

A systematic review of screening diagnostic tools for trigeminal neuralgia

THN Teshima^{1,2} , JM Zakrzewska^{1,2}  and R Potter³

British Journal of Pain
2023, Vol. 0(0) 1–12
© The Author(s) 2023



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/20494637221146854

journals.sagepub.com/home/bjp



Abstract

Background and objective: Trigeminal neuralgia (TN) is a rare chronic neuropathic pain condition of sudden and severe pain, often described as an electric shock. Diagnosis is challenging for non-expert clinicians, particularly in primary care settings. We wanted to identify and assess the diagnostic accuracy of existing screening tools for TN and orofacial pain that could be used to support the diagnosis of TN in primary care.

Databases and data treatment: We searched key databases (MEDLINE, ASSIA, Embase, and Web of Knowledge and PsycINFO) supplemented by citation tracking from January 1988 to 2021. We used an adapted version of the Quality of Diagnostic Accuracy Studies (QUADAS-2) to assess the methodological quality of each study.

Results: Searches identified five studies, from the UK, USA and Canada; three validated self-report questionnaires; and two artificial neural networks. All screened for multiple orofacial pain diagnoses, including dentoalveolar pain, musculoskeletal pain (temporomandibular disorders) and neurological pain (trigeminal neuralgia, headache, atypical facial pain and postherpetic neuralgia). The overall quality assessment was low for one study.

Conclusions: Diagnosis of TN can be challenging for non-expert clinicians. Our review found few existing screening tools to diagnose TN, and none is currently suitable to be used in primary care settings. This evidence supports the need to adapt an existing tools or to create a new tool for this purpose. The development of an appropriate screening questionnaire could assist non-expert dental and medical clinicians to identify TN more effectively and empower them to manage or refer patients for treatment more effectively.

Keywords

orofacial pain, trigeminal neuralgia, diagnosis, questionnaire, systematic review

Objectives

- To identify screening tools used to classify or diagnose trigeminal neuralgia in patients presenting with non-diagnosed orofacial pain symptoms.
- To identify studies that develop or test clinical prediction rules for the diagnosis of trigeminal neuralgia and exclude other types of orofacial pain.
- To determine the test accuracy of existing tools used to classify or diagnose trigeminal neuralgia and identify positive and negative predictive values for individual diagnostic features.

¹Department of Oral Medicine, UCL Eastman Dental Institute, London, UK

²Department of Oral Medicine and Pain Education Research Centre at ULCH NHS Foundation Trust, Royal National ENT & Eastman Dental Hospitals, ULCH NHS Foundation Trust, London, UK

³Warwick Clinical Trials Unit, Coventry, UK

Corresponding author:

THN Teshima, Department of Oral Medicine, UCL Eastman Dental Institute, Bloomsbury Campus, Rockefeller Building, 21 University Street, 3rd floor, London WC1E 6DE, UK.

Email: t.teshima@ucl.ac.uk

Introduction

Trigeminal neuralgia (TN) is a rare chronic neuropathic pain condition that causes sudden, usually unilateral, recurrent, severe pain, often described by patients as feeling like an electric shock. It can involve one or more branches of the trigeminal nerve, and the severity of pain can lead to important psychological and physical consequences if not treated appropriately.^{1,2}

The diagnosis of TN is based largely on the personal description of pain and symptoms the patient provides. There are established and very specific diagnostic criteria for TN,³⁻⁵ but diagnosis can be a challenge for non-expert clinicians especially in the early phases. Chronic orofacial pain disorders present with a significant complexity of symptoms that can overlap between diagnoses, and TN can be frequently confused with more common conditions such as dental pain or temporomandibular disorders. This leads to patients often consulting with a dental practitioner in the first instance and potentially having unnecessary irreversible dental treatments.⁶⁻⁸ Patients seeking diagnosis from medical practitioners and neurosurgical services are also often misdiagnosed.⁹ It can sometimes take years for a patient to get an accurate diagnosis and start effective treatment, which in turn increases healthcare costs and negatively impacts the patients' experience and management.⁶

Accurate screening of orofacial pain diagnosis therefore holds a crucial role in avoiding non-indicated procedures that may worsen orofacial pain and lead to inappropriate specialist referral, diagnostic delay and misdiagnosis. This is particularly important for TN, which is known to negatively impact mental health, can result in inadequate irreversible treatments, increase healthcare consultations and associated costs, and is strongly reliant on a biopsychosocial management.⁶ The need for an appropriate diagnostic tool for TN is in line with a broader perspective regarding a lack of suitable tools to assist healthcare professionals to avoid diagnostic errors in primary care. This was recently identified as a high-priority patient safety problem in primary care by the World Health Organization.¹⁰ The evidence was demonstrated by a recent large retrospective review that analysed over 2000 records across 21 general medical practices in the UK, which found an incidence of missed diagnostic opportunity of over 60%.¹¹

A study from a neurosurgical department in Germany reported that many dental procedures were performed in over 60% of patients who first sought dental assessment such as dental extractions, root canal treatments and restorative work including implants, fillings and tooth replacement procedures (82% of patients) for undiagnosed TN.⁷ This included just over fifty participants who underwent surgical treatment for

