DEVELOPMENT AND VALIDATION OF THE MALAY DIAGNOSTIC CRITERIA FOR TEMPOROMANDIBULAR DISORDERS (DC/TMD)

MOHD KHAIRUL ANWAR BIN MOHD TAHIR

FACULTY OF DENTISTRY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

DEVELOPMENT AND VALIDATION OF THE MALAY DIAGNOSTIC CRITERIA FOR TEMPOROMANDIBULAR DISORDERS (DC/TMD)

MOHD KHAIRUL ANWAR BIN MOHD TAHIR

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF CLINICAL DENTISTRY (ORAL MAXILLOFACIAL SURGERY)

FACULTY OF DENTISTRY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate : Mohd Khairul Anwar Bin Mohd Tahir

I.C/Passport No

Matric No : DGJ150002

Name of Degree : Master of Clinical Dentistry
(Oral & Maxillofacial Surgery)

Title of Research Report : Development and Validation of the Malay Diagnostic Criteria for Temporomandibular Disorders (DC/TMD)

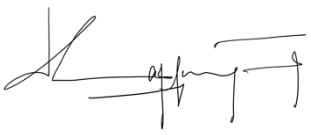
Field of Study : Oral & Maxillofacial Surgery

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright of this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work, I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 13th August 2019

Subscribed and solemnly declared before,

Witness's Signature  Date: 13th August 2019

Name: Dr Kathreena Kadir

Designation

DEVELOPMENT AND VALIDATION OF THE MALAY DIAGNOSTIC CRITERIA FOR TEMPOROMANDIBULAR DISORDERS (DC/TMD)

ABSTRACT

Introduction. The health status questionnaires measuring Oral Health Related Quality of Life has increased tremendously over the years. The Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) was developed in 2014 and has been extensively used worldwide for Temporomandibular Disorders research and clinical settings. However, for this tool to be used in Malaysian culture, the tool has to be modified from the source version through a formal cross-cultural adaptation process. This study aims to develop and validate a Malay version of DC/TMD tool through established guidelines so that it can be used as a TMD diagnostic protocol among the Malaysian population who are majority Malay language speakers. The objectives of this study are to translate the English version of DC/TMD into Malay language and to assess the psychometric properties of the Malay DC/TMD. **Methods.** The DC/TMD was translated into the Malay language using a forward-backward method. The finalized translated version was given to a total of 252 subject which consist of 165 non TMD individuals and 87 TMD patients. The psychometric properties of 2 domains which were the Malay Graded Chronic Pain Scale (GCPS) and the Malay Jaw Functional Limitation Scale (JFLS) were assessed through internal consistency reliability, test-retest reliability and validity tests. The literature on this topic was also reviewed. **Results.** The Cronbach's alpha for both domains demonstrated good internal consistency of the items (Malay GCPS = 0.95, Malay JFLS = 0.97) while the intraclass correlation coefficient for test-retest reliability was 0.98 for Malay GCPS and 0.99 for Malay JFLS. There were 3 types of validity tests that were conducted namely, concurrent validity, construct validity and discriminant validity. All validity association

were found statistically significant with good positive correlation except for the association between Malay GCPS and 3 items of Malay version of Oral Health Impact Profile (M-OHIP) which were poor negative correlations. Eleven out of fourteen constructed validity hypotheses were confirmed with 79% of the results corresponding to the itemized hypotheses. A specific OHIP tool for TMD (OHIP TMD) is recommended for a more sensitive result to be used in the future. **Conclusion.** The Malay version of DC/TMD is empirically shown to be valid and reliable for assessing TMD among Malay language speakers in Malaysia. A further study is recommended to collect more data for better supportive evidence.

Key words: DC/TMD, cross-cultural adaptation, Malay language, Temporomandibular disorders, Reliability and Validity, Translation

PENTERJEMAHAN DAN PENGESAHAN KRITERIA DIAGNOSTIK UNTUK GANGGUAN TEMPOROMADIBULAR (KD/GTM) KE BAHASA MELAYU

ABSTRAK

Pengenalan. Soal selidik status kesihatan yang mengukur kesihatan oral telah meningkat dengan pesat sejak beberapa tahun ini. Kriteria Diagnostik untuk Gangguan Temporomandibular (KD/GTM) telah dibentuk pada tahun 2014 dan telah digunakan secara meluas di seluruh dunia untuk tujuan penyelidikan gangguan temporomandibular (GTM). Walau bagaimanapun, untuk soal selidik ini digunakan dalam budaya Malaysia, ianya ini perlu diubahsuai dari bahasa Inggeris ke Bahasa Melayu melalui proses komunikasi antara budaya yang formal. Kajian ini bertujuan untuk membentuk dan mengesahkan soal selidik KD/GTM versi Bahasa Melayu melalui garis panduan yang ditetapkan supaya ianya boleh digunakan sebagai protokol diagnosis GTM di kalangan penduduk Malaysia yang majoritinya bertutur dalam Bahasa Melayu. Objektif kajian ini adalah untuk menterjemah KD/GTM versi Bahasa Inggeris ke Bahasa Melayu dan menilai sifat psikometriknya. **Kaedah.** KD/GTM diterjemahkan ke dalam Bahasa Melayu menggunakan kaedah yang ditetapkan. Soal selidik yang telah diterjemah diberikan kepada sejumlah 252 subjek yang terdiri daripada 165 subjek bukan-GTM dan 87 pesakit GTM. Ciri-ciri psikometrik daripada 2 domain iaitu *Graded Chronic Pain Scale (GCPS)* dan *Jaw Functional Limitation Scale (JFLS)* dinilai melalui *'internal consistency reliability', ujian 'test-retest reliability'* dan ujian kesahihan. **Keputusan.** Nilai '*Cronbach Alpha'* bagi kedua-dua domain menunjukkan konsistensi dalaman antara item yang baik (GCPS = 0.95, JFLS = 0.97) manakala' *intraclass correlation coefficient'* adalah 0.98 untuk GCPS dan 0.99 untuk JFLS. Terdapat 3 jenis ujian kesahihan yang dijalankan iaitu Kesahihan Bersama *(Concurrent validity),*

Kesahihan Pembuktian (*Construct validity)* dan Kesahihan Diskriminasi (*Discriminant Validity*). Kesemua ujian kesahihan adalah signifikan secara statistik dengan korelasi positif yang baik kecuali korelasi antara *GCPS* dan 3 item Profil Impak Kesihatan Mulut versi Melayu *(M-OHIP)* yang mempunyai korelasi negatif. Sebelas daripada empat belas hipotesis yang dibina adalah sahih dimana 79% daripada hasilnya bersesuaian dengan hipotesis yang telah dibentuk. Profil Impak Kesihatan Mulut versi Melayu khas utk GTM *(M-OHIP TMD)* disyorkan untuk menghasilkan data yang lebih sensitif . **Kesimpulannya.** KD/GTM versi bahasa Melayu adalah instrumentasi yang sah untuk menilai GTM di kalangan penduduk Malaysia. Kajian lanjut disyorkan untuk mengumpul lebih banyak maklumat untuk data sokongan yang lebih baik.

Kata kunci: KD/GTM, Komunikasi antara Budaya, Bahasa Melayu, Gangguan Tempromandibular, Kebolehpercayaan dan Kesahihan, Penterjemahan

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful

Alhamdulillah, I am truly grateful to be given the opportunity to complete my research project entitled "The Development of Malay Diagnostic Criteria for Temporomandibular Disorders (DC/TMD)". This report was a big leap in my learning curve as a post graduate student and was certainly not done singlehandedly.

First and foremost, I would like to express my gratitude to Allah, God Almighty for giving me the strength and determination in completing this challenging task. Indeed, there have been difficult moments in the completion of this project that would not have been overcome without divine intervention.

This research topic would not have been conceived if not for the brainchild of my main supervisor, Dr Kathreena Kadir for giving me the idea and encouraging me to pursue with this topic. She has been extremely helpful in my data collection and the write up of this report. Her guidance and support have proved invaluable.

My sincere thanks must also go to my co-supervisors, Associate Prof. Dr Zamros Yuzadi bin Mohd Yusof from the Department of Community Oral Health & Clinical Prevention as well as Professor Dr Adrian Yap from the National University of Singapore. They generously gave their time to offer me valuable comments and constructive criticism which helped me to develop a broader perspective to my research project.

My gratitude also goes out to Ms. Najihah, a supportive Statistician at our Dental Research Management Centre (DRMC) who has been very supportive and dedicative in conducting statistics clinics and consultation sessions for all of us, the post graduate students.

I cannot forget my batchmates for going through the hard times, cheering me on and celebrating each accomplishment together; Dr Muzaffar Apipi, Dr Umar Kamali, Dr Sujesh Sree and also Dr Noor Azwani Mat Nawi. Thanks for the friendship and memories.

To my family, particularly my mother and siblings; thank you for your love, support, and unwavering belief in me. Without you, I would not be the person I am today. I could not have asked for a better support system than this, I am truly blessed.

Not to forget, I would like to thank my lovely and beautiful wife, Dr Aswani binti Abdul Rashid for her love and constant support, for all the late nights and early mornings, and for keeping me sane over the past few months. Thank you for being my best friend and a big hug and love for your unfailing emotional support. To my son, Khalaf Anwar, thank you for allowing me to spend most of my time on this research and I promise, Baba will spend more time with you after this.

Last but not least, this research is dedicated to my late father who has been my constant source of inspiration. Abah, this is for you.

A big thank you once again to everyone involved in this project!

