Sore Throat Assessment Tool-10 for patients with acute pharyngo-tonsillitis

Tejs Ehlers Klug

ABSTRACT

INTRODUCTION: The aim of this study was to develop and validate a patient-based questionnaire for the assessment of sore throat symptoms. The questionnaire can be used to quantify the effects of different treatments in patients with acute pharyngo-tonsillitis (APT).

METHODS: A 13-item questionnaire was constructed using a five-point Likert scale. Adult patients with APT consulting one of the participating general practitioners answered the survey. Healthy students served as controls.

RESULTS: A total of 77 sore throat patients and 103 healthy controls were included. Ten items had moderate to high (> 0.25) inter-item correlations and high mean scores (> 3.0) among patients and were included in the final tool, the Sore Throat Assessment Tool (STAT)-10. The patients' mean STAT-10 score was 36.5 (95% confidence interval (CI): 34.8-38.2) at the time of inclusion compared with 10.4 (95% CI: 10.2-10.6) for controls (p < 0.001). Cronbach's alpha coefficient for the questionnaire was 0.87. The standard error of measurement was 2.99, the intraclass correlation 0.92 and the Spearman correlation test-retest reliability 0.87. The STAT-10 scores increased with Centor scores (p = 0.018).

CONCLUSIONS: The reliability and validity of the STAT-10 were very high. The STAT-10 is the first validated tool for measuring the intensity and duration of symptoms from the perspective of sore throat patients and for quantifying and comparing different treatment modalities in APT patients.

FUNDING: This study received funding from The Lundbeck Foundation (Grant #R185-2014-2482), Fonden for Lægevidenskabens Fremme, and Hans Skouby's Fond.

TRIAL REGISTRATION: The study was approved by the Danish Data Protection Agency (2015-57-0002).

Sore throat is a highly prevalent condition affecting children and adults. Regardless of aetiology (viral or bacterial), the infection is usually self-limiting, but costly due to absence from school and work. Complications (e.g., rheumatic fever, peritonsillar abscess) are well described, but difficult to foresee and relatively rare, which makes the number needed to treat to avoid one case of complicated disease very high [1].

The management of sore throat is controversial. Paracetamol (acetaminophen) and/or non-steroidal anti-inflammatory drugs remain the mainstay for throat pain management [2]. In addition, a number of randomised clinical trials have shown a beneficial effects of corticosteroids on pain relief [3]. However, pain has been the only symptom monitored in studies of these drugs, and steroids are not commonly used by clinicians in many countries [4]. In contrast, antibiotics are widely prescribed to sore throat patients [5, 6]. Antimicrobial treatment is used to avoid complications, limit spreading of the infection and reduce the duration of illness. However, rheumatic fever has become very rare in the Western World, and studies showing a reduced risk of abscess development were non-randomised and from a time period (1950s) with a higher prevalence of complications, and some patients in these studies were treated with antibiotics not presently used [1]. Previous studies on the effects of antibiotics (as well as steroids and other treatments) focus on a limited number of symptoms (most commonly pain) [7, 8] and measure or analyse symptoms dichotomously [7-14], typically at only one time point [8-10, 15]. Therefore, the studies provide little insight into the effects of the studied drugs on all patients' symptoms and general well-being. Because of the divergent and limited outcomes used, the literature fails to answer which treatments are the most effective in reducing the intensity and duration of symptoms. Furthermore, sore throat symptoms may seem straight forward, but it remains unexplored if patients were asked the right questions as no studies have been conducted on the symptoms to be used in sore throat patient questionnaires.

To inform future studies, the present study was undertaken to develop and validate a patient-based questionnaire for the assessment of sore throat symptoms that can be used to quantify the effects of different treatment modalities in patients with acute pharyngo-tonsillitis (APT).

METHODS

Item generation

A group of sore throat experts (ear-nose-throat specialists and general practitioners with clinical and scientific expertise in APT) and patients contributed with each 5-10 items he or she deemed to have excellent validity for the assessment of sore throat. The final number of items was achieved by reducing the initial

ORIGINAL ARTICLE

Department of Otorhinolaryngology, Head and Neck Surgery, Aarhus University Hospital, Denmark

Dan Med J 2019;66(9):A5561

high number of proposed items because of repetitions and overlaps. A list of thirteen items was chosen by the author (Sore Throat Assessment Tool (STAT)-13) (**Table 1**).

