U* test for skewed continuous data, and the significance level (α) was set at 0.05.

2.8 | The Liverpool Peritonsillar abscess Score (LPS)

In this study, we used the previously reported final iteration of the LPS (Figure 1).

3 | RESULTS

3.1 | Demographics and frequency of PTA

We included 205 patients referred with sore throat (Table 1). Of these, 54 patients were diagnosed with PTA based on the presence of aspirated pus. There was no significant difference in the median age of patient with or without a PTA ($P = .029$).Overall, 44% of patients were men and were significantly more likely to have PTA ($P < .001$).

3.2 | Accuracy of external validation

In this external validation study, the LPS continues to have a high negative predictive value (NPV; 99%) and sensitivity (98%). Its specificity and positive predictive value (PPV) are lower at 79% and 63%, respectively.

A receiver operating characteristics curve, including the area under the curve (AUROC), was calculated for the LPS. This showed that model accuracy was very high (0.888; Figure 2).

The LPS COVID-19 modification (removal of the uvular deviation component) appeared to have similar classification function based

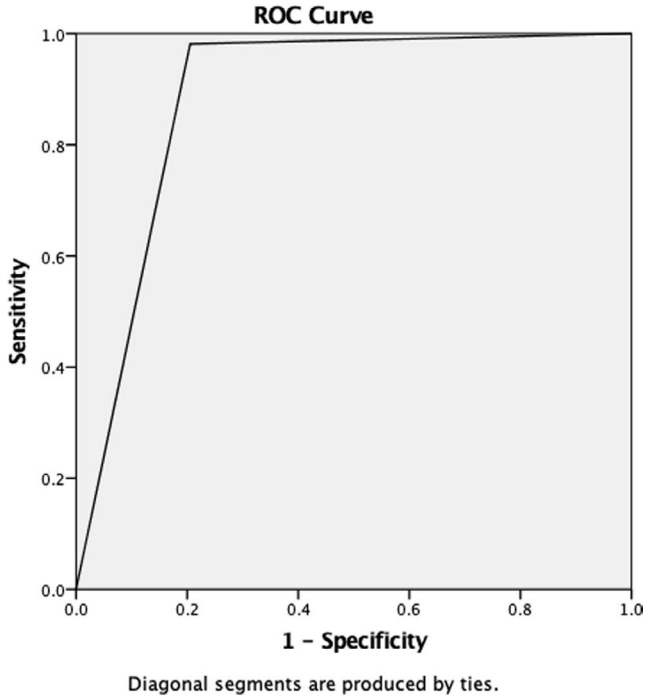

FIGURE 2 Receiver operating characteristics (ROC) curve for the external validation sample. The area under the curve (AUROC) is 0.888, indicating very good accuracy

TABLE 2 Classification function of the LPS COVID-19 modification (where uvular deviation has been removed). The cut-off value remains 4

		Observed PTA		
		Positive	Negative	
Predicted PTA	Positive	53	31	PPV 62%
	Negative	1	120	NPV 98%
		Sens 94%	Spec 79%	

on these data. Sensitivity and NPV remained high at 94% and 98%, respectively (Table 2).

4 | DISCUSSION

4.1 | Synopsis of key results

When tested against an independent geographically diverse population, the LPS demonstrates high NPV and sensitivity, as well as overall accuracy (AUROC). We hypothesise that a lower specificity and PPV are because some patients with tonsillitis will have trismus, and because we defined PTA using definite aspiration of pus in order to reduce subjectivity (Table 3). Since statistical analysis of the LPS has yielded consistent findings across both internal and external validation, further iterative development of the LPS is not indicated.

We have developed a COVID-19 modification to the LPS, which may be of use where clinicians are trying to reserve oral examination for severe cases during the COVID-19 pandemic. This modified score also demonstrates high NPV and sensitivity.

TABLE 3 Classification function of the LPS following external validation

		Observed PTA		
		Positive	Negative	
Predicted PTA	Positive	53	31	PPV 63%
	Negative	1	120	NPV 99%
		Sens 98%	Spec 79%	

Abbreviations: Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value.

4.2 | Limitations and strengths

This observational study benefits from collection of prospective data over a broad geographical area with a pre-determined sample size. Cohort characteristics are comparable to many other studies of patients with sore throats, and age and gender distributions are similar to the original observational study used to develop the LPS.[4,9] The classification function and AUROC yield high degrees of negative prediction and accuracy, respectively. In developing the LPS, we have tried to use objective measures to ensure that the resulting score is highly reproducible in the hands of many different clinicians, including non-specialists.

In this study, we have categorised diagnosis into discrete categories of "PTA" or "non-PTA" sore throat to ensure reproducibility and to reduce subjectivity: this is in line with our previous methodology.[4] This is likely to mean that the LPS does not differentiate between peritonsillar cellulitis and tonsillitis (in fact, it is likely that this is also reflected in the relatively lower specificity and PPV of the LPS). The stated aim of the LPS, however, is to avoid a missed diagnosis of PTA rather than to differentiate between tonsillitis and peritonsillar inflammatory disease. In the UK, most PTA is treated with a combination of drainage and antibiotics. By contrast, both peritonsillar cellulitis and tonsillitis can be treated medically with antibiotics and supportive advice.