Mohd Khairul Anwar Bin Mohd Tahir

June, 2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

TMD : Temporomandibular Disorders

TMJ : Temporomandibular Joint

ICHD : The International Classification of Headache Diseases

WHO : World Health Organization

DC/TMD : Diagnostic Criteria for Temporomandibular Disorders

RDC/TMD : Research Diagnostic Criteria for Temporomandibular Disorders

HRQoL : Health-Related Quality of Life

OHRQoL : Oral Health Related Quality of Life

IADR : International Association for Dental Research

GCPS : Graded Chronic Pain Scale

JFLS : Jaw Functional Limitation Scale

PHQ : Patient Health Questionnaire

GAD : Generalized Anxiety Disorders

OBC : Oral Behavior Checklist

JDL : Jaw Disability List

ICC : Interclass correlation coefficient

M-BPI : Malay version of Brief Pain Inventory

M-OHIP14 : Malay-Oral Health Impact Profile 14

CI : Confidence Interval

NOHSA : National Oral Health Survey of Adults

M-OHIP TMD: Malay-Oral Health Impact Profile for Temporomandibular Disorders

μ : Spearman's rank correlation coefficient

α : Cronbach's alpha

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Introduction

Temporomandibular Disorders (TMD) are believed to be multifactorial. In general, TMD is a disorder of the facial masticatory system associated with the temporomandibular joint (TMJ), masticatory muscles and its surrounding hard and soft tissue components. TMD consists of joint and muscle symptoms that can be certified by limited mouth opening, pain at the masticatory muscles and temporomandibular joint (TMJ), myofascial pain, headache, jaw deviations, jaw dysfunctions such as clicking, locking and dislocation, as well as other functional restrictions in the muscle and TMJ area (Jung, Kim, Park, & Jung, 2015). The International Classification of Headache Diseases (ICHD, 2004) of World Health Organization (WHO), described headache or facial pain attributed to TMD as a recurrent pain on the face or head region with at least one of the following; 1) pain that is triggered by jaw movement and/or during chewing hard or tough food, 2) abnormal jaw opening or reduced jaw mobility, 3) unilateral or bilateral noise and/or tenderness of the TMJ(s).

Based on the National Institute of Dental and Craniofacial Research, 2013, TMD affected 5%-12% of the population and has been determined as the main source of orofacial pain with musculoskeletal disorders. Nowadays, TMD is said to cause significant public health problem and this disorder is listed as the second most common musculoskeletal disorders that causes pain and disability after chronic low back pain (National Institute of Dental and Craniofacial Research, 2013).

In general, patients will only seek treatment for this disorder when the TMD symptoms are chronic and persistent. This orofacial pain might be affecting their quality of life and causing disturbances to their psychosocial status (Schiffman et al.,2014).

In order to manage TMD effectively, it is believed that a proper diagnosis and the exact TMD components of the disorders are well determined. A global development of patient care and medical research require a standard and universal tool to assess the TMD status. As TMD seems to be a complex issue, a simple, decisive and valid clinical operational tool is needed to provide physical diagnosis in both clinical and research settings. Furthermore, a bio behavioral evaluation of pain-related nature and assessment of psychosocial status of a TMD patient is important as part of the diagnostic process.

Based on the physical and psychosocial status in diagnosing TMD, an updated dual-axis tool was developed known as Diagnostic Criteria for Temporomandibular Disorders (DC/TMD). DC/TMD is an international and standardized tool that has been used worldwide in order to diagnose TMD that can be used in both clinical and research settings. Information gathered by DC/TMD is based on the evidence-based criteria for the clinicians to use in clerking their TMD patients. Thus, this will ease the communication among the clinicians and researchers regarding consultations, referrals, as well as diagnosis and prognosis of TMD (Feinstein,1967).

DC/TMD was developed and published in 2014 based on the original tool which is the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) and has been extensively used worldwide as a diagnostic protocol for TMD research. DC/TMD was first published in 1992 (Dworkin & LeResch,1992) and the initial plan of the development of RDC/TMD was expected to be a first step towards a better classification for TMD (Schiffman et al., 2014). However, over the years, the RDC/TMD consortium network stated the need for further investigation and improvements of the RDC/TMD version. The consortium board recommended that RDC/TMD needed further assessment of the accuracy and efficiency of the tool as well as to improve its validity and clinical utility (Schiffman et al., 2014). With all the hard

work and multiple consensus, an improved and validated tool of RDC/TMD was developed and finally published in 2014 known as Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) which is more evidence based.

DC/TMD consist of Axis I and Axis II. Axis I is the physical and clinical assessment while Axis II consists of the investigation of the psychosocial status. The new DC/TMD is more relevant for use in both clinical and research settings. It is more complete, straightforward and can act as a simple screening tool which allows for diagnosing of patients with a range of simple to complex TMD signs and symptoms (Schiffman et al.,2014).

The original version of DC/TMD has been translated into various languages since 2016. Among the earliest translated versions were Dutch, Spanish and Swedish. These versions were developed in contemplation of testing the competency of the English version as well as for training purposes, calibration and accuracy of field trials. Other languages to follow were Chinese, Finnish, Hungarian, Italian, Japanese and Turkish. To date, there are more than 25 on-going works that are taking place to translate the original version of DC/TMD into other languages such as Hindi, Thai, Indonesian, Arabic, Urdu, Korean, French, Hebrew, Nepali, Romanian, Polish and many more.

Guillemin et al. in 1993, stated that clinicians and researchers who are planning to diagnose and to do research of a disease without an acceptable and native health-related quality of life (HRQOL) tools that are not in their own language have two options, which are;

(1) to construct a new tool, or

(2) to adapt a tool that has been validated in a different language.

This process is known as cross-cultural adaptation procedures which consist of translation and validation works. Hence, this study was conducted to develop and validate a Malay version of DC/TMD that can be used among the Malaysian population. With the development of the Malay DC/TMD, we are hoping that it can be used for research and clinical diagnostic purposes to diagnose and to provide better management of TMD in the Malaysian posterity.

1.2 Aim

To develop and validate a Malay version of DC/TMD tools through a formal cross-cultural adaptation process.

1.3 Objectives

The specific objectives of this study are:

i. To translate DC/TMD English version into the Malay language.

ii. To assess the psychometric properties of the Malay DC/TMD.

CHAPTER 2: LITERATURE REVIEW

2.1 Temporomandibular Disorders

2.1.1 Definition

Temporomandibular disorders (TMD) are significant problems. TMD is a broad and complete terminology that is related to numerous clinical issues that refer to the temporomandibular joint (TMJ), supporting masticatory muscles and their encircling structures (Kim, Kim, Im, & Yun, 2012). Due to its involvement of the craniofacial region and its relationship to the mandible, Okeson (1997) in his publication stated that TMD can also be called as craniomandibular disorders (CMD). TMD definition can also be incorporated with myofascial pain dysfunction syndrome, synovial capsule-ligament disorder, internal disc derangement and osteoarthritis (Wright,2005). While most literatures described TMD as a dental related problem, U J Yap et al. (2003) linked TMD as an accumulation of medical and dental illnesses that affect the TMJ and/or the muscles of mastication as well as the adjacent tissue components.

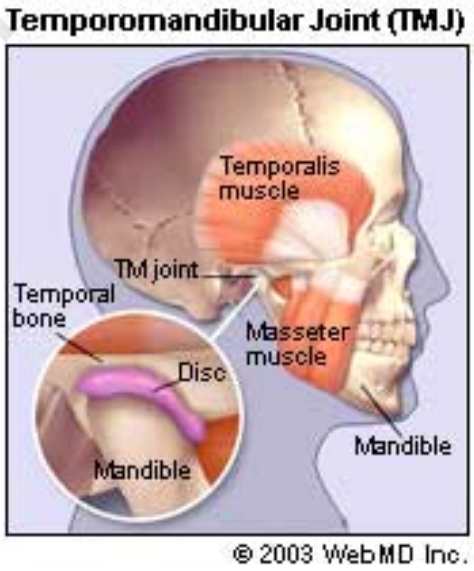


Figure 2.1: TMD is related to the TMJ, the disc, supporting muscles and its' surrounding structures. (Image adapted from web MD Website, 2003)

2.1.2 Etiology

There are many theories associated with the etiology and exact causes for various signs & symptoms of TMD. However, TMD is believed to be multifactorial (Abrahamsson et al., 2013). Some subtypes of TMD usually will be influenced by a specific etiology, but as a general condition, TMD has no familiar etiology or biological evidence (U J Yap et al., 2003).

Multifactorial components of TMD agents are believed to be in nature with biological, behavioral, environmental, social, emotional as well as intellectual and psychosocial factors (U J Yap, K Chua, Dworkin, Tan, & Tan, 2002). All these components can act in combination or as a single factor. Parafunctional habits such as bruxism, tooth clenching and gum chewing activity are also thought to be the causes of TMD.

(Jang, Kwon, Lee, Bae, & Kim, 2016) reported that TMD is also closely related to an individual's life style and job scope. Principally, individuals with a heavy usage of the jaw or mouth and persistent strain to the head and neck muscles (i.e. musicians) might have increased risk of developing TMD symptoms.

2.1.3 Signs and Symptoms of TMD

The prevalence of TMD varies in different populations across the globe. Previous studies have mentioned that the prevalence of sign and symptoms of TMD varies from 14% to 72.2% (Schellhas, Pollei, & Wilkes, 1993). In Asian countries, it is reported that the female to male ratio distribution of TMD was 3.1:1 (U J Yap et al., 2003). Based on this study, majority of the TMD patients were female aged between 25 to 44 years old.

It is believed that the disorder of this joint and its surrounding structures, can significantly affect the sufferer with clinical signs and symptoms such as severe orofacial and neck pain, headache, sleep disorders, depression as well as limitation of jaw functional activities like eating, chewing, biting, kissing and speaking (Manfredini et al., 2011). Most frequently, the type of sleep disturbances in TMD patients are sleep bruxism, insomnia and sleep apnea (Smith et al., 2009). Other than that, symptoms of TMD also includes reduced mandibular movement, joint sounds (i.e. clicking, crepitus), generalized myofascial pain and functional limitations or jaw deviation during mouth opening (Wadhwa&Kapila,2008). TMD has various familiar symptoms. However, the most common symptoms that motivate patients to seek treatment is pain in the masticatory muscle and/or the TMD surrounding area (Lei, Liu, Yap, & Fu,2015).