Patients and controls

Patients aged 15-40 years consulting one of the 25 participating general practitioners with acute (duration of symptoms < 8 days) pharyngo-tonsillitis (objective finding of intensified tonsillar colour, oedema or exudate) without suspected complications were asked to answer the on-line questionnaire at the time of inclusion and two hours later (patients responded to automatically delivered e-mails). The questionnaire was conducted in Danish. The following data were obtained in all patients: objective temperature, Centor score (one point for each of anamnestic fever, absence of cough, tonsillar exudates and tender cervical lymph nodes), duration of symptoms and results of streptococcal rapid antigen detection test. Medical students aged 15-40 years without newly (within past month) arisen throat symptoms served as controls. The exclusion criteria for both groups were history of radiotherapy, gastro-oesophageal reflux, chronic pain or daily use of analgesics (except for current throat pain); and neoplastic, haematologic, rheumatologic or neurologic disorders. The study was conducted in the period from 1 September 2017 to 1 June 2018.

Item reduction and reliability analysis

Inter-item correlations were estimated for the 13 items using Spearman correlations to evaluate item redundancy. Test-retest reliability was evaluated by administering the survey on two separate occasions to normal controls (seven days between administration) and to sore throat patients (two hours between administration). The short time interval between survey administrations to sore throat patients was chosen as a trade-off between the potential for improvement of symptoms (e.g., use of analgesics) after consulting the general practitioner and the risk of recall bias. In addition to the time points described above, patients with a positive test to *Streptococcus* group A (GAS) antigen answered the survey seven days after inclusion. The test-retest reliability was further evaluated by a Spearman correlation between the two scores. The internal consistency was assessed using Cronbach´s alpha. In most applied studies, the lowest acceptable level of internal consistency is 0.7 for group level and 0.9 for individual analysis.

Validity and responsiveness analysis

The validity of the tool was examined via an unmatched case-control contrast between sore throat patients and healthy controls. The responsiveness of the tool was assessed by examining the scores before (day one) and after (day eight) antibiotic treatment of GAS-positive patients.

Missing data were imputed from the mean of completed items, provided that > 50% of the items had been completed. Statistical analyses were performed using Student's t-test, Fisher's exact test, the Kruskal-Wallis test and analysis of variance (ANOVA). Statistical significance was defined as $p < 0.05$.

Trial registration: The study was approved by the Danish Data Protection Agency (2015-57-0002). All participants (and parents of minors) gave their informed consent prior to their inclusion in the study. According to Danish law, notification of the study to the local ethical committee was not required.

RESULTS

Patient characteristics

A total of 77 patients with APT (21 GAS-positive) and 103 healthy controls answered the questionnaire. Patients were significantly older (mean 28.3 years ± standard deviation (SD) 6.6 years) than controls (mean 25.6 years, ± SD 1.4) ($p < 0.001$, student's t-test). The proportion of females was 71% among patients and 68% among controls ($p = 0.64$, Fisher's exact test). Patient characteristics are outlined in **Table 2**.

Removal of redundant items

In the patient group, items # 9 (difficulty breathing) and # 13 (cough) had low correlations (< 0.25) with the majority of other items (**Table 3**), and they were therefore removed from the survey instrument. The other items had acceptable correlation coefficients (between 0.25 and 0.8). In addition, item # 12 (difficulty opening my mouth) was removed from the survey instrument because of a low mean score among patients (Table 1).

The mean scores for all items were very low in the control group (Table 1). The STAT-10 was formed after removing the three mentioned items (# 9, 12, and 13) (Table 1).

Reliability

Patients' mean STAT-10 scores were 36.5 (95% confidence interval (CI): 34.8-38.2) and 34.6 (95% CI: 32.6-36.5) at the time of inclusion and two hours later, respectively ($p < 0.001$, Student's t-test).

The Cronbach´s alpha coefficient for the questionnaire was 0.87 for the first measurement and 0.91 for the second measurement, which exceed the minimum acceptable value for group level analysis and suggests that each item addresses the same construct. The standard error of measurement was 2.99 (95% CI: 2.58-3.55) and the intraclass correlation was 0.92

TABLE 1

The Sore Throat Assessment Tool (STAT)-13 (all items) and STAT-10 (final questionnaire)[a]. The mean scores at the time of inclusion for 77 patients and 103 controls are given.

