BJGP OPEN

Diagnostic accuracy of FEVER-Pain and Centor criteria for bacterial throat infection in adults with sore throat: a secondary analysis of a randomised controlled trial.

Seeley, Anna; Fanshawe, Thomas; Voysey, Merryn; Hay, Alastair; Moore, Michael; Hayward, Gail

When citing this article please include the DOI provided above.

Title: Diagnostic accuracy of FEVER-Pain and Centor criteria for bacterial throat infection in adults with sore throat: a secondary analysis of a randomised controlled trial.

Authors

Dr Anna Seeley, BA MBChB[1*] - Academic Clinical Fellow

Dr Thomas Fanshawe, PhD[1] – Senior Medical Statistician

Dr Merryn Voysey, PhD[2] - Lead Statistician

Professor Alastair D Hay, PhD[3] – Professor of Primary Care

Professor Michael Moore, MSc[4] – Professor of Primary Care

Associate Professor Gail Hayward, PhD[1] – Professor of Primary Care

Affiliations

1. Nuffield Department of Primary Care Health Sciences, University of Oxford, UK
2. Oxford Vaccine Group, Department of Paediatrics, University of Oxford, UK
3. Centre of Academic Primary Care, Bristol Medical School, University of Bristol, Bristol, UK
4. Primary Care and Population Sciences Unit, University of Southampton, Aldermoor Health Centre, Aldermoor Close, Southampton, United Kingdom

Correspondence to

Anna Seeley, Nuffield Department Primary Care Health Sciences, Radcliffe Square, Woodstock Road, OX2 6HG; anna.seeley@worc.ox.ac.uk

Main text word count: 2498 **Abstract Word Count:** 250 **References:** 20 **Tables:** 2 **Figures:** 4

How this study fits in

Sore throat is one of the most common presentations in primary care, and despite the self-limiting nature, antibiotics are over-prescribed more than any other condition. There remains ambiguity as to the performance of FeverPAIN and Centor criteria in ruling out streptococcal infection. This diagnostic accuracy study from 565 participants in UK primary care, shows neither clinical prediction rule perform well in a low prevalence setting. Clinicians can be reassured that either no or a delayed antibiotic prescription is appropriate for the majority of patients.

ABSTRACT

Background: Sore throat is a common and self-limiting condition. There remains ambiguity in stratifying patients to immediate, delayed or no antibiotic prescriptions. NICE recommends two clinical prediction rules (CPRs), FeverPAIN and CENTOR, to guide decision-making.

Aim: Describe the diagnostic accuracy of CPRs in identifying streptococcal throat infections.

Design and Setting: Adults presenting to UK primary care with sore throat, who did not require immediate antibiotics.

Methods: As part of the Treatment Options without Antibiotics for Sore Throat (TOAST) trial, 565 participants, aged ≥ 18, were recruited on day of presentation to general practice. Physicians could opt to give delayed prescriptions. CPR scores were not part of the trial protocol but calculated post-hoc from baseline assessments. Diagnostic accuracy was calculated by comparing scores to throat swab cultures.

Results: 81/502 (16.1%) of patients had Group A, C or G streptococcus cultured on throat swab. Overall diagnostic accuracy of both CPRs was poor: area under receiver-operating-characteristics curve 0.62 for Centor; 0.59 for FeverPAIN. Post-test probability of a positive or negative test was 27.3% (95% confidence intervals: 6.0-61.0%) and 84.1% (80.6-87.2%) for FeverPAIN ≥4, versus 25.7% (16.2-37.2%) and 85.5% (81.8-88.7%) for Centor ≥3. Higher CPR scores were associated with increased delayed antibiotic prescriptions (χ^2 = 8.42, p=0.004 for FeverPAIN ≥4; χ^2 = 32.0, p<0.001 for Centor ≥3).

Conclusions: In those that do not require immediate antibiotics in primary care, neither CPR provide a reliable way of diagnosing streptococcal throat infection. However, clinicians were more likely to give delayed prescriptions to those with higher scores.

2

INTRODUCTION

Sore throats are a common presentation in UK primary care, accounting for around 3.5 million appointments per year (1). The majority of infections are viral, and risk of progression to serious complications is low, approximately 1% of patients. (2). Despite this, antibiotics are frequently prescribed; roughly 60-70% of consultations in 2010-2011 (1) and more frequently associated with inappropriate prescriptions than any other condition (3).