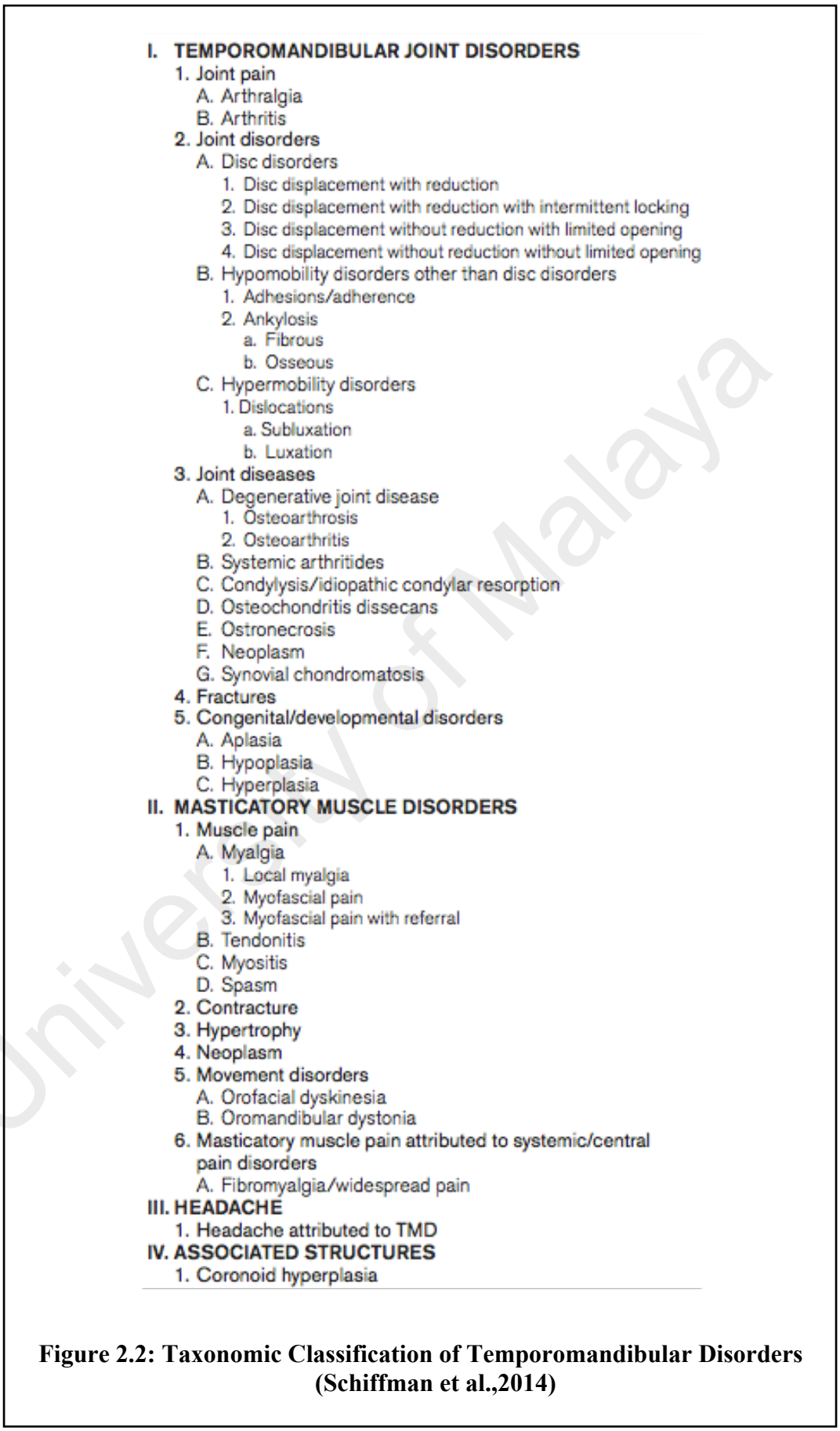
Besides physical signs and symptoms, patients with TMD also showed symptoms of psychosocial distress and multiple psychologic disturbances (UJ Yap et al.,2002). It is relatively recognized that distress symptoms such as anxiety and depression may influence, accelerate and preserve TMD signs and symptoms (UJ Yap et al., 2002). Multiple studies have showed that patients with painful TMD have greater levels of psychological and emotional distress (Filligim et al.,2012; Yap et al., 2002).

2.1.4 Diagnosis

Diagnosing TMD is quite a challenging task. The diagnosis of temporomandibular disorders is often achieved through clinical examinations with the aid of radiographic and imaging analysis (Jang et al., 2016). In addition to the clinical examinations and radiographic investigations, the functional assessment tool has also been used to diagnose TMD (Schiffman et al.,2014). Based on this study, a simple, comprehensive, good and accurate operational definition for the clinical history, examinations and imaging process are needed to provide physical diagnosis in both

clinical and research settings. Thus, taking all of these factors together, a new dual-axis tool was created and has been known as the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD).

DC/TMD was developed based on the original version of the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) which have been extensively used as a diagnostic protocol for TMD research since its publication in 1992 (Dworkin & LeResch, 1992). Based on DC/TMD **(Figure 2.2)**, the classification of TMD diagnosis can be classified into temporomandibular joint disorders, masticatory muscle disorders, headache and associated structures.



I. TEMPOROMANDIBULAR JOINT DISORDERS
1. Joint pain
A. Arthralgia
B. Arthritis
2. Joint disorders
A. Disc disorders
1. Disc displacement with reduction
2. Disc displacement with reduction with intermittent locking
3. Disc displacement without reduction with limited opening
4. Disc displacement without reduction without limited opening
B. Hypomobility disorders other than disc disorders
1. Adhesions/adherence
2. Ankylosis
a. Fibrous
b. Osseous
C. Hypermobility disorders
1. Dislocations
a. Subluxation
b. Luxation
3. Joint diseases
A. Degenerative joint disease
1. Osteoarthrosis
2. Osteoarthritis
B. Systemic arthritides
C. Condylysis/idiopathic condylar resorption
D. Osteochondritis dissecans
E. Ostronecrosis
F. Neoplasm
G. Synovial chondromatosis
4. Fractures
5. Congenital/developmental disorders
A. Aplasia
B. Hypoplasia
C. Hyperplasia
II. MASTICATORY MUSCLE DISORDERS
1. Muscle pain
A. Myalgia
1. Local myalgia
2. Myofascial pain
3. Myofascial pain with referral
B. Tendonitis
C. Myositis
D. Spasm
2. Contracture
3. Hypertrophy
4. Neoplasm
5. Movement disorders
A. Orofacial dyskinesia
B. Oromandibular dystonia
6. Masticatory muscle pain attributed to systemic/central pain disorders
A. Fibromyalgia/widespread pain
III. HEADACHE
1. Headache attributed to TMD
IV. ASSOCIATED STRUCTURES
1. Coronoid hyperplasia

Figure 2.2: Taxonomic Classification of Temporomandibular Disorders (Schiffman et al.,2014)

2.2 Diagnostic Criteria for Temporomandibular Disorders

2.2.1 Development of the DC/TMD

The Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) was developed from the Research on Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) which was published in 1992. The RDC/TMD was first developed with a thought of improving the TMD classification (Schiffman et al,2014). It has been used extensively as a diagnostic tool for TMD research since its publication (Dworkin & LeResche, 1992).

The original RDC/TMD consisted of Axis I and Axis II instruments. Axis I determine the physical diagnoses while Axis II instruments measure the psychosocial status and pain-related disability. After the usage for many years, the authors of RDC/TMD stated that the tool needed future research to revise its validity and clinical usage (Schiffman et al.,2014). Starting from 2001, multicentered validation studies of RDC/TMD with reference for standard examiners has been conducted which is known as the 'Validation Projects'. This project focused on the comprehensive assessment of the RDC/TMD reliability and validity.

In July 2008, a symposium was conducted at the International Association for Dental Research (IADR) Conference funded by the International RDC/TMD Consortium Network. A revised RDC/TMD diagnostic algorithm was presented to the international research community. An agreement was achieved after the presentation to hold a workshop for the development of new RDC/TMD.

Following that, in 2009, the International RDC/TMD Consortium Network once again conducted a conference to address 'The Validation Project's' proposal regarding

development of the new RDC/TMD known as DC/TMD that served for research and clinical purposes (Schiffman et al.,2014). The Validation Project's data and suggestions were finally published in 2010 (Look et al., 2010).

The proposal draft of the new DC/TMD was presented in July 2010 at another IADR conference. The draft was introduced to the international clinical and research community for appraisal and discussions. Further clarification of the new DC/TMD diagnoses was held in 2011. A field testing of the examiner blueprint for the Axis I assessment protocol and Axis II specifications were done from 2011-2012 in 4 countries, namely The United States of America, Sweden, Denmark as well as Germany (Schiffman et al.,2014).

