DR PASWACH WIRIYAKIJJA (Orcid ID : 0000-0002-2340-517X)

DR STEFANO FEDELE (Orcid ID : 0000-0001-9006-9412)

Article type : Original Manuscript

Title:

Patient-reported outcome measures in recurrent aphthous stomatitis: A critical assessment of quality properties

Running Title:

Patient reported outcome measures in RAS

Keywords:

Recurrent aphthous stomatitis, Patient Reported Outcome Measures, Psychometric properties, Interpretability

Paswach Wiriyakijja[1, 2*] Stefano Fedele[1, 3] Stephen Porter[1] Valeria Mercadante[1] Richeal Ni Riordain[1]

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/odi.12726

[1] UCL Eastman Dental Institute London, UK

[2] Department of Oral Medicine, Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand

[3] NIHR University College London Hospitals Biomedical Research Centre, London, United Kingdom

*Corresponding author.

E-mail address: paswach.w@gmail.com

OBJECTIVES: To analyse the range of existing patient-reported outcome measures (PROMs) used in studies of recurrent aphthous stomatitis (RAS) and to evaluate their quality properties via the assessment of psychometric properties and interpretability.

MATERIALS AND METHODS: Electronic databases were searched to identify relevant publications related to PROMs used in RAS. Publications were selected based on predefined criteria. All identified PROMs were then classified by measuring concepts and assessed for instrument characteristics and evidence for quality properties for RAS patients.

RESULTS: Twenty-eight PROMs were used in studies of RAS patients. Instruments focused upon oral symptoms (n=4), psychosocial status (n=15) and quality of life (n=9). Five PROMs (Oral Health related Quality of Life-UK, Chronic Oral Mucosal Disease Questionnaire, Oral Health Impact Profile-14, Medical Outcome Study Short Form-36, Mumcu's Composite Index) were found to have some evidence of psychometric performance. No PROM showed evidence for interpretability of their scores in RAS patients.

CONCLUSION: There was a wide range of PROMs used in clinical studies of RAS. The majority of these PROMs lack evidence of measurement properties and interpretability for RAS patients. Further studies are required to confirm whether these instruments are suitable and useful for this patient group.

INTRODUCTION

The past two decades have witnessed an increasing emphasis on measuring disease and treatment outcomes from the patient's perspective, leading to a steady rise in the development of patient-reported outcome measures (PROMs) as validated instruments to capture important outcomes directly from a patient (Black and Jenkinson, 2009; Devlin and Appleby, 2010). The aim of a PROM is to quantify patients' subjective perception of the disease and treatment impact on their day-to-day lives in a standardized way (Smith *et al*, 2005). Routine collection of PROM data in clinical settings has been proven to help inform clinical decision-making and enhance clinician-patient communication. In addition, a PROM can be a potential determinant that reflects the quality of healthcare service (Chen *et al*, 2013). From the perspective of clinical research, a critical step in the clinical trial design is to select a well-designed PROM with sufficient evidence of its fundamental quality properties including the three psychometric or measurement properties (validity, reliability, responsiveness) and interpretability to ensure that the instrument is appropriate and useful for a specific patient population (Mokkink *et al*, 2010). A psychometrically sound PROM is an instrument that is valid (able to measure what it is intended to measure), reliable (able to produce consistent scores in different occasions) and responsive (able to detect change over time if change does exists). Apart from psychometric performance of an instrument, scores or outcomes of PROMs should also be interpretable or have clinical meanings that are easily

understood by both patients and clinicians (Mokkink *et al*, 2010).

Little is known about the use of PROM in recurrent aphthous stomatitis (RAS), a very common ulcerative condition that is known to cause significant pain and discomfort of the oral mucosa (Akintoye and Greenberg, 2014). Its precise aetiopathogenesis remains unclear but is most likely multifactorial (Slebioda *et al*, 2014). RAS is characterized by recurrent eruptions of painful solitary or multiple small well-delineated round or ovoid ulcers with a yellowish or greyish centre and surrounding erythematous halo. The ulceration arises spontaneously at intervals of days (in mild cases) to months in otherwise well persons, gives rise to no systemic manifestations such as fever and heals spontaneously (Scully and Porter, 2008). Based upon its clinical presentation, RAS is classified as minor (MiRAS), major (MaRAS) and herpetiform (HU) types. Of all its clinical variants, MiRAS is the most prevalent form and is associated with relatively small ulcers (less than 1 cm) that are self-remitting and usually resolve within 7 to 14 days. Lesions of RAS usually first appear during childhood and adolescence and can hamper a patient's normal activities including food and fluid intake, speech and oral hygiene care, and this may consequently lead to psychosocial distress and impaired quality of life (QoL) (Jurge *et al*, 2006). Given the lack of definitive aetiological factors, management of RAS is usually symptomatic, and is aimed at alleviating patient's oral symptoms as well as improving patient's psychosocial status and quality of life (Baccaglini *et al*, 2011). Therefore, a patient-centred instrument that is able to capture the oral symptoms, psychosocial status and quality of life in patients with RAS is of great importance for accurate disease assessment and management.

While a clinician-centred RAS-specific clinical scoring system is available (Tappuni *et al*, 2013), very few studies have focussed upon the use and the quality properties of PROMs in patients with RAS. Three narrative reviews have reported on the use of PROMs in patients with a range of oral mucosal diseases (Ni Riordain and

McCreary, 2010; Ni Riordain *et al*, 2015; Wiriyakijja *et al*, 2017). There remains, however, no critical assessment of the quality properties of PROMs in patient with RAS. The aims of the present study are to 1) analyse the range of existing PROMs used for the measurement of oral symptoms, psychosocial status, and quality of life in patients with RAS, and 2) critically assess their psychometric properties and interpretability.