The National Institute of Clinical Excellence (NICE) guidelines on sore throat infection recommends use of two clinical prediction rules (CPR) to inform antibiotic prescribing strategies (4). The Centor score was developed in 1980s, in adults presenting to emergency departments, and allocates one point to presence of cervical lymph nodes, fever, tonsillar exudates, or absence of cough (5). NICE recommends immediate or delayed antibiotics for a score of 3 or greater, associated with a 32-56% chance of streptococcal infection. The Centor criteria have been criticised for leading to over-prescription of antibiotics in primary care settings (6) and a recent meta-analysis has highlighted it's limitations, particularly to 'rule in' infection (7). It was also developed to detect Group A streptococcus infection only. FeverPAIN was derived from UK primary care populations and is a 5 point scale with a point for fever, absence of cough and purulent tonsils, as per Centor, but also for severe tonsillar exudate and duration of symptoms for less than 3 days (8). It incorporates three clinical decisions: no antibiotics (0-1); a delayed prescription (2-3); immediate antibiotics (4-5), with scores in the latter category associated with a 60-65% chance of Group A, C or G streptococcal infection. Despite endorsement from NICE (4) it has only been tested in one RCT, outside the original study population (9). Thus, there remains uncertainty about both the performance and validity of each CPR in clinical practice.

The TOAST (Treatment Options without Antibiotics for Sore Throat) was a double-blinded randomised controlled trial, in which adults presenting to their primary care clinician with acute sore throat, were randomised to receive a single dose of dexamethasone or placebo. Those who needed immediate antibiotics were not recruited to the trial, but as part of the pragmatic nature, clinicians could opt for a

3

delayed prescription. Throat swabs were taken on initial assessment and symptom diary cards were recorded for the first 7 days. Results of the study have previously been reported (10), with steroids making no difference to symptom burden in the first 24 hours. The detailed data collection of sore throat symptomatology, duration and aetiology makes this a suitable cohort for examining the validity of each CPR in a population deemed not to require immediate antibiotics, with a focus on the delayed prescription strategy, given recent evidence that delayed prescriptions may reduce antibiotic prescribing rates without compromising on clinical safety (11).

The aims of this paper were to retrospectively analyse TOAST data to understand:

1. The diagnostic accuracy of CENTOR and FeverPAIN, compared with microbiological culture of throat swabs, in identifying streptococcal throat infections in a low-risk population.

2. The relationship between CPR and prescription decision in a cohort where this was not explicitly recorded.

METHODS

Participant recruitment and baseline assessment for TOAST

Participants aged 18 or older were recruited on the day of presentation to their GP practice with acute symptoms of sore throat (onset within the last 7 days) and odynophagia judged by the clinician to be infective in origin. Exclusion criteria included: recent (<1 month) use of inhaled or oral corticosteroids or adenotonsillectomy, recent use (<14 days) of antibiotics, or a clear alternative diagnosis. Prior to randomization, the clinician was free to offer either no antibiotics or a delayed antibiotic prescription; typically to take after 48 hours if symptoms had not improved. Participants were randomized and treated immediately with either a single dose of 10mg oral dexamethasone or matching placebo.

4

Data required for calculation of FeverPAIN and Centor scores was obtained at baseline, using standardised questionnaires, by trained clinicians, but scores were not documented. A throat swab was taken for microbiological culture.

Participants were asked to complete a symptom diary for 7 days post-randomisation. Sore throat and pain on swallowing were recorded daily on a validated (12) 7-point Likert scale. Use of over-the-counter medications and antibiotic consumption were also recorded.

<u>Identification of Streptococcal throat infections</u>

Throat swabs were collected at baseline and sent to our central laboratory for culture and sensitivity. A positive swab was reported if any Lancefield β-haemolytic Group A, C, and G streptococcus was isolated, given the overlapping clinical syndromes and sequelae (13). We performed additional analysis focused on solely Group A Streptococcal (GAS) infections, responsible for the majority of infections, and the only strain which is identified on most commercially available rapid antigen testing kits.

<u>Clinical Prediction Rules (CPRs)</u>

NICE guidelines use both CPRs to employ thresholds, below which no antibiotics are recommended and above which either a delayed or immediate antibiotic prescription should be considered (4). We chose, as our main comparison, a high probability of streptococcal throat infection for both CPRs i.e. FeverPAIN ≥ 4 or Centor ≥ 3, where delayed or immediate antibiotics may be given. We used a freely available online calculator to produce an infographic of how diagnostic accuracy results would translate in to test results for a hundred patients presenting with sore throat (14). We compared these thresholds to the reference standard of microbiological culture of bacterial throat swab. Additionally, we looked at each different score on the CPRs to see how the likelihood of streptococcal infection varied across this range. Finally, we examined the relationship between delayed prescriptions offered, and the CPR scores. We did not know how clinicians were using CPR scores.

Statistical Analysis

All data analysis was done in STATA (STATA Corp version 14). FeverPAIN and Centor scores were calculated based on data provided in patient baseline assessment. Scores at each threshold of the CPR were compared to bacterial culture of throat swab using sensitivity, specificity, and positive and negative predictive values, with 95% confidence intervals (95% CI). The overall performance of each CPR was also summarised using the area under receiver-operating-characteristic (ROC) curve, with a 95% confidence interval. The Pearson's chi-squared (χ^2) test was used to compare CPR scores with delayed prescription rates, a p value less than 0.05 was considered statistically significant.