Item #	At present, to what degree do you experience the following symptoms and are the following daily activities restricted?	Patients, mean score	Controls, mean score	Not at all	To a low degree	To a moderate degree	To a high degree	To a very high degree
STAT-13 and STAT-10								
1	My throat hurts	4.12	1.05	1	2	3	4	5
2	I have pain on swallowing	4.25	1.01	1	2	3	4	5
3	I feel uncomfortable	3.75	1.05	1	2	3	4	5
4	I have difficulty attending my work/school	3.46	1.03	1	2	3	4	5
5	I have reduced my level of activity	3.74	1.06	1	2	3	4	5
6	I have difficulty swallowing	3.86	1.02	1	2	3	4	5
7	I have a decreased appetite	3.09	1.02	1	2	3	4	5
8	I have a headache	3.16	1.17	1	2	3	4	5
10	I feel feverish	3.17	1.01	1	2	3	4	5
11	I feel ill	3.90	1.02	1	2	3	4	5
Only STAT-13								
9	I have difficulty breathing	1.88	1.08	1	2	3	4	5
12	I have difficulty opening my mouth	1.92	1.03	1	2	3	4	5
13	I have a cough	2.34	1.11	1	2	3	4	5

a) The study was performed in Danish (see supplemental online material for Danish version), the questionnaire wordings were evaluated by sore throat patients and items were arranged on a five-point Likert scale.

(95% CI: 0.84-0.95). The spearman correlation test-retest reliability was 0.87.

Normative data generation

The controls' mean STAT-10 scores were 10.4 (95% CI: 10.2-10.6) and 11.1 (95% CI: 10.5-11.8) at the time of inclusion and seven day later, respectively ($p = 0.025$, Student's t-test). The Cronbach´s alpha coefficient for the questionnaire was 0.59 for the first measurement and 0.90 for the second measurement. The standard error of measurement was 3.12 (95% CI: 2.74-3.61) and the intraclass correlation was 0.26 (95% CI: –0.80-0.49).

Validity

The differences in mean scores in the individual items of the STAT-10 between patients and controls (at the time of inclusion) were in the 1.99-3.24 range ($p < 0.001$ for all individual items, Student's t-test).

The mean STAT-10 scores for GAS-positive and GAS-negative patients were 38.7 (95% CI: 35.1-42.3) and 35.4 (33.5-37.4), respectively ($p = 0.092$, Student's t-test). The STAT-10 scores increased significantly with Centor scores (Table 2, $p = 0.018$, Kruskal-Wallis test). No statistically significant associations between STAT-10 scores and gender ($p = 0.29$, Student's t-test), duration of symptoms ($p = 0.92$, ANOVA) and age ($p = 0.46$, ANOVA) were found.

TABLE 2

Characteristics of 77 patients with acute pharyngo-tonsillitis and stratified mean STAT-10 scores at inclusion.

	Patients	Mean STAT-10 score
Age, mean (± SD), yrs	28.3 (± 6.6)	-
Gender, n		
Females	55	37.0
Males	22	34.9
Temperature, mean (± SD), °C	37.3 (± 1.0)	-
Centor score[a], n		
1	7	30.9
2	28	34.5
3	30	38.1
4	12	39.8
Duration of symptoms, mean (± SD), days	3.4 (± 1.8)	-
Streptococcal RADT, n		
Positive	21	38.7
Negative	56	35.4

RADT = rapid antigen detection test; STAT = Sore Throat Assessment Tool; SD = standard deviation.
a) Patients receive 1 point for: anamnestic fever, absence of cough, presence of tonsillar exudates, and tender cervical lymph nodes.

Responsiveness analysis

Fifteen of 21 GAS-positive patients completed the questionnaire seven days after inclusion. The STAT-10 score for these 15 patients was 40.1 (95% CI: 36.0-44.1) at time of inclusion and 11.8 (95% CI: 10.0-13.6) at fol-

TABLE 3

Inter-item correlations (Spearman) between the 13 items in Sore Throat Assessment Tool-13.