MATERIALS AND METHODS

A comprehensive review of English language articles in the literature were undertaken with the aim to identify all PROM-related clinical studies of defined population of participants with RAS.

Search strategies

A series of structured literature searches were performed on three medical databases including the MEDLINE (through PubMed), EMBASE and Web of Science Citation Index to retrieve all relevant clinical studies from the published literature from 1990 until June 2017 due to a substantial rise in the development and psychometric validation of PROM since 1990 (Garratt et al, 2002). The pre-defined search terms used for this review were comprised of disease keywords ('recurrent aphthous stomatitis' OR 'recurrent oral ulcers') combined with AND to the following keywords for each concept domain.

1. oral symptoms: 'pain' OR 'discomfort' OR 'symptom*'

2. psychosocial status: ‘psych*’ OR ‘anxiety’ OR ‘depress*’ OR ‘stress’ OR ‘mood’ OR ‘emotion*’ OR ‘social’
3. quality of life: ‘quality of life’ OR ‘oral health related quality of life’

Although this review focuses on RAS the term ‘recurrent oral ulcers’ was included in the search strategies to ensure a more extensive review of the oral ulceration literature.

Criteria for considering studies for this review

Study types and subjects

English-language, peered-review original articles involving the development, testing of psychometric properties (validity, reliability and responsiveness), documentation of interpretability and/or use of at least one validated PROM for the measurement of oral symptoms, psychosocial status and quality of life in participants with RAS were included. Clinical studies using PROMs as a screening instrument rather than for measuring outcomes, clinical studies using ad hoc instrument (instrument developed without psychometric testing), review articles, letters, commentaries, editorials or abstracts were excluded.

Study subjects with a confirmed diagnosis of RAS, diagnosed based upon previous or current history of RAS or clinical presentation of RAS-like oral ulcerations without underlying systemic conditions associated with the presence of aphthous-like oral ulceration (ALU) were included. Multi-disease studies with results stratified for participants with RAS were also included. Participants with diagnosis of the following

conditions were excluded: Behcet's disease, Reiter's syndrome, Sweet syndrome, Ulcerative colitis, Crohn's disease, Celiac disease, auto-inflammatory syndromes, haematological abnormalities (severe anaemia, cyclic or chronic neutropenia), recurrent erythema multiforme or any viral infection.

Data extraction

All identified PROMs were categorized upon their underlying concepts into oral symptom-PROMs, psychosocial-PROMs and QoL-PROMs. The number of items, subscales or domains, rating scales and score types and ranges of each identified PROM were then reviewed. These identified PROMs were subsequently assessed for their quality properties for the application in RAS patients.

The evaluation of the quality properties of the identified PROMs included

1. Validity: defined as the degree to which a PROM measures the construct(s) it purports to measure. The assessment of validity includes
   - Content validity: the extent to which the content of a PROM adequately reflects the proposed construct to be measured.
   - Construct validity: the extent to which a PROM validly measures the 'construct' or the theoretical concept that it purports to measure.
   - Criterion validity: the extent to which the scores of a PROM adequately relate to another 'criterion' measure that is considered to be a 'gold standard' in the field of study.

2. Reliability: the degree to which the measurement is free from measurement error. The assessment of reliability includes
   - Test-retest reliability: the extent to which the same results are obtained on repeated measurement of the same PROM when no change in patient's status has occurred.
   - Internal consistency reliability: the degree of inter-relatedness between the items.
3. Responsiveness: the ability of a PROM to detect change over time in the construct measured.
4. Interpretability: the degree to which one can assign qualitative meaning to a PROM's quantitative scores or change in scores (Mokkink *et al*, 2010).

RESULTS

Search results

Database searches identified 3,169 potentially relevant publications, which included duplicates and spurious references. Following the selection criteria, 129 articles were ultimately included in this review (Figure 1). Overall, a total of 28 PROMs were identified for the assessment of patient reported outcomes in patients with RAS from 129 publications (detailed in Table 1).

Oral symptom-PROM

There were 4 PROMs used for the assessment of RAS symptoms in 100 clinical studies, which included three generic instruments (Visual Analog Scale (VAS), Numerical Rating Scale (NRS), Graded Chronic Pain Scale (GCPS)) and one disease-specific instrument (Mumcu's composite index). With respect to the type and severity of RAS being reported in 100 studies, 63 were MiRAS-specific studies, 2 included patients with MaRAS (one of these included HU patients) and 35 did not report RAS subtypes in their studies.

Regarding generic oral symptom-PROMs, the vast majority of RAS studies (86/100, 86%) used VAS while NRS were used in twelve studies (12%) and only one study (Sherman et al, 2007) used GCPS. There was a wide diversity in the use of word descriptors in the VAS and NRS among included studies including 'pain' (in 59 of 86 RAS studies using VAS (68.60%); 9 of 12 RAS studies using NRS (75%)), 'pain and discomfort' (in 6 of 86 RAS studies using VAS (6.98%)) and many others (Table 2). Out of the 86 RAS studies using the VAS, only 33 studies (38.37%) provided clear and accurate information, in the relevant material and methods section, regarding the use of the instrument and the measurement of results; 34 articles (39.53%) reported unclear or incorrect information while 19 articles (22.09%) did not provide any information.