RESULTS

Trial population

The characteristics of the study population have been described previously (10). Five hundred and sixty-five participants were randomized, with a median age of 34 years. 425 (75.2%) were female and 434 (76.8%) were employed or in education. 223 participants (39.5%) received a delayed antibiotic prescription. Most participants had low or medium FeverPAIN or Centor scores; median score was 1 for both CPRs in the cohort.

Throat swab culture results were available for 502 (89%) participants. 81 (16.1%) were positive for Streptococcal group A, C or G infections, and 63 (12.5%) were positive for Group A only (Table 1).

Association between CPR scores and presence of streptococci on throat swab

Streptococci were detected on throat swabs across the range of scores of both CPRs. Prevalence was low, and positive throat swabs were less common for very low scores (0,1) (Figure 1). The overall performances of each CPR were similar and modest; Summary area under ROC curve was 0.62 (95% CI 0.56 – 0.69) for Centor, 0.59 (95% CI 0.52-0.65) for FeverPAIN (Figure 2). At the aforementioned thresholds for considering antibiotics, sensitivity of both CPRs was low (Table 2). At a FeverPAIN

6

score of ≥ 4 the specificity was 98.1%, higher than a Centor score of ≥ 3 (86.9%). However, the sensitivity of a FeverPAIN score was very low (3.7%). At a prevalence of 16.1% the post-test probability of a positive test was 25.7% (95% CI 16.2-37.2%) for Centor ≥ 3 and 27.3% (95% CI 6.0-61.0-%) for FeverPAIN ≥ 4. Figure 3 displays how this may translate into clinical practice. 3 in 100 people may have received a delayed prescription for antibiotics using FeverPAIN score of ≥ 4, with 2 of these prescriptions unnecessary. In contrast 15 people may have received a delayed antibiotic prescription, with 11 unnecessary for Centor ≥3.

The diagnostic accuracy did not improve for either CPR, at different thresholds (Table 2). The NPV estimates remained high across all scores, reflective of the prevalence of streptococcal infections in our study population. Using a FeverPAIN score of ≥ 2 or "intermediate probability" would have improved the sensitivity to 63% (51.5-73.4) but decreased the specificity to 52.5% (47.6-57.4). In clinical terms this translates into prescribing 50 in 100 people delayed antibiotics, only 10 of whom have streptococcal infection, and 6 infections in the remaining 50 people not identified.

Group A Streptococcal Infections

Focusing on only GAS infection diminished the diagnostic accuracy of FeverPAIN CPR (Supplementary Figure 1). Results largely remained the same for Centor criteria, and area under summary ROC curve between two CPRs significantly differed (χ^2 = 4.27, p=0.039). The corresponding PPV for a FeverPAIN score ≥ 4 was 18.2 % (95% CI 2.28-51.8%) compared to 20.3% (11.8-31.2%) for Centor ≥ 3.

Association between scores and likelihood of a delayed prescribing decision

Overall, 223 (39.1 %) of participants received a prescription for delayed antibiotics. Figure 4 shows a clear relationship between increasing CPR score and the proportion of patients receiving antibiotics. A score at or above the NICE threshold for considering antibiotics was significantly associated with receiving a delayed prescription; χ^2 = 32.0, p<0.001 for Centor scores ≥3 and χ^2 = 8.42, p=0.004 for FeverPAIN ≥ 4. Very few of our patients had maximal CPR scores (less than 2% had a FeverPAIN score of ≥4 and 2.5% had a Centor score of 4), suggesting these patients may have received immediate antibiotics, thus excluded from entering the trial.

However not all antibiotic prescribing was so clearly following the NICE guidance. Delayed prescriptions were frequently issued to those with low scores; 29% of those with a FeverPAIN score of 0, and 24% with a Centor score of 0.

DISCUSSION

Summary of findings

In patients presenting to primary care who are not deemed to require immediate antibiotics our results suggest that the two CPRs recommended by NICE have a limited role in deciding who has a streptococcal throat infection.

Overall, the two CPRs were equivalent, both demonstrating poor ability to discriminate between the presence/absence of streptococci, and at lower than previous estimates. For the NICE chosen thresholds for consideration of immediate or delayed antibiotics, a FeverPAIN score of ≥ 4 had a greater specificity but poorer sensitivity than Centor ≥ 3. The post-test probability of a positive swab was 6.02-61%, for FeverPAIN scores ≥ 4, a much greater range of uncertainty than 62-65% predicted by NICE. Similarly, the PPV for Centor scores of 3 or above were 16.2-37.2% compared with 32-56% expected. Given that this is a low risk population, where the clinician has already decided immediate antibiotics are not needed, using the FeverPAIN score would select a much smaller proportion of patients for potential offer of prescription 11/522 (2%) compared to Centor 74/522 (14%).