TN, of which more than 80% had their initial consultation with dentists, and from which only less than 5% included TN as differential diagnosis. Only two of these patients reported to have had an initial differential diagnosis of TN. Almost half of the patients had multiple consultations with other medical and dental specialties prior to referral to neurosurgery, which happened within 6 months of the onset of symptoms, but more than 2 years elapsed prior to neurosurgical attention in 42.5% of patients. Although no differences in outcome was reported between those receiving dental treatments or not, this highlighted the importance of screening tools to assist the initial diagnosis and inform management. Corroborating these findings, a systematic review of a rare group of trigeminal cephalalgias found that misdiagnosis resulted in irreversible dental treatments, as well as other sinusal surgeries and patients received many inappropriate medications.¹²

Conversely, a recent observational study from a neurosurgical centre in Oslo showed an overestimation of TN diagnosis in first time referral patients, where only 21% of participants (17 out of 38 in total) actually had confirmed diagnosis of TN and fulfilled the current international diagnostic criteria.⁹ It is worth noting that all of them were also seen by a dental practitioner prior to referral. Patients with misdiagnosed TN were mostly presenting with dental pain or temporomandibular disorders according to clinical history, and only one of 21 participants had TN as true differential diagnosis. Importantly, of this group, more than 60% of patients had an overestimation of neurovascular contact on the brain MR scan, which is the standard assessment to indicate the need for a microvascular decompression.

Considering the significant impact of misdiagnosis which can result in inappropriate treatments for patients, there is a need for a diagnostic tool that can support health professionals to identify and diagnose TN more effectively, and in particular, a tool that can be used by non-expert clinicians in primary care. Our scoping review identified two tools developed to diagnose TN, the first an online questionnaire to predict diagnosis of patients with different facial pains,¹³ which was validated with patients attending a neurosurgery clinic, the majority of whom already had a diagnosis of TN. The second a classification and diagnostic grading system for TN³ aimed at experts in specialist centres rather than non-expert clinicians in primary care.

The aim of this systematic review is to identify all existing diagnostic tools that support the diagnosis of TN and to assess their suitability for use in primary care. This is an essential step in identifying whether there is a need to standardise or validate an existing tool or to develop a new one that can assist primary care practitioners to identify TN more effectively.

Methods

We registered the review prospectively with the International Prospective Register of Systematic Reviews (PROSPERO, Dec 2021: CRD42021273835) and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidance (PRISMA) for the reporting of systematic reviews.

Search strategy and study selection

The search strategy was developed in collaboration with a medical librarian. We included studies that described the validation of tools intended to screen, diagnose or classify trigeminal neuralgia and that compared with a reference standard. Studies published in English from January 1988, the date the International Classification of Headache Disorders first published the diagnostic criteria and classification of TN,¹⁴ until the date of the literature search in December 2021 were included.

We excluded studies with participants below the age of 18 years, dissertations and conference papers, and any studies where the sole purpose of the study was to report the validation of a tool in a different language. We searched the key electronic databases MEDLINE, Embase, Web of Knowledge and PsycINFO in December 2021. We used free text and subject heading terms for trigeminal and orofacial pain:

‘trigeminal neuralgia’, ‘trigeminal pain’, ‘trigeminal nerve’, ‘trigeminal nerve disease’, ‘tic doreux’, ‘orofacial pain’, ‘facial pain’, and combined with ‘classification’, ‘diagnostic’, ‘diagnosis’, ‘diagnostic techniques and procedures’, ‘sensitivity’, ‘specificity’, ‘efficacy’, ‘effectiveness’, ‘efficiency’, ‘predictive value of tests’, ‘likelihood function’, ‘diagnostic odds ratio’, and ‘screening’, ‘questionnaire’, ‘survey’, ‘interview as topic’, ‘checklist’, ‘tools’, ‘instruments’ and ‘ICHD’.

Results were managed using Rayyan systematic review software. Duplicates were removed, and titles checked for relevance by two authors (THNT and RP). We sought a full copy of possible relevant papers; these were assessed independently for inclusion by the two authors and any disagreement resolved by discussion or a third reviewer (JZ) when required. We checked reference lists of relevant papers for any additional studies.

Data extraction and quality assessment

We developed a pro-forma to capture study and tool specific information: author, year of publication, title, country/setting, study design, characteristics of study participants, a brief description of the tool, test characteristics, diagnostic/classification parameters,

validation assessment (level of agreement, sensitivity, specificity, positive and negative predictive values), how the tool was used (questionnaire, online, interview) and by whom (expert, non-expert, patient).

Data were extracted independently by two authors (THNT and RP) who also independently assessed the methodological quality of each study. We used an adapted version of the Quality of Diagnostic Accuracy Studies (QUADAS-2), a validated tool used for the quality assessment of diagnostic accuracy studies.¹⁵ The QUADAS-2 evaluates risk of bias in four domains: patient selection (sampling and exclusions), index test (conduct and interpretation), reference standard (conduct and interpretation) and flow and timing (interval between index and reference standard, number receiving reference standard and included in the analysis). The authors of the QUADAS-2 stress that the tool should be tailored for each review, and the questions we included are shown in [Table 1](#).

We assessed the risk of bias as low, high or unclear for each domain and calculated an overall risk of bias dependent on the number of domains judged as high risk of bias: 0 domain = low, 1 = low/medium, 2 = medium, 3 = medium/high and 4 = high risk of bias. Where criteria used to judge risk for one of the domains was unclear, the risk was considered ‘high’, and this was denoted in the overall risk of bias.