Apart from generic instruments, there was one disease-specific instrument identified for the assessment of RAS-related pain and other different constructs: Mumcu's composite index (composite index: instrument generating single combined score of two or more individual components), which was developed for the assessment of the impact of oral ulcer activity in RAS and Behcet's disease (Mumcu *et al*, 2009) and was used for study outcome measurement in one RAS study (Soylu Özler et al,

2016). It consists of 3 different subscales including oral ulcer activity (as reflected by the presence or absence of oral ulcers in the previous month; 0 - 1 point), pain (measured by VAS; 0 – 5 points) and functional status (assessed the impacts of oral ulcers on taste, speaking, and eating/chewing/swallowing on a 5-point Likert-type scale; 0 - 4 points), with total score of 10. However, this composite index was validated for use only in Turkish population, without any evidence of translation or cross-cultural validation for other countries/languages.

Psychosocial-PROMs

A total of 15 PROMs have been used for the evaluation of psychosocial status in patients with RAS from 18 clinical studies. One study reported the inclusion of all three RAS clinical variants while the other studies did not report clinical subtypes of included participants. All of the PROMs are generic instruments (Table 3), which measure different psychological and emotional constructs including anxiety (11 studies), depression (7 studies), stress (6 studies), distress/psychological symptoms (3 studies), coping (1 study) and anger (1 study). The most frequently used psychosocial-PROMs in RAS were the Hospital Anxiety and Depression Scale (HADS; 5 studies), followed by the State-Trait Anxiety Inventory (STAI; 3 studies).

Quality of life-PROMs

A total of 9 QoL-PROMs were identified from 17 studies. Sixteen studies were non RAS subtype-specific whereas one was MiRAS-specific. Six of the instruments assessed oral health-related quality of life (OH-QoL) while three (SF-36, SF-12 and WHOQOL-BREF) examined general aspects of quality of life. Of the 6 OH-QoL-PROMs, one instrument was developed for the use in children aged 11-12 years old.

Table 4 provides characteristics of these instruments. The most frequently used QoL-PROMs in RAS population were the Oral Health Impact Profile-14 (OHIP-14; 9 studies), followed by the Oral Health-Related Quality of Life-UK (OHQOL-UK; 3 studies) and the Medical Outcome Study Short Form 36 Health Survey (SF-36; 2 studies).

Evidence for quality properties of identified PROMs

Of all identified PROMs, 5 PROMs including 4 QoL-PROMs (SF-36, OHIP-14, OHQOL-UK and COMDQ) and the Mumcu's composite index had undergone psychometric testing in RAS patients but only the COMDQ was found to have good psychometric evidence on all main psychometric properties (validity, reliability, responsiveness) in patients with RAS. Mumcu's composite index was examined for its validity and reliability in Turkish patients with RAS (Mumcu *et al*, 2009). Three other QoL-PROMs including SF-36, OHIP-14 and OHQOL-UK were investigated only for their internal consistency reliability in Turkish patients with RAS (Mumcu *et al*, 2006), with results showing high Cronbach's α coefficient (≥ 0.92) in all instruments. However, other psychometric properties including validity and responsiveness in these QoL-PROMs have yet been examined in a RAS population. Table 5 summarises the psychometric testing of the reviewed PROMs. Importantly, none of the PROMs used in RAS patients has evidence or documentation for interpretability of their scores in this patient population.

DISCUSSION

Recurrent aphthous stomatitis is a common oral ulcerative condition associated with pain and other oral symptoms, which can have a significant negative impact upon normal oral functioning, psychosocial functioning and OH-QoL in affected individuals (Llewellyn and Warnakulasuriya, 2003; Tabolli *et al*, 2009). Therefore, in the clinical evaluation of patients with RAS, it is of paramount importance to assess the effects of this oral condition and its treatment from the perspective of the patient. The selection of an appropriate instrument to measure subjective RAS-related patient-reported outcomes requires careful consideration of the psychometric properties as well as the interpretability of the instruments. The present study reviewed the use of PROMs in clinical studies of patient with RAS as well as published evidence supporting the psychometric properties and interpretability specifically for this patient population.

In the present study three generic oral symptom-PROMs including the VAS, NRS and GCPS were used in patients with RAS, with VAS being the most frequently used instrument. Nevertheless, further investigation into the use of these instruments in the RAS literature revealed inconsistencies in reporting the type of oral symptoms measured by VAS, as shown by a wide spectrum of different word descriptors for VAS including “pain”, “burning sensation”, “discomfort”, “irritation” and many others (Table 2). This study heterogeneity makes it difficult to pool VAS data for the comparison between studies and meta-analysis. Also, we observed that only 38% of studies of RAS provided clear instruction regarding the use of VAS in the methodology section, whilst the information in the remaining studies on VAS was either absent, unclear or inaccurate; for instance, 27 studies using VAS (31.40%) stated that “patients rated their symptoms on a scale from 0 to 10” , which appear to reflect NRS rather than VAS, and 7 studies using VAS (8.14%) used different range of numerical scale including 1-10 (5 studies), 0-5 (1 study) and 0-4 (1 study). Whilst

both VAS and NRS have been widely used in clinical studies of RAS, neither has been investigated for psychometric performance specifically for patients with RAS. We would suggest that further testing of the psychometric properties of VAS and NRS in the RAS population is recommended.

Regarding the assessment of psychosocial status, anxiety and depression were the most frequently evaluated concepts in RAS patients, and the HADS, STAI and BDI were the most commonly used psychosocial-PROMs in the RAS literature. All three measures have been validated in a general population (Spielberger and Gorsuch, 1983; Beck *et al*, 1988; Snaith, 2003); nevertheless, all of them lack psychometric evidence in the RAS population. There were a few instruments used for assessing other psychosocial constructs in individuals with RAS, and again there was no published evidence of their psychometric testing or interpretability in this patient population. Overall, the present findings raise concerns as to whether these instruments are indeed relevant to RAS patients and if they are suitable for assessing the psychosocial status of individuals with RAS.