Following the above move, the final estimation of reliability and validity of DC/TMD had taken place in 2013. Finally, in the year of 2014, a formal and validated DC/TMD was published. Since then, this tool has been used worldwide and multiple processes of translation and cultural adaptation were first done by the Brazilian, Dutch, Swedish and German groups (Ohrbach et al., 2013). To date, DC/TMD has been translated into various languages around the world. The major steps taken in development of the new DC/TMD is shown in **Figure 2.3.**

Year	Event
1992	**Publication of RDC/TMD** Expert-based classification of most common TMD derived from epidemiologic and clinical data Dual-axis system: Clinical conditions (Axis I) and pain-related disability and psychological status (Axis II)
2001–2008	**Validation Project** Multicenter study with reference standard examiners Comprehensive assessment of the reliability and validity of the RDC/TMD Establish need to revise RDC/TMD
2008	**Symposium at IADR‡ Conference (Toronto)** Revised RDC/TMD diagnostic algorithms presented to the international research community Published critique and recommendations to enhance use in research
2009	**International RDC/TMD Consensus Workshop at IADR Conference (Miami)** Input from the international dental and medical clinical and research community as well as from patient advocate perspective Published critique and recommendations to facilitate use in clinical and research settings
2010	**Publication of Major Findings by Validation Project** Revised RDC/TMD algorithms provided reliable and valid clinical criteria for pain-related TMD Demonstrated need for imaging for most TMJ disc displacements and degenerative joint disease Support for existing Axis II instruments Recommended development of DC/TMD with international input
2010	**Symposium at IADR Conference (Barcelona)** DC/TMD presented to the international clinical and research community Critique and comments on Axis I diagnostic algorithms for the most common TMD and Axis II assessment protocol
2011	**International RDC/TMD Consensus Workshop at IADR Conference (San Diego)** Refinement of Axis I diagnostic algorithms for common and less common TMD
2011–2012	**Field Trials of Axis I Examiner Specifications and Axis I & II Self-Report Instruments** Test sites: Buffalo (US), Minneapolis (US), Malmö (Sweden), Aarhus (Denmark), Heidelberg (Germany), and Stockholm (Sweden)
2012	**Finalization of DC/TMD at IADR Conference (Iguacu Falls)** Further input from members of national and international TMD pain organizations Review of the DC/TMD by the IADR 2009 conference participants
2013	**Final Estimates of Reliability and Validity for Axis I Diagnostic Criteria** Derived from the datasets of the Validation Project and TMJ Impact Project Finalization of DC/TMD

*Research Diagnostic Criteria for Temporomandibular Disorders.
†Diagnostic Criteria for Temporomandibular Disorders.
‡International Association for Dental Research.

Figure 2.3: Major steps from the original RDC/TMD (1992) to the new DC/TMD (2014) (Schiffman et al.,2014)

2.2.2 The new DC/TMD

2.2.2(a) The DC/TMD Axis I

Axis I of DC/TMD comprises of TMD pain screener, questionnaire on symptoms, demographics and a clinical examination form (International RDC/TMD Consortium et al., 2014). The differences of clinical form in DC/TMD is based on the tooth numbering system. The North American version uses the Universal system adopted by the American Dental Association while the International version uses the Federation Dentaire Internationale (FDI) tooth numbering system (International RDC/TMD Consortium et al., 2014). In any clinical setting, the usage of Axis I TMD Pain Screener is suggested for all patients (Gonzalez et al., 2011). If the TMD Pain

Screener is positive, a further assessment is done to conclude a specific TMD pain-related diagnosis (Schiffman et al.,2014).

2.2.2(b) The DC/TMD Axis II

The Axis II screening instruments comprise of 41 questions which are incorporated in (International RDC/TMD Consortium et al., 2014);

- Pain Drawing form,
- Graded Chronic Pain Scale (GCPS) version 2.0,
- Jaw Functional Limitation Scale (JFLS – scale 8 and scale 20),
- Patient Health Questionnaire (PHQ – scale 4, scale 9 and scale 15),
- Generalized Anxiety Disorders (GAD- scale 7), and
- Oral Behavior Checklist (OBC)

Axis II domains are recommended when screening displays a pain related disorder as per displayed by **Figure 2.4**. In addition, it should also be made compulsory in patients who are presented with pain that lasts for 6 months or longer or in the evidence of failure of a previous treatment received by the patient (Schiffman et al.,2014).

Domain	Instrument	No. of items	Screening	Comprehensive
Pain intensity	Graded Chronic Pain Scale (GCPS)	3	✓	✓
Pain locations	Pain drawing	1	✓	✓
Physical function	Graded Chronic Pain Scale (GCPS)	4	✓	✓
Limitation	Jaw Functional Limitation Scale—short form (JFLS) Jaw Functional Limitation Scale—long form (JFLS)	8 20	✓	✓
Distress	Patient Health Questionnaire-4 (PHQ-4)	4†	✓	
Depression	Patient Health Questionnaire-9* (PHQ-9)	9†		✓
Anxiety	Generalized Anxiety Disorder-7 (GAD-7)	7†		✓
Physical symptoms	Patient Health Questionnaire-15* (PHQ-15)	15		✓
Parafunction	Oral Behaviors Checklist (OBC)	21	✓	✓

*The RDC/TMD depression and nonspecific physical symptoms instruments could be substituted for the PHQ-9 and PHQ-15, respectively, if continuity with legacy data is important.
†Each of the PHQ-4, PHQ-9, and GAD-7 include one additional item beyond the number listed above; the additional item is a global reflective question regarding functional interference due to any of the endorsed symptoms on that instrument.

Figure 2.4: The instruments listed in the Axis II (screening and comprehensive) of DC/TMD with their domains which include all areas of biopsychosocial assessment. (Schiffman et al,2014)

(A) *Axis II screeners* consist of 5 simple self-reported screening instrument which are the GCPs, Pain drawing, JFLS, PHQ and OBC. These screeners are used for detection of pain-relevant psychosocial and behavioral functioning (Schiffman et al.,2014). The Patient Health Questionaire-4 is a valid screening tool for detecting clinical patients of "psychological stress" secondary to anxiety and/or depression (Kroenke, Spitzer, Janet, Williams, & Lö, 2009). The Graded Chronic Pain (GCPS) is a precise, decisive and valid instrument that measures the pain intensity and pain-related impairment (Von Korff, Ormel, Keefe, & Dworkin, 1992). The next instrument is a pain drawing form. This pain drawing instrument allows patients to reflect their pain located at the head, jaw and body region (International RDC/TMD Consortium et al., 2014). A comprehensive and diffuse pain drawing will suggest the need for an overall and complete evaluation of the patient. The fourth instrument is the Jaw Functional Limitation Scale (JFLS). JFLS records the overall constraint of the mastication process, jaw movement, and verbal as well as emotional expression (Ohrbach, 2008). Lastly, the prevalence of oral parafunctional habits are measured by the Oral Behaviors Checklist (OBC) (Ohrbach, 2008).

(B) Comprehensive Axis II Instruments are implemented by the clinical specialists or researchers in respect of gathering a more complete assessment of psychosocial functioning of a patient/subject (Schiffman et al.,2014). Comprehensive Axis II Instruments also assess pain intensity, pain-related impairment and the physical as well as emotional functioning just like the Axis II screeners components. In addition, the new DC/TMD also includes new measures for a more complete evaluation of emotional functioning (Schiffman et al.,2014) via the usage of PHQ-9 for depression and GAD-7 for assessing anxiety. To conclude the Axis II instruments, the new DC/TMD remains a measure for physical symptoms by using the PHQ-15. As overall, the Axis II screening

tools determine the limitation to treatment responses, causes of chronic TMD, and aims for further additional interventions (Schiffman et al.,2014).

2.2.3 DC/TMD Axis I Diagnostic Algorithms

The new DC/TMD is more comprehensive and reliable (Schiffman et al.,2014). As compared to the original RDC/TMD (Dworkin & LeResch, 1992), the new DC/TMD is able to determine the pain-related TMD including producing a valid and reliable Axis I diagnostic algorithms for the most familiar pain-related TMD and for most common intra-articular TMD. The diagnostic criteria for most common intra-articular TMD include;

(i) Disc displacement with reduction,

(ii) Disc displacement with reduction with intermittent locking,

(iii) Disc displacement without reduction with limited opening,

(iv) Disc displacement without reduction without limited mouth opening,

(v) Degenerative joint disease, and

(vi) Subluxation

Following the development of DC/TMD, there are 12 common temporomandibular disorders that have been identified which include arthralgia, myalgia, local myalgia, myofascial pain, myofascial pain with referral, four-disc displacement disorders, degenerative joint disease, subluxation and headache attributed to TMD (Schiffman et al.,2014).

Based on the taxonomic classification for TMD, the temporomandibular disorders can be categorized to temporomandibular joint disorders, masticatory muscle disorders, headache attributed to TMD and TMD related to associated structures such as coronoid hyperplasia (Schiffman et al.,2014). Using the DC/TMD, multiple diagnoses are able to be conducted; muscle related diagnoses, diagnoses for each joint pain, joint disease or for joint disorders, headache contributed to TMD as well as diagnosis of the TMJ associated structures. Treatment plans are based on these diagnoses which should be implemented carefully considering all hazards and advantages of the associated treatment plan (Schiffman et al.,2014). The application of DC/TMD is summarized in **Figure 2.5.**

	Axis I: Physical diagnosis		Axis II: Psychosocial status	
	Pain diagnoses	Joint diagnoses	Distress and pain disability	
Application	Clinical or research		Clinical	Clinical or research
Screening test	TMD pain screener	DC/TMD for disc displacements, degenerative joint disease, and subluxation	PHQ-4 and GCPS	PHQ-9, GAD-7, PHQ-15, and GCPS
Confirmatory test	DC/TMD for myalgia, arthralgia, and headache attributed to TMD	Imaging: MRI for disc displacements, CT for degenerative joint disease, and panoramic radiographs, MRI, or CT for subluxation	Consultation with mental health provider	Structured psychiatric or behavioral medicine interview

Patient Health Questionnaire-4 (PHQ-4), Graded Chronic Pain Scale (GCPS), Patient Health Questionnaire-9 (PHQ-9), Generalized Anxiety Disorder-7 (GAD-7), Patient Health Questionnaire-15 (PHQ-15).