Results

Study selection

The systematic search identified 3314 citations after removing duplicates ([Figure 1](#)). From these, initial screening excluded 3275 titles, the main reasons for exclusions were conditions other than TN and/or orofacial pain, other study design or publication type and articles in other languages than English. Full-text copies of thirty-nine studies were then assessed for eligibility, and a further 34 articles excluded due to publication type ($n = 4$), other conditions ($n = 11$), validation of same tool in different language ($n = 2$), non-validation studies ($n = 10$) and non-diagnostic tool studies ($n = 7$). The final five studies were included in the qualitative synthesis.^{13,16–19}

Study characteristics

A summary of overall study characteristics of the selected studies is presented in [Table 2](#). Studies were published between 1994 and 2018; three studies were conducted in the USA,^{13,18,19} one in Canada¹⁷ and one in the United Kingdom.¹⁶ Four studies included multiple orofacial pain diagnoses and one was focused only on neuropathic

Table 1. Summary of quality assessment according to QUADAS-2 tool within each study and specific questions included in each domain.

Domains of risk assessment							
	Patient selection	Index test	Reference standard	Flow and timing			
Reference	Was there consecutive or random sample of patients?	Index test results blinded to the results of the reference standard?	Is the reference standard likely to correctly classify the target condition?	Were the reference results interpreted without knowledge of the results of the index test?	Did patients receive the same reference standard?	Were all patients included in the analysis?	Overall
Jafree et al.	Case-control design avoided?	Inadequate exclusions avoided?	Pre-specified threshold?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Have all patients received reference standard?	Were all patients included in the analysis?	Low/medium
Limonadi et al.	Low	Low	Low	Low	High	Low	Low
McCartney et al.	Low	Low	High	Low	High	Medium	Medium
Babiloni et al.	High	High	Low	Low	High	High	Medium/high
Hapak et al.	Low	Low	Low	Low	High	High	Low/medium

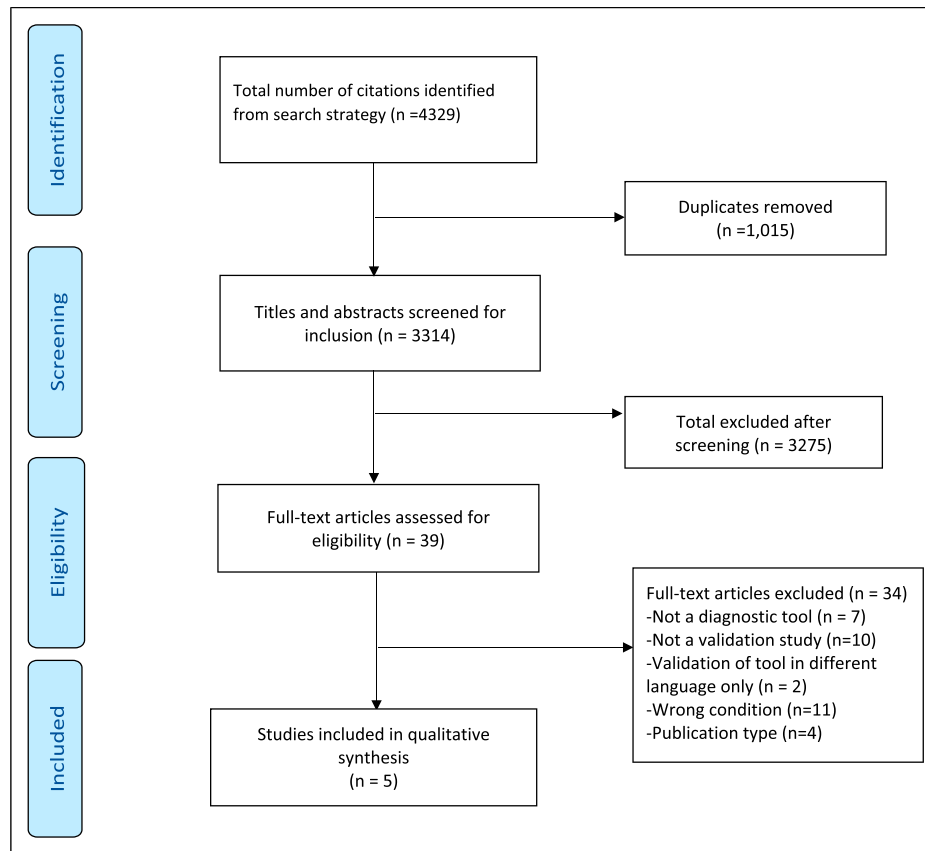


Figure 1. PRISMA flow chart of article selection.

conditions.¹⁸ The number of participants included in the validation studies ranged from 43 to 607.

The types of validated tools included three self-complete questionnaires: the PainDETECT (PD-Q),¹⁶ the modified version of the Leeds Assessment of Neuropathic Symptoms and Signs pain scale (S-LANSS)¹⁸ and a screening questionnaire based on McGill scale.¹⁷ Two studies validated artificial neural networks (ANNs),^{13,19} an online-based system composed of artificial intelligence trained to provide orofacial pain diagnoses according to responses of self-reported questionnaires.