Evaluation of QoL in patients with RAS is also crucial. QoL-PROMs used in clinical studies of RAS population can be classified into OH-QoL-PROMs and generic QoL-PROMs. The present study identified 6 OH-QoL PROMs, but only three have had their psychometric properties tested in the RAS population: the OHIP-14, OHQOL-UK and COMDQ. OHIP-14 is the most frequently used PROMs for the assessment of QoL in the RAS literature. This PROM was initially developed for use in older Australian adults and is a shortened version of the original OHIP-49. It contains 14 items with a subset of 2 questions for each of the 7 domains of OH-QoL, based upon Locker's conceptual framework of oral health (Locker, 1988; Slade and Spencer,

1994). The development of OHQOL-UK was based on an adult UK population's perceptions of how oral health affects quality of life (McGrath and Bedi, 2003). Therefore both OHIP-14 and OHQOL-UK were developed without the input from patients with RAS and therefore may not be able to capture all relevant aspects associated with the disease and related treatment. COMDQ is an oral medicine-specific PROM developed for the assessment of quality of life in patients with chronic oral mucosal disease including RAS (Ni Riordain *et al*, 2011). It is the only validated QoL-PROM with input from patients with RAS during its development process. In addition, COMDQ has the highest number of psychometric studies for patients with RAS compared to the other OH-QoL PROMs. Regarding the measurement of general aspect of quality of life, only two PROMs have been used in studies of RAS patients including SF-36, SF-12 and QHOQOL-BREF, with only the SF-36 having some psychometric evidence tested for use in patients with RAS.

We also identified one oral ulcer composite index, which aims to determine the impact of oral ulcer activity in patients with RAS and BD (Mumcu *et al*, 2009). This index, however, has not been widely adopted for use in clinical research of RAS, and apart from Turkish original language, there is no evidence of translation nor of cultural validation of this index. In addition, the rationale behind weights of three subscale scores to generate total composite index score appears to be unclear. Further validation studies for this composite index are recommended.

We found that there are no studies reporting the interpretability of any PROM used in

clinical research of patients with RAS, and this casts doubts on the clinical meaningfulness of the PROM results in clinical studies of patients with RAS. Interpretability refers to what the scores or change scores mean in clinical context, which facilitates better understanding of PROM results (Mokkink *et al*, 2010). The numerical scores produced from PROMs should be easily translated into clinically meaningful information, relevant to patients, clinicians and researchers. An interpretability parameter such as the minimal important difference (MID), the smallest magnitude of change in PROM scores that is important or meaningful to patients, can therefore facilitate clinical interpretation of these scores. Although MID for improvement of OHIP-14 in Behcet's disease was previously determined in a Turkish study by Hayran *et al* (2009), this parameter has yet to be determined in group of patients with RAS. There is thus a need for further studies determining interpretability of PROMs in patients with RAS.

The goal for RAS management is usually to minimize oral symptoms and improve patient's oral functioning and quality of life. Although different groups of medications are available for RAS patients, there is currently no robust evidence supporting the efficacy of any of these medications, and future larger randomized placebo-controlled trials (RCTs) are required. These RCTs will require the careful selection of validated outcome measures, both clinician-centred and PROMs. Although the present study identified that some PROMs showed appropriate psychometric properties for use in clinical studies of RAS, there is currently a lack of uniformity regarding the choice of outcome measures including both PROMs and clinical scoring systems across the RAS literature (Brocklehurst *et al*, 2012). The comprehensive development of a core outcome set (COS) for clinical trials of RAS has been initiated and presented by Taylor *et al* at the recent European Association of Oral Medicine (EAOM) conference in 2016. The methodology incorporated both patients with RAS (n=6) and experts

(n=70), leading to a COS of 13 core outcomes for interventional studies in RAS. This COS includes all 6 key outcomes highlighted by patients namely, ulcer size, ulcer duration, frequency of ulcer attack, number of ulcers, pain and diet. The use of COS for RAS will improve the quality and uniformity of data in future clinical trials, allowing comparison between treatments and data pooling in systematic reviews and meta-analyses.

In conclusion, there was a wide diversity of PROMs used in clinical studies of RAS, which include instruments for oral symptoms, psychosocial status and QoL. The majority of these PROMs lack evidence of measurement properties and interpretability for RAS patients. Further studies are required to confirm whether these instruments are suitable and useful for this patient group.

ACKNOWLEDGEMENT

This project received no financial support.

CONFLICT OF INTEREST

None to declare.

REFERENCES

Akintoye, S., & Greenberg, M. (2014). Recurrent Aphthous Stomatitis. *Dental Clinics of North America, 58*, 281-297. doi: 10.1016/j.cden.2013.12.002

Baccaglini, L., Lalla, R. V., Bruce, A. J., Sartori-Valinotti, J. C., Latortue, M. C., Carrozzo, M., & Roger III, R. S. (2011). Urban legends: recurrent aphthous stomatitis. *Oral Diseases*, *17*, 755-770. doi: 10.1111/j.1601-0825.2011.01840.x.

Beck, A., Steer, R., & Carbin, M. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*, 77-100. doi: https://doi.org/10.1016/0272-7358(88)90050-5

Black, N., & Jenkinson, C. (2009). Measuring patients' experiences and outcomes. *BMJ, 339*, b2495. doi: https://doi.org/10.1136/bmj.b2495

Brocklehurst, P., Tickle, M., Glenny, A. M., Lewis, M. A., Pemberton, M. N.,

Taylor, J., …Yates, J.M. (2012). Systemic interventions for recurrent aphthous

stomatitis (mouth ulcers). *The Cochrane Database of Systematic Reviews*, *12*. doi: 10.1002/14651858.CD005411.pub2

Chen, J., Ou, L., Hollis, S. J. (2013). A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Services Research*, *13*, 211. doi: 10.1186/1472-6963-13-211.