Figure 2.5: The DC/TMD can be applied to the clinical and research settings. Its assessment protocol has both screening and confirmatory test for the most common physical diagnoses (Axis I) and the contributing factors (Axis II). (Schiffman et al.,2014)

2.3 Cross-cultural Adaptation and Translation of Health-Related Quality of Life Measures

A research that connects the language and cultural perimeters certainly needs direct attention to the usage of local language and cultural factors. This is important when the data collection involves verbal expression and comprehension which is expected to display comparable reliability and validity across the linguistic and cultural boundaries (Ohrbach et al.,2013).

A worldwide integration of patient care, medical and dental research require a standard international tool to evaluate health status of an individual (John, Hirsch, Reiber, & Dworkin, 2006). Usually, such instruments are developed in a specific cultural environment. Over the years, the development of health-related quality of life (HRQoL) instruments are increasing worldwide to determine the impact of medical intervention on quality of life (QoL) (Guillemin et al.,1993). Commonly, most of the measures are developed in English and are designed for the usage in English-speaking countries. Thus, there is a need for the measures to be specifically intended for the usage in non-English-speaking countries and also among immigrant population. Therefore, in order to achieve that need, clinicians and researchers have two options (Guillemin et al.,1993);

(i) to develop a new instrument in their own language, or

(ii) to alter an instrument previously validated in another language which is known as a cross-cultural adaptation process.

There are three approaches to develop instruments into a second language proposed by (Bullinger, Anderson, Cella, & Aaronson, 1993) sequential, parallel and simultaneously.

(i) *The sequential approach*: an original instrument in a source language being translated into a target language. It is then re-translated into the source language which is known as back-translation. An equivalency is carried out on an adaptation in the target language. This approach is regularly done worldwide.

(ii)The parallel approach: culturally related information is presented in full length during the development process of the instrument. International consensus of each item is relied upon in order to establish a single set of items. The items should be appropriate to the measurement of the construct in each of the development settings.

(iii) The simultaneously approach: with this approach, a hypothesis is made that both universal and cultural specific assessments are required to exist. In such manner, a given version of a particular language consists of both general items that exist in all language versions as well as items specific to that culture, in relation to the respective construct. However, this method is uncommonly used due to the large scale of resources and coordination which is required for concurrent instrument development in more than one location at one time.

The cross-cultural adaptation process involves a course of translation work and an assessment of the psychometric properties (e.g.; reliability and validity) of the translated instrument in the new cultural environment (John et al., 2006). Based on Last (1995), validity is defined as 'the degree to which a measurement measures what it purports to measure'. A cross-cultural equivalence of an instrument is illustrated by the

successful usage of an instrument in another culture and differ from its original language and culture (John et al., 2006). This process grants support for its construction.

2.3.1 Phase I: Translation theory and cultural adaptation

The aim of cross-cultural translation is to achieve agreement and equality between two different languages. The aim of a consistent cross-cultural translation process is to achieve content and conceptual equivalence. A well-known method for cross cultural research is introduced by Brislin in 1970 known as 'The Brislin' model. This model is used for instrument translation (Jones, Lee, Phillips, Zhang, & Jaceldo, 2001).

Based on this 'The Brislin' model, a bilingual translator (sometimes referred to as 'forward-translator') translates (the instrument from its source language into a target language. The other bilingual translator will then retranslate to source language which is known as back-translation. To make sure the agreement and equality of the translated instruments, the back translation is done by blinding the back-translator from the source document. Both versions (the source and the back-translated version) are then compared for accuracy. Any uncertain items are then determined, and the process of back-translation is done again by another blinded bilingual translator. The process is repeated a few times until the content and the meaning of the translated documents is commonly recognized to be equivalent and distinct. These preferred translation processes are advised to be done by at least two independent bilingual translators. Guillemin et al. (1993), proposed a translation guideline **(Figure 2.6)** that consist of five steps which include:

(i) *Translation: To produce multiple translation by qualified translator.*

The translation process should be done by at least two translators. This is important for the discovery of inaccuracy and misleading of interpretations of multiple

items in the source instruments. The quality of translation will be higher if the task of translating is done by teams rather than single individuals, who are more likely to introduce personal characteristics. A qualified translator also plays an important role in a quality translation which is based on the individual's characteristics and his/her qualifications. A highly-educated person may not be culturally ideal for the target community (Sercherst,1972). The translator should preferably translate into their native language and some of them should also be aware of the aim and objectives of the instruments to be translated as well as the idea involved in the project so that the translated work will be more reliable.

(ii) Back translation by qualified individuals: To produce as many back-translations as forward translation.

Each forward translation should be back translated separately from each other. Divergent of interpretations in the forward translation may be intensified in the back-translation, and thereby revealed. Failure to comply to the target cultural background and uncertainty in the source version can also be revealed during this step.

(iii) Committee review of those translation and back translation: A multidisciplinary team to compare source and final versions.

A multidisciplinary committee should consist of expert (in the disease(s) of the related study) individuals. The committee should be established in order to construct a finalized version based on the input from the forward and back-translation. The committee review members should also analyze the introduction and instruction to the questionnaire as well as to evaluate the response scale of each question.

Based on Guillemin et al., (1993), the roles of the committee review are as below:

(a) to use structured techniques to clear up inconsistency.

(b) to make sure the introduction to the measurement and the instruction to answer the questionnaire is correctly translated in order to maintain the source originality and reproducibility of the measurement. The team can modify the instruction/format or to decline any improper items and may generate replacement for better fitting of the cultural target situation while preserving the general concept of the edited items.

(c) to make sure that the translation is fully understandable. A comprehensible translation should be understandable to the majority of a community which can be understood by 10-12 year old children (Brislin et al., 1973).

(d) to verify a cross equivalence of source and final versions which include;

- semantic equivalence - compatibility in the meaning of words, vocabulary and grammar,
- idiomatic equivalence - translatable idioms and colloquialisms which support crucially in the emotional and social aspect,
- experiential equivalence - situation illustrated in the source version should fit the target cultural environment, and
- conceptual equivalence – specify to the validity of the concept explored and experienced in the target culture. In some situations, the items might be similar in sematic meaning but not similar conceptually.

(v) *Pretesting: To check for equivalence in source and finalized versions.*

The pre-testing method is done purposely to check for any mistake and discrepancy in the translation. It should be done in the same population who answered the questionnaire initially. According to Guillemin et al. (1993), pretesting can be done in two ways – either to use a probe technique or to submit both source and finalized versions to bilingual lay individuals.

Probe technique. The questionnaire is distributed to a group of patients. Patients are then asked the probe question ("What do you mean?") after each answer. Patients are encouraged to explain in detail his/her interpretation of the item in an open-ended manner. This is to make sure that the final item is well and correctly understood as equivalent to the source item.

Submit both source and finalized versions to bilingual lay individuals. The source and finalized versions of the measure can be distributed to a group of bilingual lay individuals. This is to identify any possible errors and deviations of the translation. Following that, they are also asked to rate the equivalence between the two versions. Items with low level of equivalence or rated discrepantly by different individuals need to be revised.

(vi) Re-evaluation of weighting of scores (if necessary).

A scoring method using weight is used in order to integrate the information in an index or in several profiles. The re-evaluation can be done by the judgement method or by using a mathematical approach.

- Judgement method – the cross-cultural validity of the weighting of items is re-evaluated by experts (health care providers, patients or lay individuals).
- Mathematical approach – the data collected form the patients are analyzed by various statistical studies for scalability via Gutmann analysis or dimensionally via factor analysis.

1. *Translation*
   Produce several translations
   Use qualified translators

2. *Back-translation*
   Produce as many back-translations as translations
   Use appropriate back-translators

3. *Committee review*
   Constitute a committee to compare source and final versions
   Membership of the committee should be multidisciplinary
   Use structured techniques to resolve discrepancies
   Modify instructions or format, modify/reject inappropriate items, generate new items
   Ensure that the translation is fully comprehensible
   Verify cross-cultural equivalence of source and final versions

4. *Pre-testing*
   Check for equivalence in source and final versions using a pre-test technique
   Either use a probe technique
   Or submit the source and final versions to bilingual lay people
   Immigrants: Choose the language of administration or use a dual-format measure

5. *Weighting of scores*
   Consider adapting the weights of scores to the cultural context

*It is not always possible to follow all the steps described here due to the design of the measure (e.g. no

Figure 2.6: Translation guidelines proposed by Guillemin et al.,1993

As conclusion made by Guillemin et al. in his publication in 1993, he stated about the benefits of cultural adaptations to a target population as below;

- it produces a common instrument for the evaluation of health-related quality of life in various cultural backgrounds,
- it allows a standard and ideal instrument for use in worldwide research;

- it recognizes comparisons between cultural societies based on a standard and ideal instrument to investigate the phenomenon cross-culturally,
- it is a cheaper and faster method than developing a new instrument as cross-cultural adaptation involves a large scale of the population.

2.3.2 Phase II: Translation Validation and Documentation

2.3.2.1 Sample size

Sample size calculation is very important in any study. In general, the basis of sample size calculation and common factor analysis is "the more the subjects, the better the study will be" (Floyd & Widaman ,1995). According to literatures, for a validation study of an instrument, the sample size calculation is calculated based on the item to patients' ratio which is 5-10 patients: 1 item (Floyd & Widaman, 1995). A validation study of the Chinese version of the oral health impact profile for TMDs (OHIP-TMDs-C) calculated their sample size based on 5:1 ratio (He & Wang, 2015). The same calculation was also used by (He, Wang, Wang, & Deng, 2012) in their validation study of Halitosis Associated Life-quality Test (HALT) questionnaire. Of note, both studies involved Oral Health Related Quality of Life (OHRQoL) measurement.