It is also possible that some of the patients diagnosed as having non-PTA sore throat had PTAs which did not yield an aspirate, either due to small size or due to inexperience on the part of the admitting clinician. However, no patient, once formally diagnosed with non-PTA sore throat, required a repeat trial of aspiration. These factors may therefore affect relatively small number of patients and are unlikely to affect the reliability of the score.

4.3 | Comparison with other studies

Almost in parallel with our original study, Spiekermann*et al* reported a PTA predictive score which included six variables.[9] Score parameters included presence of halitosis, uvular oedema, unilateral swelling of the palatal arches and trismus, as well as presence

of an inflammatory biomarker in both serum and saliva (S100 A8/A9 complex). The study was retrospective in nature and included a smaller sample of 75 patients (24 acute tonsillitis; 51 peritonsillar inflammation). The authors aimed to determine whether they could predict if a patient with peritonsillar inflammation would benefit from medical treatment or "abscess relief" with the primary aim being able to differentiate between peritonsillar abscess from cellulitis. Demographics and key definitions were similar between that study and our own. The medical context, however, appeared to be different in that a significant number of patients underwent abscess tonsillectomy, which is now rarely performed in the UK. In addition, the use of S100 A8/A9 complex as a biomarker for peritonsillar inflammation is not in widespread use at present, is limited to the research sphere and requires additional laboratory resources and funding.

4.4 | Generalisability

External validation now provides a measure of confidence in the reliability and utility of the LPS. The LPS does not seek to alter the threshold for admitting patients with sore throat (when they cannot swallow enough to remain hydrated and to take medication). Instead, its clinical applicability lies in the referral interface between primary or emergency care and ENT, and in stratification of a patient's PTA risk.

The primary method of transmission of COVID-19 is thought to be via droplet spread: clinicians may be at greater risk when examining the oropharynx.[10] Updated ENT UK guidance attempts to limit clinician exposure during patient assessment, and the COVID-19 modification of the LPS could be used to complement this guidance.[6]

5 | CONCLUSION

In this study, we have validated the LPS externally in a new patient cohort, yielding findings consistent with our original work in which this score was developed. We have demonstrated that LPS is a robust and easy-to-use tool for colleagues seeking to stratify the risk of PTA in their patients.

CONFLICTS OF INTEREST

None.

DATA AVAILABILITY STATEMENT

The data supporting this study are available from the corresponding author upon reasonable request.

ORCID

David Selwyn https://orcid.org/0000-0003-0028-3859
Elliot Heward https://orcid.org/0000-0002-9692-646X
Richard Siau https://orcid.org/0000-0002-6950-4026
Hussein Walijee https://orcid.org/0000-0002-6838-4499
Tom Hampton https://orcid.org/0000-0002-6515-8058
Dorota Chudek https://orcid.org/0000-0001-5174-8359
Waqas Din https://orcid.org/0000-0003-1416-8181
Andrew S. Lau https://orcid.org/0000-0002-6489-204X

REFERENCES

1. Douglas CM, Altmyer U, Cottom L, Young D, Redding P, Clark LJ. A 20-year observational cohort of a 5 million patient population – tonsillectomy rates in the context of two national policy changes. *Clin Otolaryngol*. 2018;44:7-13.
2. Tachibana T, Orita Y, Abe-Fujisawa I, et al. Prognostic factors and effect of early surgical drainage in patients with peritonsillar abscess. *J Infect Chemother*. 2014;20:722-725.
3. Lau AS, Upile NS, Wilkie MD, Leong SC, Swift AC. The rising rate of admission for tonsillitis and neck space abscesses in England, 1991–2011. *Annals R Coll Surg Engl*. 2014;96(4):307-310.
4. Lau AS, Selwyn DM, Yang D, et al. The Liverpool Peritonsillar abscess Score: Development of a predictive score through a prospective multicentre observational study. *Clini Otolaryngol*. 2019;44(3):293-298.
5. Wells PS, Anderson DR, Rodger M, et al. Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis. *N Engl J Med*. 2003;349(13):1227-1235.
6. ENT UK. (2020). ENT UK COVID-19 Adult Tonsillitis & Quinsy Guidelines. Updated March 2020. Available at: entuk.org/ covid -19 [Accessed 6 April 2020]
7. Von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370(9596):1453-1457.
8. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med*. 2016;35:214-226.
9. Spiekermann C, Roth J, Vogl T, Stenner M, Rudack C. Potential of the novel PTA score to identify patients with peritonsillar inflammation profiting from medical treatment. *Dis Mark*. 2018;2018:2040746.
10. Vukkadala N, Qian ZJ, Holsinger FC, Patel ZM, Rosenthal E. COVID-19 and the otolaryngologist – preliminary evidence-based review. *Laryngoscope*. Published online 26 March, 2020. https://doi.org/10.1002/lary.28672. [Epub ahead of print]

How to cite this article: Selwyn D, Yang D, Heward E, et al. A prospective multicentre external validation study of the Liverpool Peritonsillar abscess Score (LPS) with a no-examination COVID-19 modification. *Clin Otolaryngol*. 2021;46:229–233. https://doi.org/10.1111/coa.13652