Devlin, N., & Appleby, J. (2010). *Getting the most out of PROMs.* London: King's Fund.

Garratt, A. (2002). Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ*, *324*, 1417. doi: https://doi.org/10.1136/bmj.324.7351.1417

Hayran, O., Mumcu, G., Inanc, N., Ergun, T., & Direskeneli, H. (2009). Assessment of minimal clinically important improvement by using Oral Health Impact Profile-14 in Behçet's disease. *Clinical and Experimental Rheumatology*, *27*, S79-84.

Jurge, S., Kuffer, R., Scully, C. & Porter, S. (2006). Number VI Recurrent aphthous stomatitis. *Oral Diseases*, *12*, 1-21. doi: 10.1111/j.1601-0825.2005.01143.x

Li, M., & He, S. (2013). Reliability and validity of the Chinese version of the Chronic Oral Mucosal Diseases Questionnaire. *Journal of Oral Pathology and Medicine*, *42*, 194-199. doi: 10.1111/j.1600-0714.2012.01157.x.

Llewellyn, C., & Warnakulasuriya, S. (2003). The impact of stomatological disease on oral health-related quality of life. *European Journal of Oral Sciences*, *111*, 297-304. doi: 10.1034/j.1600-0722.2003.00057.x

Locker, D. (1988). Measuring oral health: a conceptual framework. *Community Dental Health*, *5*, 3-18.

McGrath, C., & Bedi, R. (2001). An evaluation of a new measure of oral health related quality of life--OHQoL-UK(W). *Community Dent Health, 18*, 138-43.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., … de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*, 737-45. doi: 10.1016/j.jclinepi.2010.02.006.

Mumcu, G., Hayran, O., Ozalp, D., Inanc, N., Yavuz, S., Ergun, T., & Direskeneli, H. (2007). The assessment of oral health-related quality of life by factor analysis in patients with Behcet's disease and recurrent aphthous stomatitis. *Journal of Oral Pathology and Medicine*, *36*, 147-152. doi: 10.1111/j.1600-0714.2007.00514.x

Mumcu, G., Inanc, N., Ergun, T., Ikiz, K., Gunes, M., Islek, U., ... Direskeneli, H. (2006). Oral health related quality of life is affected by disease activity in Behcet's disease. *Oral Diseases*, *12*, 145-151. doi: 10.1111/j.1601-0825.2005.01173.x

Mumcu, G., Sur, H., Inanc, N., Karacayli, U., Cimilli, H., Sisman, N., … Direskeneli, H. (2009). A composite index for determining the impact of oral ulcer activity in Behçet's disease and recurrent aphthous stomatitis. *Journal of Oral Pathology and Medicine*, *38*, 785-791. doi: 10.1111/j.1600-0714.2009.00803.x.

Ni Riordain, R., Hodgson, T., Porter, S. R., & Fedele, S. (2016). Validity and reliability of the Chronic Oral Mucosal Diseases Questionnaire in a UK population. *Journal of Oral Pathology and Medicine*, *45*, 613-616. doi: 10.1111/jop.12425.

Ni Riordain, R., & McCreary, C. (2012). Further reliability and responsiveness of the Chronic Oral Mucosal Diseases Questionnaire. *Oral Diseases*, *18*, 60-66. doi: 10.1111/j.1601-0825.2011.01844.x.

Ni Riordain, R., & McCreary, C. (2010). The use of quality of life measures in oral medicine: a review of the literature. *Oral Diseases*, *16*, 419-430. doi: 10.1111/j.1601-0825.2009.01647.x.

Ni Riordain, R., & McCreary, C. (2011). Validity and reliability of a newly developed quality of life questionnaire for patients with chronic oral mucosal diseases. *Journal of Oral Pathology and Medicine*, *40*, 604-609. doi: 10.1111/j.1600-0714.2011.01021.x.

Ni Riordain, R., Meaney, S., & McCreary, C. (2011). A patient-centered approach to developing a quality-of-life questionnaire for chronic oral mucosal diseases. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, And Endodontology*, *111*, 578-586.e2. doi: 10.1016/j.tripleo.2010.12.011.

Ni Riordain, R., Shirlaw, P., Alajbeg, I., Al Zamel, G. Y., Fung, P. L., Yuan, A. D., …

Brennan, M. T. (2015). World Workshop on Oral Medicine VI: Patient-reported outcome measures and oral mucosal disease: current status and future direction. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology*, *120*, 152-60. doi: 10.1016/j.oooo.2015.01.023.

Scully, C., & Porter, S. R. (2008). Oral mucosal disease: Recurrent aphthous stomatitis. *British Journal of Oral and Maxillofacial Surgery*, *46*, 198-206. Doi: https://doi.org/10.1016/j.bjoms.2007.07.201

Sherman, J. J., Barach, R., Whitcomb, K. K., Haley, J., & Martin, M. D. (2007). Pain and pain-related interference associated with recurrent aphthous ulcers. *Journal of Orofacial Pain*, *21*, 99–106.

Slade, G. D., & Spencer, A. J. (1994). Development and evaluation of the Oral Health Impact Profile. *Community Dental Health*, *11*, 3-11.