2.3.2.2 Quality criteria for assessment of questionnaires: Reliability & Validity

Over the years, there were a tremendous number of published systemic reviews associated with available measurements that measure specific concept in a target population (Terwee at al.,2007). In majority of these systemic reviews, the measurement's properties and their contents are compared. To determine the methodological quality of studies on the development and assessment of the health-related instruments, there are some criteria that have been proposed to evaluate those

questionnaires/instruments. The Scientific Advisory Committee (SAC) of the Medical Outcomes Trust (2002), has proposed 8 characteristics of the instrument elements that warrant consideration in evaluation (Aaronson et al., 2002).

The 8 quality criteria for measurements properties of health-related questionnaires are:

i. conceptual and measurement model,
ii. validity,
iii. reliability,
iv. responsiveness,
v. interpretability,
vi. respondent and administrative burden,
vii. alternative forms, and
viii. language and cultural adaptations via translation.

Of all these criteria, reliability and validity are the main benchmark of the quality of a measuring instrument (Kimberlin & Winterstein,2008).

2.3.2.3 Reliability

The criteria of reliability (Kimberlin & Winterstein,2008) was adopted to assess the stability of measures that are distributed to the same individuals at two different times. Reliability can also be used with the same standard via test-retest reliability. In test-retest reliability, the main focus is the degree of providing same or identical answers in given repeated measurements in a stable individual. (Terwee et al.,2006). The value of reliability coefficients varies from 0.00 to 1.00. The interclass correlation coefficient (ICC) is the most appropriate and most accepted reliability criterion for

continuous measure while weighted Cohen's Kappa coefficient should be used for ordinal measures. Higher coefficients demonstrate higher levels of reliability. Terwee et al. (2006), classified a positive rating reliability if the ICC or weighted Kappa is minimally valued at 0.70 in a sample size of at least 50 subjects.

Other than that, reliability is also used to evaluate the similarity series of items from the same test which is known as internal consistency. Based on Terwee et al. (2007), internal consistency refers to the extent to which items in a measurement subscale are integrated. In such a way, all the items are measuring the same concept.

The hypothesis of an internal consistency is that items evaluating the same concept should correlate to each other (Kimberlin & Winterstein, 2008). Internal consistency reliability can be estimated by Cronbach's alpha which is a function of the median interrelationship of items and the number of items in the scale. Cronbach's alpha is considered sufficient enough to measure internal consistency and each of the subscale should be done independently with good Cronbach's alpha values of more than 0.70 (Terwee et al.,2007). Subtypes of reliability tests are shown in **Figure 2.7**

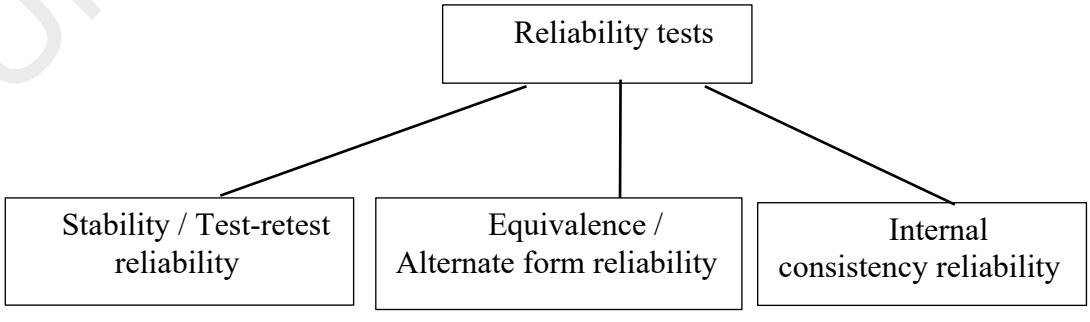


Figure 2.7: Subtypes of reliability tests (Bolarinwa,2015)

2.3.2.4 Validity

Validity refers to the ability to which an instrument measures what it is supposed to measure. **Figure 2.8** shows graphical representation of the subtypes of various forms of validity tests.

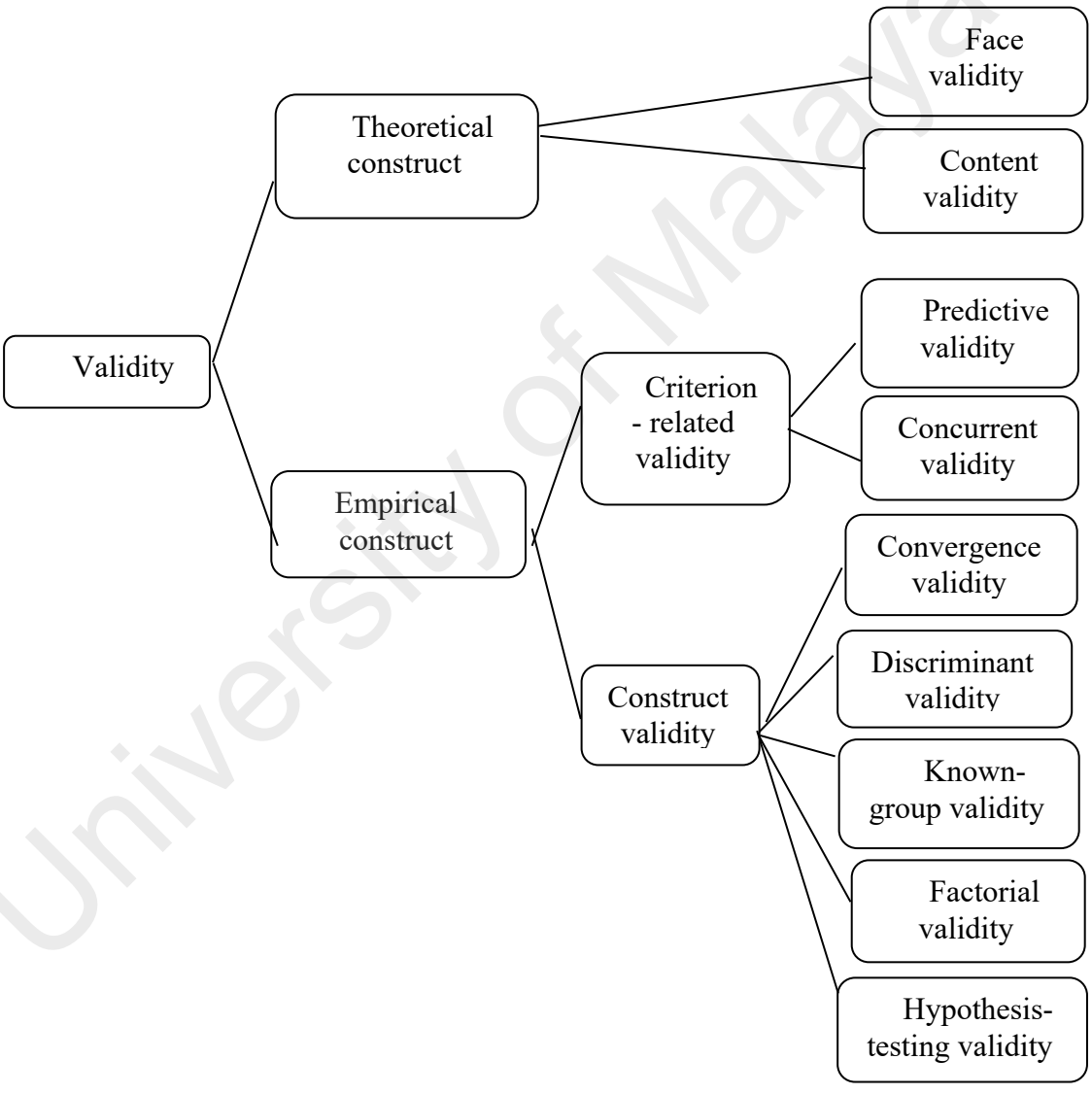


Figure 2.8: Subtypes of validity (Bolarinwa,2015).

2.3.2.4a Theoretical Construct

Theoretical construct can be established using a panel of experts which can be explored in two forms namely, face validity and content validity (Bolarinwa,2015).

Face validity can be obtained when an expert on the research topic evaluating the questionnaires of the particular instrument concludes that it assesses the characteristic of interest (Bölenius et al.,2012). Face validity is a casual and brief validity process.

On the other hand, content validity refers to which the instrument measures the construct of interest (Bölenius et al.,2012). Rating of content validity will be given by the experts to indicate whether an item is +1 (which refers to 'favourable') or +0 which indicate 'unfavourable' (Sangoseni, Hellman, & Hill, 2013). However, over the years, a Likert rating has been developed for content validity. This item rating and scale level will be calculated and a level of agreement between raters will be determined (Sangoseni et al., 2013) known as 'Scale-level Content Validity Indices' (S-CVI). A S-CVI value ≥0.78 is considered a significant level for inclusion of an item into the study (Sangoseni et al., 2013). However, this type of validity is also considered as highly subjective like the face validity and this condition is considered a major drawback for content validity (Bolarinwa,2015).

2.3.2.4b Empirical Construct: Criterion-related validity

Criterion validity is a type of validity when scores of a particular measurement is compared to a standard instrument (Terwee et al.,2007). There are two types on criterion-related validity (Kimberlin & Winterstein,2008);

(a) *predictive validity* – the evaluation is retrieved at some point after the test has been distributed. This validity also evaluates the ability of the test to precisely predict the criterion.

(b) *concurrent validity* – a type of validity that measures the scores of an instrument and correlate them with scores of other instruments that measure the same concept.

This assessment can be done concurrently in the same individuals which will benefit in terms of cost, time and efforts that is involved in the validation of a new instrument.