The tools assessed a range of orofacial pain disorders, including musculoskeletal pain such as temporomandibular disorders (TMD), dentoalveolar pain and/or persistent dentoalveolar pain (PDAP), TN and variants (TN1, TN2), postherpetic neuralgia (PHN), glossopharyngeal neuralgia (GPN), atypical facial pain (AFP, trigeminal neuropathic pain (TNP), trigeminal deafferentation pain (TDP), symptomatic TN (multiple sclerosis, STN), nervus intermedius neuralgia (or geniculate neuralgia, NIN), cluster headache, migraine headache

and temporal arteritis, compared to pain-free controls.

Three studies used a well-established reference test based on previously published classification criteria.^{16–18} The remaining two reference tests were based on neurosurgical clinician expertise.^{13,19}

The number of questions from each questionnaire included in this review ranged from 7 to 22 questions. None of the questionnaires addressed all the standards needed to fulfil the minimum diagnostic criteria for classical TN according to current classification by ICOP-1 and ICHD-3.^{1,4} Table 3 summarises which diagnostic characteristics were addressed by each questionnaire and if it excludes other possible common orofacial pain conditions.

The method of data collection by each study was mostly prospective. The exception was from the modified S-LANSS study, in which this was unclear and recruited participants already had a previous diagnosis of orofacial pain.¹⁸ In the study by Hapak et al. it was not clear whether it was prospective but implies that participants would have completed the index questionnaire prior to the first assessment.¹⁷

Table 2. Characteristics of included validation studies of screening tools for diagnosing trigeminal neuralgia and/or other orofacial pain.

Reference	Year	Country	Name of index test	Number of questions	Who administered the index test	Who diagnosed clinically	How test was provided	Number of participants	Reference test
Jafree et al.	2018	UK	painDETECT (PD-Q)	8 questions + 4 pain course patterns as pictures	Postgraduate specialty dentists on site	Independent facial pain clinicians, blinded	Paper questionnaire	251 new patients	International Classification of Headache Disorders (ICHD) + CD/TMD classification for TMD 2014 Expert clinicians in the field
Limonadi et al.	2006	US	Facial pain questionnaire (none specified) and artificial neural network (ANN)	18	Senior investigator (author)	Same senior investigator, blinded to questionnaire (author)	Paper questionnaire + expert entered electronic data into the ANN	143 new patients	Expert clinicians in the field
McCartney et al.	2014	US	Facial pain questionnaire (none specified) and artificial neural network (ANN)	22	TN specialist (author) online	Same TN specialist (author), unclear if blinded	Self-administered online questionnaire + expert entered electronic data into the ANN	607 facial pain patients	Expert clinicians in the field
Babiloni et al.	2017	US	Modified S-LANSS for intraoral pain	Modified S-LANSS for intraoral pain (non-specified)	Sent to patient prior to clinical assessment and some performed after clinic	Two individual facial pain experts, unclear if blinded	Self-completed paper or electronic questionnaire (patient choice)	136 facial pain patients	International Classification of Headache Disorders (ICHD) + CD/TMD classification for TMD
Hapak et al.	1994	Canada	None specified	21 + digital pain scale and pain diagram based on McGill scale	Neurologist with expertise in facial pain (author)	One of four investigators (2 neurologists and 2 dentists), unclear if blinded	Self-administered paper questionnaire	92 new patients	International Headache Society (IHS) criteria 1988, and TMD diagnoses 1979

Table 3. Summary of diagnostic subjects included in each questionnaire in order to fulfil minimum diagnostic criteria for classical trigeminal neuralgia (features according to current international diagnostic criteria by ICOP-1 and ICHD-3).

Questionnaire and reference	Recurrent paroxysmal pain	Unilateral	Innocuous stimuli trigger pain	Exclusive to trigeminal nerve distributions	Pain character (electric shock-like, shooting, stabbing or sharp)	Exclusion of other common condition	Short duration of attack (up to 2 min)	Severe intensity
PD-Q Jafree et al.	Yes	No	Yes	Yes, but diagram is small and unclear for face	No	No	No	No
ANN Limonadi et al.	Yes	Yes	Yes	Yes	Yes	Yes, ENT or TMJ related conditions	Yes	No
ANN McCartney et al.	Yes	Yes	Yes	Yes	Yes	Yes, ENT or TMJ related conditions	Yes	No
Modified S-LANSS Babiloni et al.	Yes	No	Yes	Yes, but diagram is small and unclear for face	Yes	No	No	Yes
McGill-Based Hapak et al.	Yes	Yes	Yes	No	Yes	Yes, headache/migraine and TMJ related conditions	No (non-specific, excludes constant pattern)	Yes, via digital pain scale of current and worst severity

Quality assessment

An adapted version of the Quality of Diagnostic Accuracy Studies (QUADAS-2) was used to assess the risk of bias according to pre-specified domains as described and summarised in Table 1.

According to the pre-defined criteria used to judge the overall risk of bias, only one study was considered low risk,¹⁹ two as low/medium risk^{16,17} and two as medium¹³ or medium/high risk.¹⁸ For patient selection and index test, most studies were assessed as low risk of bias with exception of the modified S-LANSS study with high risk for both domains.¹⁸ Similarly, the reference standard was low risk for almost all studies, apart from the study by McCartney et al., which demonstrated high risk for this domain.¹³ In contrast, the flow and timing domain of the assessment was reported high risk for all studies, except the study by Limonadi et al., which presented low risk. These were due to unclear information whether all participants received the same

reference standard, had appropriate interval between tests and inclusion of all participants responses in the analysis. In addition, it is worth noting that the settings of each study were different, where three were conducted in orofacial pain units,^{16–18} whereas one in a neurosurgical setting.^{13,19}

Diagnostic accuracy

The diagnostic accuracy measures including sensitivity, specificity, positive and negative predictive value, and level of agreement are summarised in Table 4.