Ślebioda, Z., Szponar, E., & Kowalska, A. (2013). Etiopathogenesis of Recurrent Aphthous Stomatitis and the Role of Immunologic Aspects: Literature Review. *Archivum Immunologiae et Therapiae Experimentalis*, *62*, 205-215. doi: 10.1007/s00005-013-0261-y

Smith, S. C., Cano, S., Lamping, D. L., Staniszewska, S.,Browne, J., Lewsey, J., … Black, N. (2005). *Patient reported outcome measures (PROMs) for routine use in treatment centres: recommendations based on a review of the scientific evidence*. London: Health Services Research Unit, London School of Hygiene and Tropical Medicine.

Snaith, R. P. (2003). The Hospital Anxiety And Depression Scale. *Health Quality of Life Outcome*, *1*, 29. doi: 10.1186/1477-7525-1-29

Spielberger, C., & Gorsuch, R. (1983). *Manual for the State-trait anxiety inventory (form Y) ("self-evaluation questionnaire")* (1st ed.). Palo Alto, CA: Consulting

Psychologists Press.

Soylu Özler, G., & Akoğlu, E. (2016). The Relationship Between Neutrophil to Lymphocyte Ratio and Recurrent Aphthous Stomatitis. *Journal of Clinical and Analytical Medicine,* 7, 152-154. doi: 10.4328/JCAM.2619

Tabolli, S., Bergamo, F., Alessandroni, L., Di Pietro, C., Sampogna, F., & Abeni, D. (2009). Quality of Life and Psychological Problems of Patients with Oral Mucosal Disease in Dermatological Practice. *Dermatology*, *218*, 314-320. doi: 10.1159/000196973.

Tappuni, A., Kovacevic, T., Shirlaw, P., & Challacombe, S. (2013). Clinical assessment of disease severity in recurrent aphthous stomatitis. *Journal of Oral Pathology and Medicine*, *42*, 635-641. doi: 10.1111/jop.12059.

Taylor, J., Glenny, A. M., Walsh, T., Brocklehurst, P., & Pemberton, M. N. (2016, September). *Core Outcome Set for Recurrent Aphthous Stomatitis – a collaborative approach.* Poster session presented at the 13th Biennial Congress European Association of Oral Medicine, Turin, Italy.

Wiriyakijja, P., Fedele, S., Porter, S. R., Mercadante, V., & Ni Riordain, R. (2017). Patient-reported outcome measures in oral lichen planus: A comprehensive review of the literature with focus on psychometric properties and interpretability. *Journal of Oral Pathology and Medicine.* Advance online publication. doi: 10.1111/jop.12604

Table 1 Types (by concepts measured), acronyms and frequency of use of PROMs in clinical studies of patients with RAS

Instrument type and name	frequency of use
PROMs assessing oral symptoms	
Symptoms	
Visual Analog Scale (VAS)	86
Numerical Rating Scale (NRS)	12
Graded Chronic Pain Scale (GCPS)	1
Pain, functional status and oral ulcer activity	
Mumcu's composite index (CI)	1
PROMs assessing psychosocial status	
Anxiety (only)	
State-Trait Anxiety Inventory (STAI)	3
Beck Anxiety Inventory (BAI)	1
Self-Rating Anxiety Scale (SAS)	1
Depression (only)	
Beck Depression Inventory (BDI)	2
Quick Inventory of Depressive Symptomatology (QIDS-SR16)	1
Stress (only)	
Recent Life Change Questionnaire (RLCQ)	2
Lipp's Inventory of Stress Symptoms of Adults (LISS)	1
Symptoms of Stress List (SSL)	1
Test of Recent Experience (TRE)	1
Anxiety and depression	
Hospital Anxiety and Depression Scale (HADS)	5
Anxiety and anger	
State-Trait Personality Inventory (STPI)	1
Distress/psychological symptoms	
General Health Questionnaire-12 (GHQ-12)	1
General Health Questionnaire-28 (GHQ-28)	1
Symptom Checklist (SCL-90)	1
Coping	
Ways of Coping Questionnaire (WCQ)	1

PROMs assessing quality of life	
Oral health related quality of life	
Oral Health Impact Profile-14 (OHIP-14)	9
Oral Health-Related Quality of Life-UK (OHQOL-UK)	3
Oral Health Impact Profile-49 (OHIP-49)	1
Oral Impacts on Daily Performances (OIDP)	1
Oral health related quality of life specific to chronic oral mucosal diseases	
Chronic Oral Mucosal Disease Questionnaire (COMDQ)	1
Oral health related quality of life specific to children	
Child Oral Impacts on Daily Performances (Child-OIDP)	1
General health related quality of life	
Medical Outcome Study Short Form 36 Health Survey (SF-36)	3
Medical Outcome Study Short Form 12 Health Survey (SF-12)	1
The World Health Organization Quality of Life-BREF (WHOQOL-BREF)	1

Table 2 Word descriptions used in VAS and NRS in the studies assessing oral symptoms of RAS

Word descriptors	**frequency**
Visual analog scale (VAS)	
pain	59
pain and discomfort	6
stimulated/challenged/evoked pain [A]	5
contact pain [B]	4
idiopathic/non-contact/spontaneous pain	4
discomfort	3
pain and burning sensation	3
pain before meals	3
soreness	3
pain before bedtime	2

discomfort during chewing	1
discomfort during speaking	1
functional complication [C]	1
irritation	1
oral symptoms [D]	1
pain and irritation	1
tingling	1
Numerical rating scale (NRS)	
pain	9
pain during tooth brushing	1
spontaneous pain	1
stimulated pain [E]	1

A: stimulated pain (N = 5; orange juice-stimulated pain [3], citric acid-stimulated pain [1], hot saline-stimulated pain [1])

B: contact pain (N = 4; pain immediately after laser treatment [2], pain after irritating ulcer with periodontal probe [2])

C: functional complication = interruption of aphthous ulcers with normal daily activities i.e. speaking, chewing and brushing