2.3.2.4c Empirical Construct: Construct validity

Construct validity is a type of validity in which an analysis is made based on the aggregation of data or scores taken from multiple studies using a particular measurement. In construct validity, the theoretically derived hypotheses are expected to be consistent with the scores in other instrument that measure the same concepts (Streiner & Norman,2003). By recommendation by Terwee et al. (2007), a construct validity should be evaluated by experimenting preconstructed hypotheses which need to be very specific. For instance, lower score of quality of life is expected in chronically ill patients compared to the healthy individuals.

2.4 Validation of TMD-related Instruments

There are many published articles regarding the validation of TMD-related instruments especially the RDC/TMD and OHIP-TMD tools. All these tools have been translated and validated into different languages such as Germany, Portuguese, Chinese, Spanish as well as Malay.

2.4.1 Validation of the German Version of the RDC/TMD

The RDC/TMD has been developed and validated into the German language by John et al. in 2006. The reliability and validity tests of the German-language RDC/TMD were assessed in 2 domains via ;(i) the Jaw Disability List (JDL) and (ii) the Graded Pain Scale (GCPS). The sample size for this study was 378 TMD patients.

(i) Jaw Disability List (JDL)

A test-retest reliability study was conducted using a time interval of 1-2 weeks which involved 27 adult patients during the second questionnaire assessment. Calculation of the Interclass correlation coefficient (ICC) was conducted for the summary score of all limited mandibular functions. In addition, the internal consistency was calculated using Cronbach's alpha.

For the validation test, the construct validity assessment was conducted by evaluating the correlation between the sum of the jaw disability items. All these correlations were measured via Spearman's rank correlations and the items are as below;

- Self-reported oral health (very good, good, fair, poor);
- Functional limitations scale of oral health related quality of life measured by the German-language Oral Health Impact Profile (OHIP-G);

- Limited mouth opening (which is described as dynamic mouth opening less than 40mm without pain); and
- Self-reported oral behavior such as nails, tongue, lip, cheek or object biting.

(ii) Graded Pain Scale (GCPS)

For this scale, John et al. (2006) conducted a misclassification matrix evaluation for the test-retest (rows are for the initial GCPS scores while the column is for the repeated GCPS scores) and used the ICC for the calculation measures. The internal consistency was also established via Cronbach's alpha. Construct validity was assessed via measuring the correlations between;

- dysfunctional chronic pain (determined as GCPS grade III or IV) and self-reported general health which is classified as very good, good, fair, and poor).
- dysfunctional chronic pain and scales of Oral Health Related Quality of Life (part of OHIP-German) which consist of 'psychological discomfort', 'physical disability', 'psychological disability', 'social disability', and 'handicap'.

The association between global rating of general health / OHRQoL scores and GCPS grades were calculated via Spearman's rank correlations.

2.4.2 Validation of the Portuguese Version of the RDC/TMD

A published paper by Lucena, Kosminsky, Costa, & Góes, (2006) aimed at establishing the validity of the Portuguese version of RDC/TMD questionnaire. This study was conducted on 155 TMD patients. The validation procedures were made up of:

(i) Internal consistency: which was evaluated from the 12 items of the limitation scale on mandibular functions. The degree of association was calculated between several DC/TMD Axis II domains, in consideration of psychological factors, graded chronic pain severity and limitation scale on mandibular functions which are the RDC/TMD subscales. Internal consistency was calculated by Cronbach's alpha test with satisfactory value of >0.5.

(ii) Reliability and reproducibility: Test-retest study was administrated on 45 subjects who are randomly selected. Time interval between the two tests was 2 weeks. None of the selected patients had received any TMD physical and pharmacological therapy. The Cohen Kappa scale and Spearman's rank correlation were done to evaluate the test-retest reliability.

(ii) Concurrent validity: In this type of validity, the Portuguese version of RDC/TMD was simultaneously tested against the gold standard Oral Impacts on Daily Performances as well as the Oral Health Impact Profile-14 (OHIP14) instruments. The concurrent validity (convergent type) was obtained by correlating the domains result of RDC/TMD (limitation scale on mandibular functions) with other two instruments that measure the same concept which are Oral Impacts on Daily Performances and OHIP14. In addition, the scores from Graded Chronic Pain Severity (GCPS) of RDC/TMD were also correlated with the designated TMD scores from the Simplified Anamnestic Index.

The concurrent and convergent validity were calculated via Spearman's rank correlation test. This comparison was statistically significant with $\mu = 0.5$ and *p* value was < 0.01.

2.4.3 Validation of the Spanish Version of the RDC/TMD

The aim of this study was to develop the cross-cultural adaptation or equivalence of the Spanish-language RDC/TMD (Gonzalez et al., 2013). This study was conducted on 33 bilingual individuals with and without TMD. The validation approach was as follows;

(i) Reliability: The analysis of reliability was done through the test-retest study which involved the Axis I and II of RDC/TMD. For all of the categorical variables, a chi-square analysis and descriptive statistics were conducted. Furthermore, the reliability was also measured by calculating the intraclass correlation coefficient (ICC) which was assessed for the data with continuous scales, while Cronbach's alpha was analyzed with categorical variables. The description values were based on the "κ-value". Values of $0.8 \leq \kappa \leq 1.0$ were considered excellent reliability, while values of $0.60 \leq \kappa < 0.80$ were translated as acceptable reliability. Lastly, a moderate reliability was valued as $0.40 \leq \kappa < 0.60$.

(ii) Internal consistency (validity): This validity involved the Axis II RDC/TMD which was the 12-items jaw disability checklist as well as the Graded Chronic Pain Scale and psychological factors. The correlations between the different variables were calculated via Spearman's Correlation Coefficient. The Spearman's score was interpreted as high ($r \geq 0.8$), moderate ($r \geq 0.4$) and weak ($r < 0.4$).

2.4.4 Validation of the Malay Version of the RDC/TMD

The Malay RDC/TMD was developed and validated by Khoo, U Jin Yap, Chan, & Bulgiba (2008). The Malay version of RDC/TMD was evaluated for internal consistency and validity. For the aim of this study, a total of 80 subjects were used which consisted of 40 TMD individuals and 40 pain-free individuals without any sign and symptoms of TMD. The data collected for this study was used for the evaluation of internal consistency and validity as below:

(i) Internal consistency: were tested using a total of 40 individuals with TMD symptoms. The tests were classified into 3 main domains (Graded Chronic Pain Scale, Non-Specific Physical Symptoms inclusive of pain and non-pain items, and Depression). All of the domains were evaluated of its internal consistency via Cronbach's alpha.

(ii) Test-retest reliability: the scores and the level were collected in the time interval of 1 week. The data collected at the initial phase is labelled as baseline data. The second data which was collected from the same 40 TMD individuals was compared with the baseline. The data (scores and level) were then evaluated using interclass correlation coefficient (ICC) and Spearman's rho correlation respectively. The values of both tests ranged from '1' which indicates perfectly reliable to '0' which suggests totally unreliable.

(iii) Discriminant validity: this type of validity was done for the 3 main domains. It was calculated by the difference in value of mean between TMD patients (n = 40) and non TMD individuals (n= 40*)*. Discriminant validity was tested using independent samples *t* test.

CHAPTER 3: MATERIALS AND METHODS

3.1 Phase I: Translation and Cultural Adaptation

There were two phases involved in the development of the Malay version of DC/TMD which are phase I and phase II. Phase I involved the process of translation and cultural adaptation while phase II was carried out to assess the psychometric properties of the Malay DC/TMD. The components for translation were based on the 'Translation and Adaptation of the DC/TMD Protocol' prepared by International RDC/TMD Consortium et al.(2014). The components of DC/TMD that were translated are as below;

(i) DC/TMD symptoms questionnaire,

(ii) TMD pain screener,

(iii) DC/TMD demographics,

(iv) Pain drawing,

(v) Graded Chronic Pain Scale Version 2.0 (GCPS – version 2),

(vi) Jaw Functional Limitations Scale (JFLS) 8-item and 20-item versions,

(vii) Patient Health Questionnaire (PHQ) 4-item, 9-item and 15-item versions,

(viii) Generalized Anxiety Disorder 7 item (GAD-7), and

(ix) Oral Behaviours Checklist (OBC)

Phase I involved the process of the English translated DC/TMD to the Malay version which was described as below (Based on 'Guidelines for Establishing Cultural Equivalency of Instruments, Ohrbach et al.,2013):

(i) *Forward translation*: The English version of DC/TMD (referred as the source instrument) was translated into the target language which is the Malay language. Three forward translators who were bilingual individuals were recruited/employed in the first step of forward translation which was done independently. The bilingual translators were Malay language speakers whose Malay is their native language, while the source language is their second language. One of the forward translators was naïve to the instrument's concept, while the other two translators were experts in the content of the disorder. This was important to preserve the cultural representation of the target language population who will be using the instrument.

(ii) *Synthesis and determination of inconsistency:* The three forward translations were then synthesized into a single version of the forward translation. The elements to be considered during the merging were conceptual equivalence, conversational and common language used as well as specific language term accuracy. A final review of the synthesis translation was conducted among the team before it was pursued for a back-translation.

(iii) Back-translation: A blinded independent back-translation was conducted to maintain the quality of the instrument development. Back-translation process was done by two bilingual back-translators who were experts on the source language. The back - translators were from the English Language Department,

University of Malaya who are both naive to the subject content of the instrument in order to scale down the information biasness as well as to highlight any unanticipated meanings.

(iii) *Expert review, revision and consolidation:* The back-translated versions were then reviewed independently by individuals (forming a review team that consists of experts of the targeted concept) that were not involved in the forward or back-translation processes. The reviewers' scopes were to rectify the discrepancies between back-translation and the source instrument. This process of review, appraisal and approval was followed by another cycle of translation and review until all translation units were in good achievement and in agreement of the translation. The translated units were compiled to produce a consolidated version of translation which is known as pre-final instrument.