Whilst we did not know if, or how, clinicians were formally using either CPR, there was some concordance with clinical decision-making, as participants with higher FeverPAIN or Centor scores were more likely to be given delayed antibiotics. However, there was heterogeneity in practice, as antibiotics were prescribed across the range of CPRs, without a clear stepwise change at a single threshold. Overall, many more patients were given delayed prescriptions for antibiotics than either had streptococcal throat infections, or indicated by high probability CPR scores.

8

Strengths and limitations

The main strength of the study is that we had detailed data on patient characteristics, and microbiological diagnosis through throat swabs, for patients presenting to primary care with sore throat. The study was not focused on, and did not require calculation of, a CPR which may have improved the ability to capture 'natural' behaviour of clinicians. We had laboratory data on all three common isolates of streptococcus, rather than only Group A strains.

There are limitations to our findings. We had a similar number of participants to the original FeverPAIN trial but this was a secondary data analysis, not specifically powered to validate clinical prediction rules. Eleven percent of culture results were missing, thought to be at random as lost in transit to laboratory by postal services, but may have altered our results. There may have been some differences in the study population by virtue of the fact that they were recruited into a trial of oral steroids. The spectrum of disease severity in TOAST was low, due to the inclusion criteria, only 2% of patients had a high probability FeverPAIN score, compared to 8.4% in original derivation study (2). In other cohorts 96.9% of participants with FeverPAIN score 4-5 were given immediate antibiotics (15). The diagnostic accuracy of both CPRs is likely to be better in a broader spectrum of patients (7,13). Clinicians may also have been more likely to prescribe delayed antibiotics, knowing half of participants would receive immunosuppressant medication (16). Our chosen reference standard of microbiological throat swab culture is a widely accepted method of identifying streptococcal throat infections, but cannot differentiate between active disease and asymptomatic carriage (17).

Comparison to other literature

The summary ROC for both Centor and FeverPAIN was lower than in the original conception studies (5,13). A recent meta-analysis of Centor scoring across a range of clinical settings, also found a low summary area under ROC curve (0.69) and poor calibration (7). The sensitivity of a Centor score of 0 was 96.4-97.8%, higher than in our population (81.5-95%), suggesting greater confidence to rule out bacterial

9

infection. But above this the PPV remained low, encouraging the over-prescription of antibiotics in most primary care settings. FeverPAIN, during development in the Primary Care Streptococcal Management (PRISM) studies had better diagnostic accuracy compared to our study, with summary area under ROC curve across two cohorts 0.71 (0.661-0.758) and 0.735 (0.69-0.78) (9). Better performance in PRISM may be due to differences in the trial populations; prevalence of streptococci in PRISM was 34%, as those who required immediate antibiotics were not excluded, compared to 16% in TOAST (13). Children over 5 years, who may have different features of symptomatic streptococcal throat infection (18) and have greater asymptomatic carriage (17,19), were also included. The finding that a high proportion of those at low risk still receive the offer for antibiotics has been reported elsewhere (15).

Implications research and practice.

The prevalence of streptococci infections was low. Coupled with the low rate of serious complications of upper respiratory tract infections in UK primary care, this should help reinforce confidence in the clinical decision that when immediate antibiotics are not deemed necessary, they are unlikely to be required. This may be applicable not just to practitioners in GP surgeries, but also other primary care providers expected to see patients with low severity symptoms e.g. pharmacies or minor illness centres. Future research should focus on deprescribing strategies. This may include rapid antigen testing for high CPR scores, as recently piloted in Wales (20), or CPR validation studies in broader populations.

References

1.   Gulliford MC, Dregan A, Moore M V., Ashworth M, Van Staa T, McCann G, et al. Continued high rates of antibiotic prescribing to adults with respiratory tract infection: Survey of 568 UK general practices. BMJ Open. 2014 Oct 1;4(10):e006245.

2.   Little P, Stuart B, Richard Hobbs FD, Butler CC, Hay AD, Campbell J, et al. Predictors of suppurative complications for acute sore throat in primary care: Prospective clinical cohort study. BMJ. 2013 Nov 25;347.

3.   Smieszek T, Pouwels KB, Dolk FCK, Smith DRM, Hopkins S, Sharland M, et al. Potential for reducing inappropriate antibiotic prescribing in English primary care. J Antimicrob Chemother. 2018 Feb 1;73(suppl_2):ii36–43.

4.   the National Institutes of Health and Care Excellence. Overview | Sore throat (acute): antimicrobial prescribing | Guidance | NICE [Internet]. NICE; 2018.