D: oral symptoms i.e. oral pain, difficulty to eat and difficulty to sleep

E: pain after swabbing ulcer with a saturation of sodium chloride and distilled water

Table 3 Characteristics of PROMs assessing psychosocial status in clinical studies of patients with RAS

Name	Items (N)	Concept	Subscale (N items)	Rating scale	Score types and range		
					Subscales	Total	Others
BAI	21	Anxiety	Anxiety (21)	4-point scale (0-1-2-3)		0-63	
BDI, BDI-II	21	Depression	Depression (21)	4-point scale (0-1-2-3)		0-63	
GHQ-12	12	Distress	Distress (12)	4-point scale (0-0-1-1 or 0-1-2-3)		0-12 0-36	
GHQ-28	28	Distress	Somatic symptoms (7); Anxiety and insomnia (7); Social dysfunction (7); Severe depression (7)	4-point scale (0-0-1-1 or 0-1-2-3)	0-7 0-21	0-28 0-84	
HADS	14	Anxiety, depression	Anxiety (HADS-A) (7); Depression (HADS-D) (7)	4-point scale (0-1-2-3)	0-21		
LISS	56	Stress	Phase: Alert (Q1) (16); Resistance amd Near-exhaustion (Q2) (16); Exhaustion (Q3) (24)	2-point scale (0-1)	0-15 (Q1, 2) 0-23 (Q3)		
QIDS-SR16	16	Depression	Depression (16)	4-point scale (0-1-2-3)		0-27	
RLCQ	68	Stress	Stressful life events (91)	2-point scale (0-life change units)		✓ (total life change units)	No of events
SAS	20	Anxiety	Anxiety (20)	4-point scale (1-2-3-4)		20-80	

SCL-90	90	Psychological symptoms	Somatisation (SOM); Obsessive-compulsive behavior (O-C); Interpersonal sensitivity (I-S); Depression (DEP); Anxiety (ANX); Hostility (HOS); Phobic anxiety (PHOB); Paranoid ideation (PAR); Psychoticism (PSY)	5-point scale (0-1-2-3-4)	✓		GSI* PST* PSDI*
SSL	59	Stress	Stress (59)	4-point scale (0-1-2-3)		0-177	
STAI	40	Anxiety	State anxiety (STAI-S) (20); Trait anxiety (STAI-T) (20)	4-point scale (1-2-3-4)	20-80		
STPI	80	Anxiety, anger	State anxiety (10); Trait anxiety (10); State anger (10); Trait anger (10); State curiosity (10); Trait curiosity (10); State depression (10); Trait depression (10)	4-point scale (1-2-3-4)	10-40		
TRE	42	Stress	Vital events (42)	2-point scale (0-life change units)		0-600	
WCQ	66	Coping	Confrontive coping (6); Distancing (6); Self-controlling (7); Seeking social support (6); Accepting responsibility (4); Escape-Avoidance (8); Planful problem solving (6); Positive reappraisal (7)	4-point scale (0-1-2-3)	✓		

**Abbreviation:* GSI = Global Severity Index (mean of all subscale scores); PST = Positive Symptom Total (number of items with score > 0); PSDI = Positive Symptom Distress Index (the sum of all item

values divided by PST)

Table 4 Characteristics of PROMs assessing quality of life in clinical studies of patients with RAS

Name	Items (N)	Concept	Subscale (N items)	Rating scale	Score types and range		
					Subscales	Total	Others
Child-OIDP	8	OHQOL specific to children	Eating (1); Speaking (1); Cleaning teeth (1); Smiling (1); Emotional stability (1); Relaxing (1); Doing schoolwork (1); Social contact (1)	4-point scale on frequency and severity (0-1-2-3)		Total (0-100) (each item score: frequency x severity x 100/72)	
COMDQ	26	OH-QOL specific to COMD	Pain & function limitation (PF) (9); Medication & treatment (MT) (6); Social & emotional (SE) (7); Patient support (PS) (4)	5-point scale (0-1-2-3-4)	0-36 for PF 0-24 for MT 0-28 for SE 0-16 for PS	0-104	
OHIP-14	14	OH-QOL	Functional limitation (FL) (2); Physical pain (PhyP) (2); Psychological discomfort (PsyD) (2); Physical disability (PhyDis) (2); Psychological disability (PsyDis) (2); Social disability (SDis) (2); Handicap (H) (2)	5-point scale (0-1-2-3-4)		0-56 (Severity)	Extent*
OHIP-49	49	OH-QOL	Functional limitation (FL) (9); Physical pain (PhyP) (9); Psychological discomfort (PsyD) (5); Physical disability (PhyDis) (9); Psychological disability (PsyDis) (6); Social disability (SDis) (5); Handicap (H) (6)	5-point scale (0-1-2-3-4)	0-36 for FL, PhyP, PhyDis 0-24 for PsyDis, H 0-20 for PsyD, SDis	0-196	
OHQOL-UK	16	OH-QOL	Physical effects/impacts (Phy-E/I) (6); Social effects/impacts (S-E/I) (5); Psychological effects/impacts (Psy-E/I) (5)	5-point scale (1-2-3-4-5 for effects and 0-1-2-3-4 for impacts)	6-54 for Phy-E/I 5-45 for S-E/I, Psy-E/I	16-144	
OIDP	8	OH-QOL	Eating (1); Speaking and pronouncing clearly (1); Cleaning teeth (1); Sleeping and relaxing (1); Smiling without embarrassment (1); Maintaining emotional state (1); Enjoying contact with other people (1); Carrying out major school work (1)	6-point scale on frequency and severity		Total (0-100) (each item score:	