	Item #												
Item #	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.0												
2	0.60	1.0											
3	0.41	0.23[a]	1.0										
4	0.58	0.38	0.64	1.0									
5	0.29	0.27	0.50	0.73	1.0								
6	0.54	0.79	0.28	0.32	0.25	1.0							
7	0.42	0.47	0.45	0.43	0.31	0.51	1.0						
8	0.17[a]	0.11[a]	0.45	0.42	0.41	0.13[a]	0.35	1.0					
9	0.12[a]	0.14[a]	0.39	0.21[a]	0.18[a]	0.17[a]	0.38	0.36	1.0				
10	0.31	0.23[a]	0.49	0.49	0.47	0.29	0.41	0.45	0.24[a]	1.0			
11	0.49	0.40	0.66	0.70	0.62	0.40	0.47	0.43	0.32	0.75	1.0		
12	0.35	0.41	0.35	0.31	0.23[a]	0.45	0.40	0.28	0.40	0.27	0.43	1.0	
13	−0.12[a]	0.02[a]	0.09[a]	0.01[a]	0.04[a]	−0.04[a]	0.03[a]	0.16[a]	0.43	0.03[a]	0.06[a]	0.20[a]	1.0

a) Low (< 0.25) correlation.

low-up seven days later ($p < 0.001$, Student's t-test). The mean follow-up STAT-10 scores for GAS-positive patients was insignificantly different from the mean follow up STAT-10 scores of controls ($p = 0.45$, Student's t-test).

DISCUSSION

The reliability of final questionnaire, the STAT-10, was found to be excellent based on the calculated Cronbach´s alpha coefficient, intraclass correlation, Spearman correlation test-retest reliability and standard error of measurement. Though the mean STAT-10 scores decreased significantly (by 1.9 points) from the time of inclusion to the second answer two hours later, both the initial and the second scores were very much higher in patients than controls (mean differences: 26.1 and 23.5 points, respectively) and, hence, the reliability of the survey remains intact. The validity of STAT-10 was also found to be good based on the highly significant differences in the individual items and summarised scores between patients and controls. Furthermore, mean STAT-10 scores increased significantly with Centor scores [16], and there was a trend towards higher scores in GAS-positive patients than in GAS-negative patients, suggesting that the intensity of symptoms was higher in patients with bacterial infection than in patients with viral infection. It is noteworthy that the intensity of symptoms, including Centor score 1, was rather high in all patient categories.

The Cronbach´s alpha coefficient and the intraclass correlation were low for the control group, which was expected as the survey items were designed to reflect symptoms in sore throat patients and not in healthy individuals.

The items of the STAT-10 can be categorised into three domains: 1) Throat symptoms (items # 1, 2, and 6), 2) Other symptoms (items # 7, 8, and 10) and 3) General (lack of) well-being (items # 3, 4, 5, and 11). Not surprisingly, the mean scores of throat symptoms were higher (4.08) than those of general well-being (3.71) and other symptoms (3.14). However, patients scored relatively high in all three domains. Combined with the fact that inter-item correlations were found to be appropriately high, the findings stress that more items are needed to describe patients' illness and for the evaluation of treatment efficacy.

The STAT-10 was developed for the utility in future studies on the effects of different treatments (e.g., antibiotics, glucocorticoids) of sore throat. Using the area the under the curve for diagrams with survey scores on the y-axis and time on the x-axis, the STAT-10 constitutes a sensitive instrument for measuring treatment efficacy relative to intensity and duration of symptoms. Furthermore, the development of a validated tool will allow for inter-study comparisons and reliable meta-analyses.

The study carries several limitations. The study was conducted in Denmark and the questionnaire answered by the patients was prepared in Danish. Hence, the survey wordings in Table 1 is an English translation, and it is possible that the questions would be interpreted and answered slightly differently by English-speaking patients. Adolescents and adults aged 15-40 years were included in the study and symptoms may be different in children. Children were disregarded to avoid potentially conflicting symptoms between adults and children and to eliminate the risk of children misunderstanding the questions. Patients were significantly

older (average 2.7 years) than controls. However, no association between age and STAT-10 scores was found, and it seems unlikely that this age difference had a significant impact on study results. The number of participants was relatively limited and subgroup findings and comparisons should be interpreted with caution.