5.   Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The Diagnosis of Strep Throat in Adults in the Emergency Room. Med Decis Mak. 1981;1(3):239–46.

6.   McIsaac WJ, Kellner JD, Aufricht P, Vanjaka A, Low DE. Empirical Validation of Guidelines for the Management of Pharyngitis in Children and Adults. J Am Med Assoc. 2004 Apr 7;291(13):1587–95.

7.   Willis BH, Coomar D, Baragilly M. Comparison of Centor and McIsaac scores in primary care: a meta-analysis over multiple thresholds. Br J Gen Pract. 2020 Mar 9;

8.   Little P, Moore M, Hobbs FDR, Mant D, McNulty C, Williamson I, et al. Primary care Streptococcal Management (PRISM) study: Identifying clinical variables associated with Lancefield group A β-haemolytic streptococci and Lancefield non-Group A streptococcal throat infections from two cohorts of patients presenting with an acute sore throat. BMJ Open. 2013;3(10).

9.   Little P, Richard Hobbs FD, Moore M, Mant D, Williamson I, McNulty C, et al. Clinical score and rapid antigen detection test to guide antibiotic use for sore

throats: Randomised controlled trial of PRISM (primary care streptococcal management). BMJ. 2013 Oct 26;347(7930).

10. Hayward GN, Hay AD, Moore M V, Jawad S, Williams N, Voysey M, et al. Effect of Oral Dexamethasone Without Immediate Antibiotics vs Placebo on Acute Sore Throat in Adults: A Randomized Clinical Trial. JAMA. 2017;317(15):1535–43.

11. Spurling GKP, Del Mar CB, Dooley L, Foxlee R, Farley R. Delayed antibiotic prescriptions for respiratory infections. Vol. 2017, Cochrane Database of Systematic Reviews. John Wiley and Sons Ltd; 2017.

12. Watson L, Little P, Moore M, Warner G, Williamson I. Validation study of a diary for use in acute lower respiratory tract infection. Fam Pract. 2001;18(5):553–4.

13. Little P, Hobbs FDR, Mant D, McNulty CAM, Mullee M, PRISM investigators. Incidence and clinical variables associated with streptococcal throat infections: a prospective diagnostic cohort study. Br J Gen Pract. 2012 Nov;62(604):e787-94.

14. Medical Test Calculator [Internet].

15. Stuart B, Brotherwood H, Van'T Hoff C, Brown A, Van Den Bruel A, Hay AD, et al. Exploring the appropriateness of antibiotic prescribing for common respiratory tract infections in UK primary care. J Antimicrob Chemother. 2020 Jan 1;75(1):236–42.

16. Pouwels KB, Dolk FCK, Smith DRM, Smieszek T, Robotham J V. Explaining variation in antibiotic prescribing between general practices in the UK. J Antimicrob Chemother. 2018 Feb 1;73(suppl_2):ii27–35.

17. Shaikh N, Leonard E, Martin JM. Prevalence of streptococcal pharyngitis and streptococcal carriage in children: A meta-analysis. Pediatrics. 2010 Sep 1;126(3):e557–64.

18. Mcisaac WJ, Md †, White D, Tannenbaum D, Low DE. A clinical score to reduce unnecessary antibiotic use in patients with sore throat. CMAJ 1998; 158(1):75-83.

19. Oliver J, Malliya Wadu E, Pierse N, Moreland NJ, Williamson DA, Baker MG. Group A Streptococcus pharyngitis and pharyngeal carriage: A meta-analysis.

PLoS Negl Trop Dis. 2018 Mar 19;12(3).

20.   Mantzourani E, Evans A, Cannings-John R, Ahmed H, Hood K, Reid N, et al. Impact of a pilot NHS-funded sore throat test and treat service in community pharmacies on provision and quality of patient care. BMJ Open Qual. 2020 Feb 1;9(1):e000833.

Diagnostic accuracy of FEVER-Pain and Centor criteria for bacterial throat infection in adults with sore throat: a secondary analysis of a randomised controlled trial.

Tables

Table 1: Throat swab culture results

	All β-Lancefield Group Streptococci†			Group A Streptococci			Group C & G Streptococci		
	Negative swab	**Positive swab**		**Negative swab**	**Positive swab**		**Negative swab**	**Positive swab**	
	(n=421)	(n=81)	**p**	(n=439)	(n=63)	**p**	(n=484)	(n=18)	**p**
Centor ≥ 3 (n=74)	55 (13%)	19 (24%)	0.01	59 (13%)	15 (24%)	0.03	77 (16%)	4 (22%)	0.33
FeverPAIN 2-3 (n=266)	192 (46%)	48 (60%)	0.02	205 (47%)	35 (56%)	0.16	253 (52%)	13 (72%)	0.03
FeverPAIN ≥4 (n=11)	8 (2%)	3 (3.7%)	0.15	9 (2%)	2 (3%)	0.42	10 (2%)	1 (6%)	0.09

Legend

† All β-Lancefield Group Streptococci to include Group A, C and G Streptococci strains isolated on agar plates after 48 hours culture.