The sensitivity and specificity of diagnosing TN for all tools were reported within large groups of different conditions, making it difficult to discriminate the diagnostic potential for TN alone. The exception was the automated system (ANN) that focused on adjusting the results that differentiate the typical TN pattern (TN1) from the variant (TN2).¹³ This system was able to demonstrate the highest

Table 4. Summary of diagnostic accuracy measures, including sensitivity, specificity, positive and negative predictive values, and level of agreement.

	Level of agreement	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Jaffree et al.	Not assessed	-59.6% [cut-off 11.5] -28.9% [cut-off 19.5]	-56.9% [cut-off 11.5] -83.25% [cut-off 19.5]	-62.4% [cut-off 11.5, prevalence 45.4%] -58.9% [cut-off 19.5, prevalence 54.6%]	-53.5% [cut-off 11.5, prevalence 45.4%] -58.5% [cut-off 19.5, prevalence 54.6%]
Limonadi et al. ^a	Group 1: 95% Group 2: Not assessed	Group 2 (n = 48): -TN1 = 84% -TN2 = 50% -TNP = 33% -TDP = 100%	Group 2: -TN1 = 83% -TN2 = 95% -TNP = 85% -TDP = 100%	Not assessed	Not assessed
McCartney et al. ^a	87.1% (529/607)	-TN1 = 92.4% -TN2 = 62.5% -TNP = 86.7% -TDP/PHN/STN = 0% -NIN = 50% -AFP = 0% -TMJ/GPN = 0%	-TN1 = 87.8% -TN2 = 96.4% -TNP = 95.2% -TDP/PHN/STN = 100% -NIN/AFP/TMJ = 99% -GPN = 100%	Not assessed	Not assessed
Babiloni et al. ^b	Not assessed	PDAP + TN = 52% PDAP = 64%	PDAP + TN = 70% PDAP = 63%	PDAP + TN = 35% PDAP = 23%	PDAP + TN = 82% PDAP = 91%
Hapak et al. ^c	A = 58% B = 50% C = 41%	A = 78.7% B = 78.9% C = 37.5%	A = 81.5% B = 78.2% C = 97%	A = 88.7% B = 55.6% C = 60%	A = 68.8% B = 91.5% C = 92.8%

^aGroup 1 = 100 participants from first recruitment. Group 2 = 43 participants included in second recruitment. TN1 = trigeminal neuralgia with spontaneous onset of pain >50% episodic, TN2 = trigeminal neuralgia with spontaneous onset of pain >50% constant, TNP = trigeminal neuropathic pain, TDP = trigeminal deafferentation pain (intentional trigeminal injury), PHN = postherpetic neuralgia from trigeminal herpes zoster outbreak, STN = symptomatic trigeminal neuralgia (multiple sclerosis), NIN = nervus intermedius neuralgia (or geniculate neuralgia), GPN = glossopharyngeal neuralgia, TMJ = temporomandibular joint disorder, AFP = atypical facial pain.

^bPDAP = persistent dentoalveolar disorder, TN = trigeminal neuralgia.

^cGroup A = Musculoligamentous pain (TMJ pain and myofascial pain); Group B = Neurologically based pain (trigeminal neuralgia, atypical facial pain, cluster headache, muscle contraction headache, migraine headache); Group C = Dentoalveolar pain (tooth or periodontal pain).

sensitivity levels for TN, especially for the episodic pattern (TN1). This was even higher in the follow-up study, although this study only included pre-selected patients with known orofacial pain.

The sensitivity and specificity levels reported by Jaffree et al. were described according to cut-off scoring determined at 11.5 based on previous original publications:²⁰ a PD-Q score ≤ 12 indicates an unlikely neuropathic component, a score ≥ 19 indicates likely neuropathic component and PD-Q scores between 12 and 19 may indicate presence of neuropathic pain.¹³ The sensitivity and specificity levels were reported above 50% (59.6% and 56.9%, respectively) when the cut-off scoring was determined at 11.5.¹² When the cut-off was increased to 19.5, the sensitivity was decreased to 28.9%, whereas specificity was increased (83.2%). The positive and negative predictive values in turn were similar regardless

of the cut-off and prevalence (62.4 and 58.9% positive, and 53.5 and 58.5% negative).

When analysing the diagnostic accuracy of other studies reported by Babiloni and Hapak et al.,^{14,15} there was no cut-off level utilised and the results showed improved sensitivity in diagnosing TN. These were however not specific only for TN as they included larger groups described as neurological or neuropathic conditions. Babiloni et al. who analysed the S-LANSS screening questionnaire aimed to compare 3 groups: a neuropathic group including TN plus PDAP and another neuropathic group including only PDAP, both against non-neuropathic groups including TMD and acute dental pain.¹⁴ The results were based on a threshold of $\geq 80\%$ (both specificity and sensitivity) to be considered target values for adequate accuracy in this study. All results were however below this

threshold, where sensitivity for neuropathic conditions (TN and PDAP) was 52% against 64% for only PDAP, whereas specificity was slightly higher for the mixed group (70%), and similar level for PDAP only (63%). The positive predictive value of both groups was low at 35 (TN + PDAP) and 23% (PDAP alone), whereas much higher negative predictive values were reported in 82 and 91% in that order.