However, the level of statistical significance between patients and controls for all individual items of STAT-10 was extremely high ($p < 0.001$), and the addition of more patients is very unlikely to alter the conclusions concerning the reliability and validity of the instrument. The same argument opposes the potential bias introduced by sociocultural differences between patients and controls. As previously mentioned, the mean STAT-10 scores decreased significantly from the time of inclusion (36.5) to follow up two hours later (34.6), which likely reflects the effects of analgesics recommended by the general practitioner and the reduced distress after receiving doctoral care and assessment. The rapidly reducing symptoms underline the limited time interval for re-testing.

In conclusion, the present study is the first to develop and validate a questionnaire to be used in future treatment studies of patients with APT. The STAT-10 is a reliable and sensitive tool for measuring the intensity and duration of symptoms from the perspective of sore throat patients and for evaluating and comparing different treatment modalities.

CORRESPONDENCE: *Tejs Ehlers Klug.* E-mail: tejsehlersklug@hotmail.com

ACCEPTED: 3 June 2019

CONFLICTS OF INTEREST: none. Disclosure forms provided by the authors are available with the full text of this article at Ugeskriftet.dk/dmj

AKNOWLEDGEMENTS: *Therese Koops Grønborg*, Aarhus University, for statistical assistance, *Maria Rusan,* Aarhus University Hospital, for English translation of the questionnaire.

LITERATURE

1. Spinks A, Glasziou PP, Del Mar CB. Antibiotics for sore throat. Cochrane Database Syst Rev 2013;11:CD000023.
2. Thomas M, Del Mar C, Glasziou P. How effective are treatments other than antibiotics for acute sore throat? Br J Gen Pract 2000;50:817-20.
3. Sadeghirad B, Siemieniuk RAC, Brignardello-Petersen R et al. Corticosteroids for treatment of sore throat: systematic review and meta-analysis of randomised trials. BMJ 2017;358:j3887.
4. Andre M, Gröndal H, Strandberg EL et al. Uncertainty in clinical practice – an interview study with Swedish GPs on patients with sore throat. BMC Fam Pract 2016;17:56.
5. Run Sigurethardottir N, Nielsen AB, Munck A et al. Appropriateness of antibiotic prescribing for upper respiratory tract infections in general practice: Comparison between Denmark and Iceland. Scand J Prim Health Care 2015;33:269-74.
6. Linder JA, Chan JC, Bates DW. Evaluation and treatment of pharyngitis in primary care practice: the difference between guidelines is largely academic. Arch Intern Med 2006;166:1374-9.
7. Zwart S, Sachs AP, Ruijs GJ et al. Penicillin for acute sore throat: randomised double blind trial of seven days versus three days treatment or placebo in adults. BMJ 2000;320:150-4.
8. Nelson JD. The effect of penicillin therapy on the symptoms and signs of streptococcal pharyngitis. Pediatr Infect Dis 1984;3:10-3.
9. De Meyere M, Mervielde Y, Verschraegen G et al. Effect of penicillin on the clinical course of streptococcal pharyngitis in general practice. Eur J Clin Pharmacol 1992;43:581-5.
10. Dagnelie CF, van der Graaf Y, De Melker RA. Do patients with sore throat benefit from penicillin? A randomized double-blind placebo-controlled clinical trial with penicillin V in general practice. Br J Gen Pract 1996;46:589-93.
11. Little P, Gould C, Williamson I et al. Reattendance and complications in a randomised trial of prescribing strategies for sore throat: the medicalising effect of prescribing antibiotics. BMJ 1997;315:350-2.
12. Middleton DB, D'Amico F, Merenstein JH. Standardized symptomatic treatment versus penicillin as initial therapy for streptococcal pharyngitis. J Pediatr 1988;113:1089-94.
13. Petersen K, Phillips RS, Soukup J et al. The effect of erythromycin on resolution of symptoms among adults with pharyngitis not caused by group A streptococcus. J Gen Intern Med 1997;12:95-101.
14. Marlow RA, Torrez AJ, Haxby D. The treatment of nonstreptococcal pharyngitis with erythromycin: a preliminary study. Fam Med 1989;21:425-7.
15. Howe RW, Millar MR, Coast J et al. A randomized controlled trial of antibiotics on symptom resolution in patients presenting to their general practitioner with a sore throat. Br J Gen Pract 1997;47:280-4.
16. Centor RM, Witherspoon JM, Dalton HP et al. The diagnosis of strep throat in adults in the emergency room. Med Decis Making 1981;1:239-46.