Table 2: Diagnostic Accuracy at each threshold of Clinical Prediction Rule

Threshold	n above / below	Sensitivity	Specificity	ROC	PPV	NPV
Centor ≥ 1	424/78	90.1 (81.5, 95)	16.6 (13.2, 20.5)	0.53 (0.5, 0.57)	17.2 (13.7, 21.2)	89.7 (80.8, 95.5)
Centor ≥ 2	211/298	61.7 (50.3, 72.3)	61.8 (56.9, 66.4)	0.62 (0.6, 0.68)	23.7 (18.1, 30)	89.3 (85.2, 92.6)
Centor ≥ 3	74/428	23.5 (14.8, 34.2)	86.9 (83.3, 90)	0.55 (0.50, 0.60)	25.7 (16.2, 37.2)	85.5 (81.8, 88.7)
Centor ≥ 4	12/490	3.7 (0.77, 10.4)	97.9 (96, 99)	0.51 (0.49, 0.53)	25 (5.49, 57.2)	84.1 (80.5, 87.2)
FeverPAIN ≥ 1	434/68	93.8 (86.2, 98)	15 (11.7, 18.7)	0.54 (0.51, 0.58)	17.5 (14.1, 21.4)	92.6 (83.7, 97.6)
FeverPAIN ≥ 2	251/251	63 (51.5, 73.4)	52.5 (47.6, 57.4)	0.58 (0.52, 0.64)	20.3 (15.5, 25.8)	88 (83.4, 91.8)
FeverPAIN ≥ 3	93/409	22.2 (13.7, 32.8)	82.2 (78.2, 85.7)	0.52 (0.48, 0.57)	19.4 (11.9, 28.9)	84.6 (80.7, 88)
FeverPAIN ≥ 4	11/491	3.7 (0.77, 10.4)	98.1 (96.3, 99.2)	0.51 (0.49, 0.53)	27.3 (6.02, 61)	84.1 (80.6, 87.2)
FeverPAIN ≥ 5	2/500	1.29 (0.031, 6.69)	99.8 (98.7, 100)	0.51 (0.49, 0.52)	50 (1.26, 98.7)	84 (80.5, 87.1)

Legend
n = number of participants
ROC = Area under Receiver operating characteristics curve
PPV = positive predictive value
NPV = negative predictive value

Figures

Figure Legends

Figure 1: Proportion with throat swabs positive for Streptococcal † infection

Proportion of those with Streptococcal infection within each clinical prediction score, with 95% confidence intervals represented by black bars. Lines on the graph added for thresholds of Centor criteria ≥ 3 and FeverPAIN ≥ 4 indicating high probability of infection, where NICE supports consideration of antibiotic prescription. Data below the graph indicate the number of participants at each score, and number who were throat swab positive for Group A, C or G streptococci.

† All β-Lancefield Group Streptococci to include Group A, C and G Streptococci strains isolated on agar plates after 48 hours culture.

Figure 2: ROC Curves for diagnostic accuracy in identifying streptococcal infections† for FeverPAIN and Centor Clinical Prediction Rules.

† All β-Lancefield Group Streptococci to include Group A, C and G Streptococci strains isolated on agar plates after 48 hours culture.

Receiver Operating Characteristic (ROC) Curve analysis showing sensitivity versus 1-specificity each threshold. The different scores are shown next to a point on the graph. Each colour represents the diagnostic accuracy of one clinical prediction rule for streptococcal infection in adults with sore throat.

Figure 3: Implications of diagnostic accuracy of FeverPAIN and Centor CPR in diagnosing streptococcal throat infections in 100 adults presenting with sore throat not deemed to require immediate antibiotics.

The diagram illustrates how both CPR scores would operate if used on 100 people presenting with sore throat, based on a similar prevalence (16%) of streptococcal infection, as found in our study. Different colours indicate possible outcomes: true positive (blue), false positive (yellow), true negative (grey) and false negatives (red).

Figure 4: Delayed antibiotics prescribed in TOAST according to each clinical preduction rule score. The graph displays the percentage of participants in each Clinical Decision Score who were prescribed delayed antibiotics at presentation. The numbers within each bar indicate the number of participants who received a delayed prescription at each score.