In a similar way, Hapak et al. reported the results in groups of conditions, with three groups: musculoligamentous pain, neurologically based pain and dentoalveolar pain.¹⁵ These were grouped according to the number of inclusionary questions answered to classify the conditions with pre-defined number for each one (>50% threshold). The digital pain scale was used in addition if no diagnosis was concluded, and polar graphs were pre-defined to determine the classification in these situations. TN was included in the group of neurologically based pain, which demonstrated similar levels of sensitivity and specificity of 78.9% and 78.2%, respectively. It has also showed a high level of sensitivity and specificity for the musculoligamentous (78.7%). The sensitivity of dentoalveolar pain was in turn lower (37.5%) than other groups, but with the highest specificity (97%). The positive predictive value of group A was low compared to its negative predictive value (88.7% against 68.8%), while groups B and C showed higher negative prediction compared to positive one (55.6 and 60% against 91.5 and 92.8%, respectively). Although high accuracy levels were demonstrated for some groups, this subclassification interfered with the diagnostic accuracy for TN particularly and so could result in negative treatment outcomes.

The other two studies that analysed the automated tool (ANN) validation^{13,19} were the only studies that reported TN and other conditions as separate entities in the analysis. TN was subdivided into TN1 and TN2, depending on whether there was >50% episodic or constant pain, respectively. The initial study only evaluated group 2, which demonstrated a level of sensitivity >50% for TN1, TN2 and TDP, and only TNP had 33% of sensitivity. High levels of specificity were reported for all conditions, from 83 to 100%.

The follow-up study¹⁷ included many more diagnostic groups (PHN, NIN, STN, AFP, GPN and TMJ), and had higher sensitivity for TN1 and TNP over 80%, and lower levels for TN2 (62.5%), NIN (50%) and TDP, PHN, STN, AFP, TMJ and GPN all 0%. The specificity in turn was over 95% for all conditions, with exception of TN1 that resulted in 87.8%. There were no predictive values reported for the ANN studies.

Two questionnaire studies reported level of agreement.¹⁵⁻¹⁷ The study by Hapak et al.¹⁵ identified a

moderate kappa agreement between the screening questionnaire and gold standard clinical assessment ranging between 0.4 to 0.58. The neurological group that includes TN showed a moderate level of 50%, whereas the musculoskeletal group had the higher level of 58% and the dentoalveolar pain the lowest level of 40%.

The two studies that validated the ANN^{13,19} reported an overall moderate-to-high level of agreement when assessing the ANN and McGill-based questionnaires, ranging from 50 to 95% between the index test and clinical diagnoses.^{13,17,19} The initial study from Limonadi et al. demonstrated 95% of agreement between the diagnosis from the machine learning device against clinical examination from group 1, whereas group 2 was not reported. The follow-up study by McCartney et al. corrected the difference between the TN groups (TN1 and TN2), and in turn demonstrated lower overall agreement of 87.1%.¹⁷

Discussion

This systematic review demonstrates limited availability of validated screening questionnaires to diagnose TN and other orofacial pain conditions. Only four questionnaires have been utilised in validation studies specifically for diagnosing these conditions, including the PD-Q, ANN, modified S-LANSS and adapted McGill scale. In terms of diagnostic accuracy, the ANN tool has the highest levels of sensitivity and specificity for diagnosing episodic TN of above 80%^{13,19} and includes almost all relevant questions to diagnose TN from the current diagnostic criteria.^{1,4} It is also a tool that can be easily accessed online by patients and healthcare professionals. However, participants in the validation study were recruited from a single neurosurgical department and therefore not necessarily representative of the general population. Moreover, the tool does not provide more than one diagnosis, which can be common for chronic orofacial pain conditions. This could potentially lead to missing a diagnostic overlap that could require additional management and referral.

Given that the aim of this review was to address whether there are tools suitable for use in primary care by non-experts as an initial assessment, it was important for us to assess how patient selection was conducted in the studies. Most of the studies included in the review recruited patients with a known pain diagnosis pre-defined by a pain expert or from a pain-specialised centre rather than a primary care clinician from the general population. The only exception was the study by Jafree et al. where new patients referred by general practitioners or oral surgery department were invited to participate.¹⁶ It was unclear in the study by Hapak et al.

how patients had been referred to the pain clinic and whether there was a prior pain diagnosis.¹⁷

None of the questionnaires included all the current diagnostic criteria for classical TN from ICOP-1 and ICHD-3.^{1,4} These include recurrent paroxysmal attacks exclusive to the trigeminal nerve distributions, unilateral episodes, provoked by innocuous stimuli and pain character of electric shock-like, shooting, stabbing or sharp pain that lasts up to 2 min of severe intensity. The automated questionnaire (ANN) was the closest to addressing all the criteria, with exception of severity of pain.^{13,19} It does ask about pre-existing diagnosis of multiple sclerosis, history of trauma and whether the patient remembers the exact moment of onset of symptoms, which is very pertinent to TN. It also enquires into response to both medical and surgical therapies; however, these would not be relevant in a patient with new symptoms. In contrast, the study by Jafree et al. analysing PD-Q has the fewest diagnostic questions about TN as it was originally used to screen back pain. However, the pain timing pattern descriptor represented by simple graphs to choose from to discriminate constant, episodic pain and existence of pain-free periods is useful. Some questionnaires attempt to exclude other common pain conditions such as temporomandibular disorders, ENT related conditions and headache/migraine diagnoses.^{13,17,19}