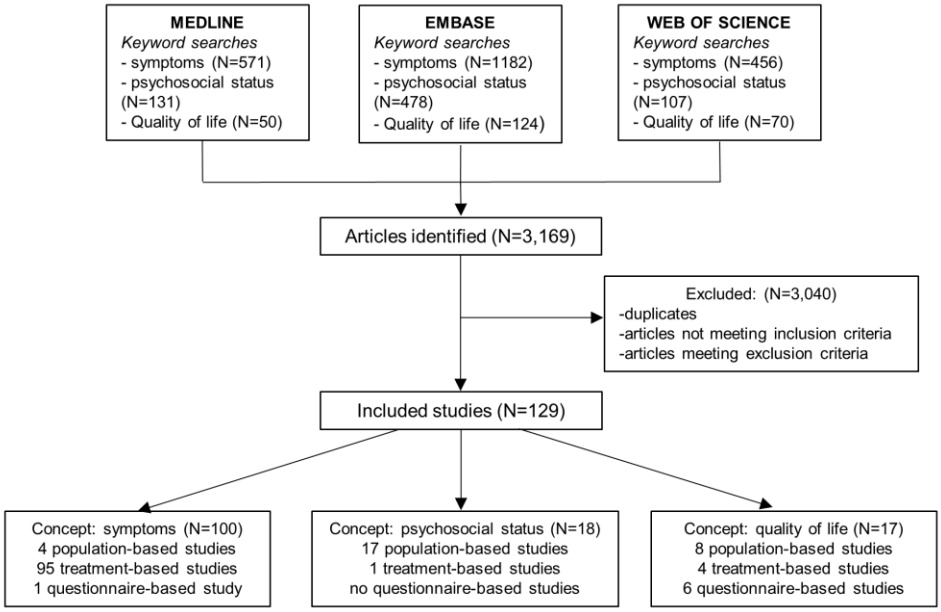
				(0-1-2-3-4-5)		frequency x severity x 100/150)	
SF-12	12	GH-QOL	Physical functioning (PF) (2); Role physical (RP) (2); Bodily pain (BP) (1); General health (GH) (1); Vitality (VT) (1); Social functioning (SF) (1); Role emotional (RE) (2); Mental health (MH) (2)	2- to 6-point scale			PCS-12 MCS-12
SF-36	36	GH-QOL	Physical functioning (PF) (10); Role physical (RP) (4); Bodily pain (BP) (2); General health (GH) (5); Vitality (VT) (5); Social functioning (SF) (2); Role emotional (RE) (3); Mental health (MH) (5); Health transition (HT) (1)	2- to 6-point scale	0-100 (transformed from raw score)	0-100 (transformed from raw score)	PCS* MCS*
WHOQOL-BREF	26	GH-QOL	Overall quality of life and general health (Ov) (2); Physical health (Ph) (7); Psychological (Ps) (6); Social relationships (So) (3); Environment (En) (8)	5-point scale (1-2-3-4-5)	7-35 for Ph 6-30 for Ps 3-15 for So 8-40 for En		

*Note: Extent = N of items reported fairly often (3)/very often (4); GH-QOL = general health related quality of life; OH-QOL = oral health related quality of life; PCS = Physical Component Summary; MSC = Mental Component Summary

Table 5 Summary of psychometric properties of identified PROMs in clinical studies of patients with RAS

Authors	PROMs	Questionnaire language/country	Main Methods of Evaluation	No of patients	Major reported outcomes
Mumcu *et al*, 2006	OHIP-14	Turkish/Turkey	Internal consistency	24	Cronbach's α = 0.95
	OHQOL-UK	Turkish/Turkey	Internal consistency	24	Cronbach's α = 0.97
	SF-36	Turkish/Turkey	Internal consistency	24	Cronbach's α = 0.92
Mumcu *et al*, 2007	OHIP-14	Turkish/Turkey	Structural validity	28	FA revealed three subscales and explained 66.49% of overall variance in patients with active oral ulcers
Mumcu *et al*, 2009	Mumcu's composite index (CI)	Turkish/Turkey	Convergent validity (correlation with VAS for pain, number of oral ulcers, frequency of relapses); Discriminant validity between patients with oral ulcers (RAS and BD) and patients with dental infections; Internal consistency	31	Moderate to good convergent validity of total CI score with VAS for pain (r = 0.90), number of oral ulcers (r = 0.77) and frequency of relapse (r = 0.51); No CI score in patients with dental infections; Cronbach's α for functional disability score = 0.75
Ni Riordain and McCreary, 2011	COMDQ	English/Ireland	Convergent validity (correlation with VAS for pain and OHIP-14), Discriminant validity between patients with and without COMD, Internal consistency	12	Good convergent validity with VAS for pain (r = 0.883) and OHIP-14 (r = 0.819); Significant difference in COMDQ scores between patients with and without COMD; Cronbach's α = 0.929

Ni Riordain and McCreary, 2012	COMDQ	English/Ireland	Test-retest reliability, Responsiveness to change	?	Good test-retest reliability (ICC = 0.81); COMDQ is responsive to changes in the patient's overall conditions
Li and He, 2013	COMDQ	Chinese/China	Structural validity; Internal consistency; Test-retest reliability	84	EFA extracted four factors (consistent with original english version) and all items demonstrated adequate factor loadings; Cronbach's α = 0.894; ICC of total COMDQ scores = 0.83
Ni Riordain *et al*, 2016	COMDQ	English/UK	Convergent validity (correlation with VAS and OHIP-14), Internal consistency	42	Moderate to good convergent validity with VAS and OHIP-14; Cronbach's α = 0.93



Note: questionnaire-based study related to the development and/or psychometric testing of PROMs for the use in RAS patients