Figure 1

Proportion

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

■ FeverPAIN
■ Centor

Centor
'High likelihood
of
Streptococcal
infection'

FeverPAIN
'High likelihood
of
Streptococcal
infection'

	0	1	2	3	4	5	
Throat Swab positive	5 8	25 21	33 31	15 16	2 3	1 -	
Total number	68 78	183 213	158 137	82 62	9 12	2 -	

Figure 2

A.Seeley et al. Sore Throat CPR secondary analysis TOAST RCT BJGP Figures & Tables
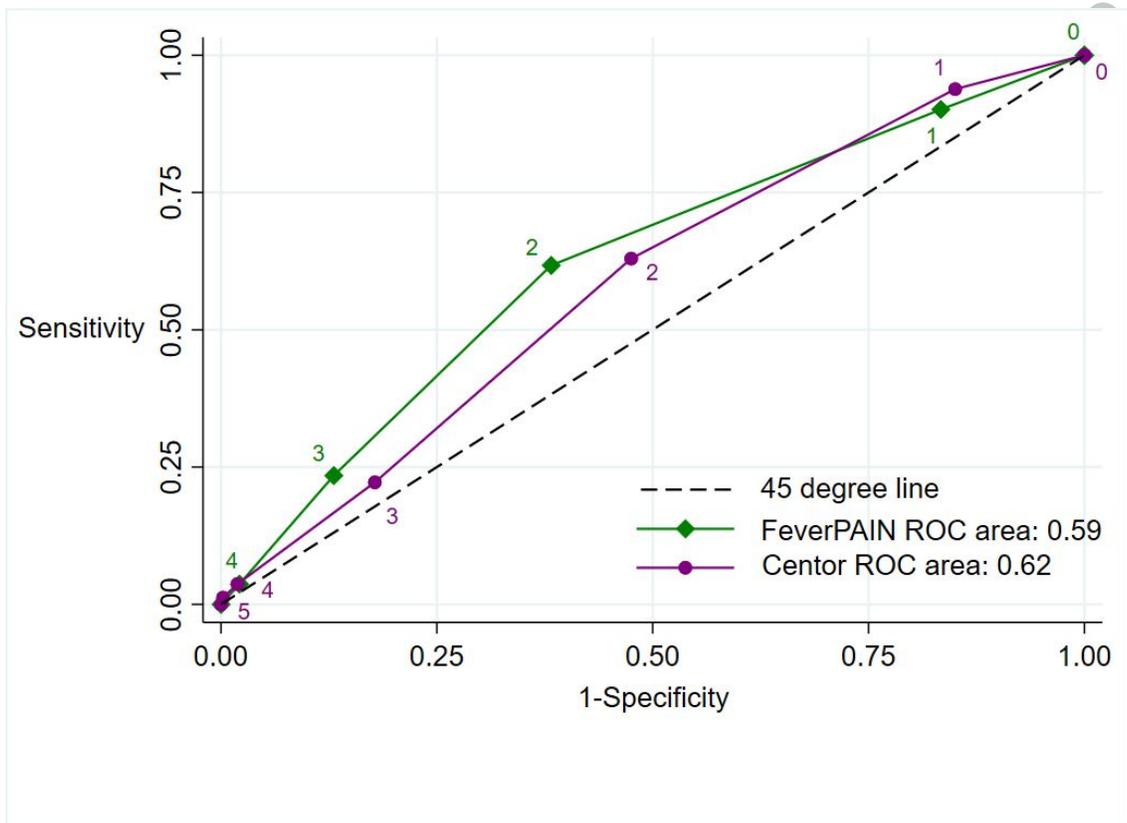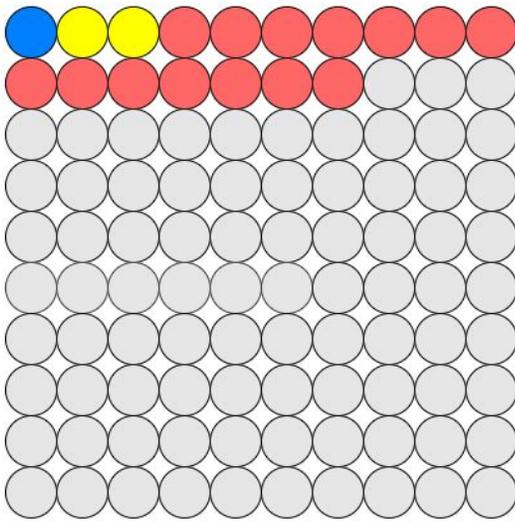
5
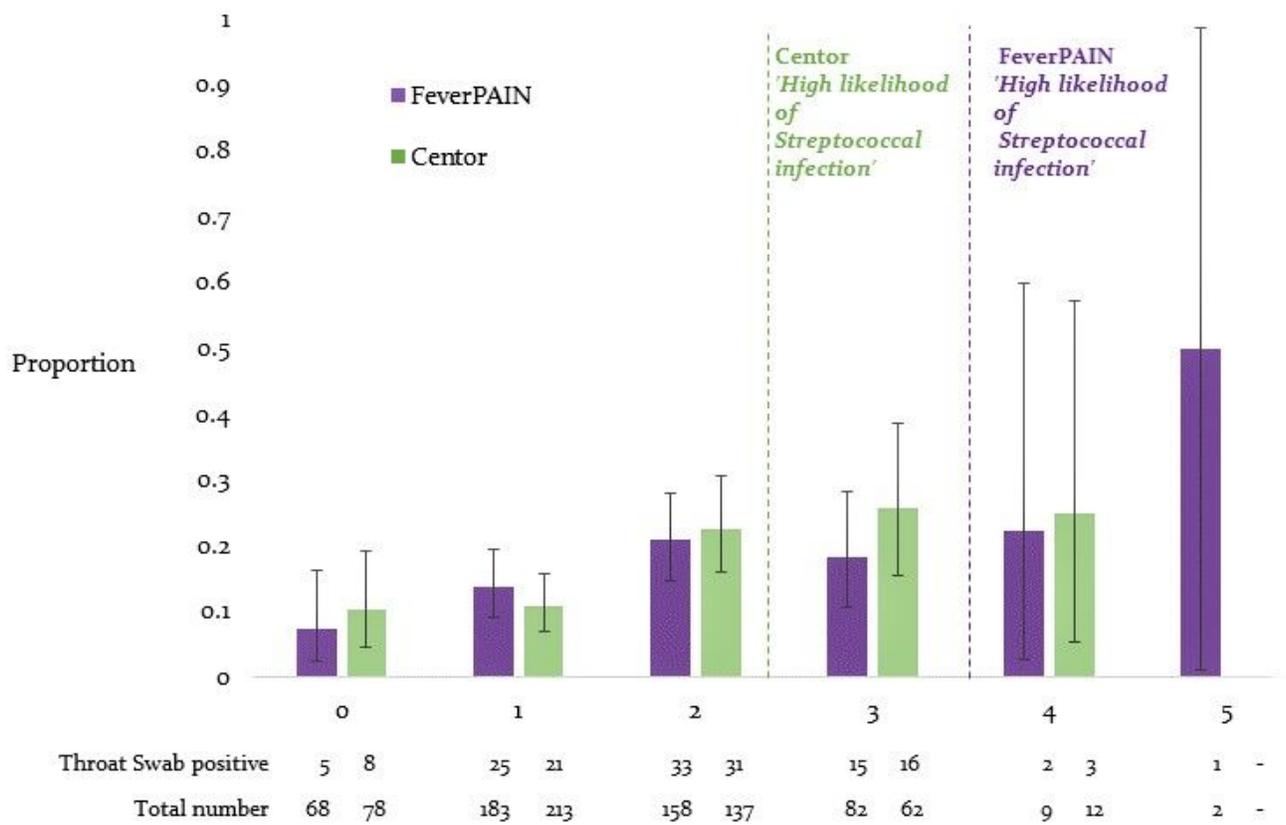
Figure 3