There are examples of other good quality validated screening questionnaires for oral health conditions. However, these either targeted the classification of conditions rather than diagnostic screening or did not include TN as one of the analysed diagnoses, and were therefore excluded from this study. For example, Aggarwal et al. reported a questionnaire-based tool that would enable classification of idiopathic orofacial pain in the general population in the UK assessed by general medical practitioners in primary care.²¹ A large and robust patient recruitment was employed of over 4000 random subjects, and the classification questionnaire was developed for the reporting of unexplained symptoms, psychologic factors, disability related to orofacial pain and questions for classification of orofacial pain. The study revealed important pain descriptors to classify the diverse conditions that may be a useful tool to inform screening of orofacial pain in population-based studies where clinical examination is not possible.

Other studies looking specifically at temporomandibular disorders are useful examples of validating brief and effective screening questionnaires.^{22,23} These analysed a short and long version of a TMD pain screener to compare participants with orofacial pain with healthy controls, headache and odontalgia diagnoses. They aimed to analyse validity and reliability of a brief

questionnaire according to diagnostic criteria (Research Diagnostic Criteria for TMD, and ICHD-3 for headaches). Both versions had a sensitivity of 99% and a specificity of 97% for correct classification of the presence or absence of TMD in the larger study of over 500 participants.²³ In contrast, the similar study by van der Meer et al. with a smaller cohort (66 participants) revealed a sensitivity of 85% for the short version and specificity of 64%, whereas the full version resulted in 83% sensitivity and similar 82% specificity to differentiate pain-related TMD.²¹ It may be useful to use some of these questions to improve differentiation of TN from temporomandibular disorders especially if the latter is unilateral.

More recently, studies have focused on innovations that involve automated electronic tools to predict diagnosis, coded medical records. These also bring the potential to create algorithms of recommended treatments and are particularly interesting for services covering vast number of potential diagnoses with no definite diagnostic test such as chronic orofacial pain. Studies on diagnosing multiple orofacial pain diagnoses, temporomandibular disorders and headaches have been reported.²⁴⁻²⁸ Two studies included in this review about the automated diagnostic tool for orofacial pain (ANN) were promising in diagnosis of TN and other orofacial pain conditions.^{13,14} This tool is available online and is unsupervised but does contain a disclaimer that it needs validating by a professional. Moreover, it demonstrated zero sensitivity for diagnosing TMD likely due to the limited number of questions used in assessing this condition. It is unusual that a large cohort of over 600 participants had no one diagnosed with TMD. This is far common than all the other orofacial pain conditions, representing a third prevalence only after headache and back pain.^{29,30} It only allows for the diagnosis of one single condition, and patients may present an overlap of different diagnosis that require to be addressed separately.

This systematic review has some limitations. It is possible that restricting our search criteria to include TN excluded some potential questionnaires that could be helpful in screening orofacial pain conditions. Moreover, comparing the values of diagnostic accuracy was challenging due to the diversity of selected diagnostic groups that have not necessarily followed the current international diagnostic classification.

This study is the first to identify the limited evidence for validated and robust screening tools with reliable diagnostic accuracy to diagnose orofacial pain conditions, focussing on TN. This is particularly important as TN requires a completely different medical and surgical treatment compared to other chronic orofacial pain diagnoses. Our results showed that no available

tool is able to support the diagnosis of TN specifically, and there is still the need for a tool validated outside of a specialist facility and suitable for use in primary care. Considering these findings, we aim to co-produce a new tool to help primary care practitioners to identify TN more effectively and refer patients for treatment when appropriate. The availability of a reliable screening tool for TN could reduce misdiagnosis, inadequate referrals, and avoid unnecessary, irreversible treatments that can be detrimental to patient management and experience.

Author contributions

TT and RP analysed the databases, drafted the manuscript, and contributed to the final version. JZ reviewed and contributed to the final manuscript. All authors discussed the results and commented on the manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: JZ undertook this work at UCL/UCLHT who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centre funding scheme. RP received funding from the Trigeminal Neuralgia Association UK. THNT (Clinical Lecturer, CL-2019-18-009) is funded by Health Education England (HEE)/National Institute for Health Research (NIHR) for this research project. The views expressed in this publication are those of authors and not necessarily those of the NIHR, NHS or the UK Department of Health and Social Care.