FeverPAIN Scoring: Using a threshold of equal or above 4 to offer delayed antibiotics

In 100 people tested:

🔵 1 person with streptococcal sore throat will be offered a delayed antibiotic prescription (true positive)

🟡 2 people will be given a delayed antibiotic prescription who do not have a streptococcal infection (false positive)

🔴 15 people with streptococcal infections will not be offered delayed prescriptions (false negative)

⚪ 82 without streptococcal infections will not be offered antibiotics (true negative)

Centor Scoring: Using a threshold of equal or above 3 to offer delayed antibiotics
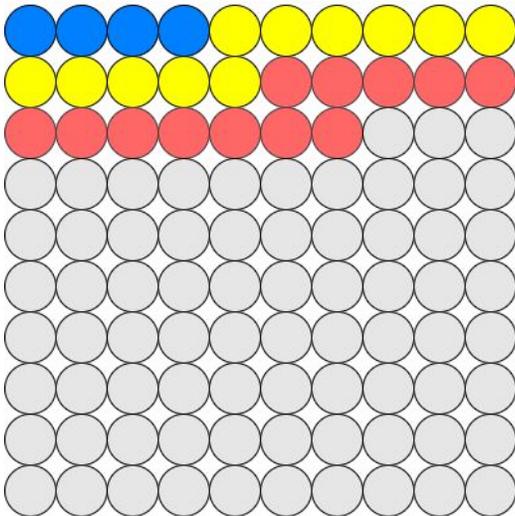
In 100 people tested:

🔵 4 people with streptococcal sore throat will be offered a delayed antibiotic prescription (true positive)

🟡 11 people will be given a delayed antibiotic prescription who do not have a streptococcal infection (false positive)

🔴 12 people with streptococcal infections will not be offered delayed prescriptions (false negative)

⚪ 73 without streptococcal infections will not be offered antibiotics (true negative)

Figure 4

A.Seeley et al. Sore Throat CPR secondary analysis TOAST RCT BJGP Figures & Tables

7