ORCID iDs

THN Teshima  <https://orcid.org/0000-0003-2978-7081>

JM Zakrzewska  <https://orcid.org/0000-0001-7805-5851>

References

1. International classification of orofacial pain, 1st edition (ICOP). *Cephalalgia* 2020; 40: 129–221.
2. Olesen J. Headache classification committee of the International Headache Society (IHS) The international classification of headache disorders, 3rd edition. *Cephalalgia* 2018; 38: 1–211.
3. Cruccu G, Finnerup NB, Jensen TS, et al. Trigeminal neuralgia: new classification and diagnostic grading for practice and research. *Neurology* 2016; 87(2): 220–228.
4. Olesen J, Bes A, Kunkel R, et al. The international classification of headache disorders, 3rd edition (beta version). *Cephalalgia* 2013; 33(9): 629–808.
5. Benoliel R, Svensson P, Evers S, et al. The IASP classification of chronic pain for ICD-11: chronic secondary visceral pain. *Pain* 2019; 160(1): 69–76.
6. Zakrzewska JM, Wu J, Mon-Williams M, et al. Evaluating the impact of trigeminal neuralgia. *Pain* 2017; 158(6): 1166–1174.
7. von Eckardstein KL, Keil M and Rohde V. Unnecessary dental procedures as a consequence of trigeminal neuralgia. *Neurosurg Rev* 2015; 38(2): 355–360.
8. Drangsholt M and Truelove EL. Trigeminal neuralgia mistaken as temporomandibular disorder. *J Evid Based Dent Pract* 2001; 1(1): 41–50.
9. Slettebø H. Is this really trigeminal neuralgia? Diagnostic re-evaluation of patients referred for neurosurgery. *Scand J Pain* 2021; 21(4): 788–793.
10. Singh H, Schiff GD, Graber ML, et al. The global burden of diagnostic errors in primary care. *BMJ Qual Saf* 2017; 26: 484–494.
11. Cheraghi-Sohi S, Holland F, Reeves D, et al. The incidence of diagnostic errors in UK primary care and implications for health care, research, and medical education: a retrospective record analysis of missed diagnostic opportunities. *Br J Gen Pract* 2018; 68(suppl 1): b18X696857.
12. Viana M, Tassorelli C, Allena M, et al. Diagnostic and therapeutic errors in trigeminal autonomic cephalalgias and hemicrania continua: a systematic review. *J Headache Pain* 2013; 14(1): 1–11.
13. McCartney S, Weltin M and Burchiel KJ. Use of an artificial neural network for diagnosis of facial pain syndromes: an update. *Stereotact Funct Neurosurg* 2014; 92(1): 44–52.
14. Classification Committee of the International Headache Society. Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Headache classification committee of the international headache society. *Cephalalgia* 1988; 8(Suppl 7): 1–96.
15. Whiting PF, Rutjes AWS, Westwood ME, et al. Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529.
16. Jafree DJ, Zakrzewska JM, Bhatia S, et al. Accuracy of the painDETECT screening questionnaire for detection of neuropathic components in hospital-based patients with orofacial pain: a prospective cohort study. *J Headache Pain* 2018; 19(1): 103.
17. Hapak L, Gordon A, Locker D, et al. Differentiation between musculoligamentous, dentoalveolar, and neurologically based craniofacial pain with a diagnostic questionnaire. *J Orofac Pain* 1994; 8(4): 357–368.

18. Babiloni AH, Nixdorf DR, Law AS, et al. Initial accuracy assessment of the modified S-LANSS for the detection of neuropathic orofacial pain conditions. *Quintessence Int (Berl)* 2017; 48(5): 419–429.
19. Limonadi FM, McCartney S and Burchiel KJ. Design of an artificial neural network for diagnosis of facial pain syndromes. *Stereotact Funct Neurosurg* 2006; 84(5–6): 212–220.
20. Freynhagen R, Baron R, Gockel U, et al. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. *Curr Med Res Opin* 2006; 22(10): 1911–1920.
21. Aggarwal VR, McBeth J, Lunt M, et al. Development and validation of classification criteria for idiopathic orofacial pain for use in population-based studies. *J Orofac Pain* 2007; 21(3): 203–215.
22. van der Meer H, Bakke M, Schytz H, et al. Validation of the temporomandibular disorder pain screener in a specialized headache center. *J Oral Facial Pain Headache* 2021; 35(2): 150–156.
23. Gonzalez Y, Schiffman E, Gordon S, et al. Development of a brief and effective temporomandibular disorder pain screening questionnaire: reliability and validity. *J Am Dent Assoc* 2014; 140(12): 1485–1493.
24. Monreal APV, Veas N and Clark G. An artificially intelligent (or algorithm-enhanced) electronic medical record in orofacial pain. *Jpn Dental Sci Rev* 2021; 57: 242–249.
25. Bianchi J, Ruellas A, Prieto JC, et al. Decision support systems in temporomandibular joint osteoarthritis: a review of data science and artificial intelligence applications. *Semin Orthod* 2021; 27(2): 78–86.
26. Potter R, Probyn K, Bernstein C, et al. Diagnostic and classification tools for chronic headache disorders: a systematic review. *Cephalalgia* 2019; 39(6): 761–784.
27. Kwon J, Lee H, Cho S, et al. Machine learning-based automated classification of headache disorders using patient-reported questionnaires. *Scientific Rep* 2020; 10: 1–8.
28. Nocera L, Vistoso A, Yoshida Y, et al. Building an automated orofacial pain, headache and temporomandibular disorder diagnosis system. *AMIA Annu Symp Proc* 2020; 2020: 943–952.
29. Dworkin SF. The OPPERA study: act one. *J Pain* 2011; 12(11): T1–T3.
30. Slade GD, Fillingim RB, Sanders AE, et al. Summary of findings from the OPPERA prospective cohort study of incidence of first-onset temporomandibular disorders: implications and future directions. *J Pain: Official J Am Pain Soc* 2013; 14(12): T116–T124.