3.2 Phase II: Validation

Phase II involved the validation process following the translation works. To assess the reliability and validity of the translated instrument, there were two domains of DC/TMD that were used in this study, (i) the Graded Chronic Pain Scale version 2 (GCPS) and (ii) the Jaw Functional Limitation Scale – 20 items (JFLS).

3.2.1 Study Population

Convenient subjects recruited for this study were selected among registered patients who attended the Faculty of Dentistry, University of Malaya for any other dental treatments. Among others, we have also recruited the undergraduate dental students of University Malaya as well as random eligible publics based on the inclusion

and exclusion criteria proposed for this study. Data collection duration was from September 2016 to February 2018.

3.2.2 Sample size calculation

The sample size calculation was based on the subjects to item ratio which were 5 to 10 subjects per item of each questionnaire (Floyd &Widaman,1995; He et al., 2012; He & Wang, 2015). A total of 252 subjects were involved for the first convenient sample, consisting of 165 non TMD individuals and 87 TMD patients. From the 252 subjects, a second convenient sample of 40 individuals (20 non TMD individuals and 20 TMD patients) were asked to answer the same questionnaire after 2 weeks to investigate test-retest reliability (Terwee et al.,2007; John et al.,2006; Lucena et al., 2006). TMD patients were identified clinically based on the existing signs and symptoms of TMD supported by the Axis I of DC/TMD (Pain Screener Questionnaire).

3.2.3 Sample selection

This study recruited subjects using convenience sampling technique of those who fulfilled the inclusion and exclusion criteria. The associated criteria were as below;

A) Inclusion criteria were divided into TMD cases and non-TMD cases.

(i) **<u>for TMD cases;</u>**

a) individuals who were 18 years old and above,

b) each subject must be presented with at least one sign of TMD. The signs include either presence of pain in the jaw, TMJ area and adjacent structures regardless at rest or during jaw action,

c) subjects should meet/comply with the DC/TMD symptoms questionnaire in Axis I of DC/TMD,

d) subjects should be able to understand and answer the Malay-DC/TMD instrument.

(ii) **For non TMD cases**;

a) individuals who were 18 years old and above,

b) subjects should be able to understand and answer the Malay-DC/TMD instrument.

B) The exclusion criteria for this study were;

a) subjects with the presence of organic pathology related to the TMJ area or history of TMJ trauma,

b) individuals with major medical history or diagnosed psychiatric disorders.

3.2.4 Study tools

In this study, the psychometric properties in the evaluation instrument abided on reliability and validity. The measurement for the Malay-DC/TMD was determined using two domains:

a) **Graded Chronic Pain Scale (GCPS)** which consists of 'Characteristic of Pain Intensity' subdomain (item no 2 to 4), 'Interference of Activities' subdomain (item no 6 to 8) and 'Days with Interference' subdomain (item no 5). However, item no 1

(perception of pain within 6 months) was excluded from the statistical analysis due to the range of scores for this item differed widely as compared to the other 7 items (Pena et al, 2016), in which the perception of pain is just within 3 months.

For GCPS, the response options for item no 1 to 4 were scaled from 0 (which indicates 'no pain') to 10 (which indicates 'pain as bad as could be'), while for item no 6 to 8, the scale ranged from 0 which indicates 'no interference' to maximum of 10 which indicates 'unable to perform any activities'. However, the response option for item no 5 in GCPS was open-ended where the subjects were required to answer the item in terms of the number of days that they had facial pain that interfered with their usual activities within the last 30 days. GCPS scoring are as below;

i) 'Characteristic Pain Intensity'- (CPI) subdomain: mean of items no 2 to 4 and multiply by 10.

ii) 'Interference of Activities' subdomain: mean of items no 6 to 8 and multiply by 10. Interference mean 0-29 (no point), 30-49 (1 point), 50-69 (2 points) and 70+ (3 points).

iii) 'Days with Interference' subdomain: assigned points were based on the number of days with interference; 0-1 day (no point), 2 days (1 point), 3-5 days (2 points) and 6+ days (3 points).

Total Disability Points = Points for 'Days with Interference' subdomain + Points for Interference of Activities' subdomain

Determination of Chronic Pain Grade;

Grade	Label	CPI	Disability Points
0	None	0	N/A
I	Low intensity pain, without disability	< 50	< 3
II	High intensity pain, without disability	≥ 50	< 3
III	Moderate limitation	N/A	3 – 4
IV	Severe limitation	N/A	5 – 6

b) **Jaw Functional Limitation Scale (JFLS)** consists of 3 types of functional limitation items which are 'Mastication' subdomain (item no 1 to 6), 'Vertical Jaw Mobility' subdomain (item no 7 – 10) and 'Verbal & Non-Verbal Communication' subdomain (item no 13 to 20). For JFLS, the response options were in continuous scale ranged from 0 (no limitation) to 10 (severe limitation). A single global score of "jaw functional limitation" was computed as mean of the 3 subdomain scores. Subdomain scores for each type of the functional limitation of the jaw can be determined by:

- 'Mastication' subdomain: mean items no 1-6
- 'Mobility' subdomain: mean items no 7-10
- 'Verbal and non-verbal communication' subdomain: mean items no 13-20

c) The **Malay version of Brief Pain Inventory (M-BPI)** consists of 2 items which are 'Characteristic Pain Intensity' and 'Interference of Function'. The Malay GCPS was concurrently validated with M-BPI based on these 2 items. The 'Days with Interference' item was excluded. The 'Characteristic Pain Intensity' of M-BPI was determined by a composite of the four pain items (item no 3 to 6) and a mean severity score was obtained (Cleeland, 2009). The response options for those 4 items ranged

from 0 (no pain) to 10 (pain as bad as you can imagine). Meanwhile, the M-BPI 'Interference of Function' item was typically measured as the mean of the 7 interference items (item no 9 to 15) with response options scaled from 0 (does not interfere) to maximum of 10 (completely interfere). Other items of M-BPI (item no 1,2,7 and 8) were not used as those items do not exhibit its psychometric properties.

d) **The Malay-Oral Health Impact Profile 14 (M-OHIP14)**, consists of 7 subdomains which are 'Functional Limitation' (item no 1 and 2), 'Physical Pain' (item no 3 and 4), 'Psychological Discomfort' (item no 5 and 6), 'Physical Disability' (item no 7 and 8), 'Psychological Disability' (item no 9 and 10), 'Social Disability' (item no 11 and 12), and 'Handicap' (item no 13 and 14). The subjects were asked to answer based on a five-point frequency Likert scale (very often, quite often, sometimes, once a while, and never). The severity of impact was calculated by adding up the response codes for each item. Higher scores of M-OHIP14 indicated poorer OHRQoL. The mean of the additive score was used for statistical analysis.

e) **Self-reported Global Oral Health** consists of 3 general questions that evaluate on the subject's:

i. Perception of the oral/jaw health status – 'In your opinion, how do you rate your oral/jaw health status?'. The response was provided based on a 5-point Likert scale as 'excellent', 'very good', 'good', 'moderate', and 'very poor'.

ii. Perception on satisfaction level of the oral/jaw health status – 'In general, are you satisfied with the health of your mouth and jaw?'. The response options were 'very satisfied', 'quite satisfied', 'moderately satisfied', 'not satisfied', and 'very dissatisfied'.

iii. Perception of the need for treatment - 'In your opinion, do you need any treatment related to your mouth and jaws?'. Options for response were 'yes', 'no', and 'don't know'.

The mean of the additive score for every question was used for statistical analysis.

f) **Limited mouth opening** item is defined as painless active mouth opening which is less than 4cm (John et al, 2006). The constructed question for this item was, "In your opinion, can you open your mouth wide enough (4cm/3 finger breadths) without feeling pain at your jaw?". The options for this question were 'Yes' or No'.

3.2.5 Conduct of the study

Phase II involved the validation processes following the translation works. The phase II processes were as follow;

(i) Pre-test and review: The pre-final instruments (finalized translated version of phase I) has been exposed to a series of assessments to evaluate the comprehensibility, practicability and recognition of the instruments at the item level. In the pre-test step, the instrument was administered to a small number of targeted samples that focus on the uncertainty of item construction and differences in understanding in the intended meaning of the items. In this study, the instrument was pre-tested via clinical sampling that consisted of 10 subjects. All the 10 subjects were verbally and independently interviewed following administration of the instrument. The interview was conducted immediately after the subjects have answered the questionnaire.

In the interviews, subjects were asked regarding the intended meaning of items and requested them to;

(a) explain their understanding of the instruments' concept;

(b) produce a substitute phrasing for the items if any ambiguities occurred;

(c) provide an overall feedback of the instrument.

Following the pre-test, any discrepancies of the items were identified and improved. The revision of the forward and back-translations were conducted at a certain point of the translated version to correct the errors based on the responses of the pre-test subjects. Finally, a complete working draft (known as the final instrument) was produced. The set of questionnaires were self-administered to 252 subjects and was completed in a quiet room. The subjects could consult the researcher for any uncertainty.

(ii) Internal reliability – a convenient sample of 40 individuals were selected for the test-retest reliability. All the 40 individuals (consisted of 20 non-TMD subjects and 20 TMD patients) were instructed to answer the same questionnaire 2 weeks after the initial administration to reduce recall bias. The first data collection was labelled as baseline data (T1) while retest data was marked as retest (T2).

(iii) Validity – to establish validity, correlations between scores on the DC/TMD domains (GCPS and JFLS) and scores on other related instruments were used. The types of validity that were accomplished are described as below:

(a) Concurrent validity: to measure how well a new instrument is compared to a well-established tool.

In concurrent validity;

- The domain of the **Malay GCPS** was concurrently validated against **the Malay version of Brief Pain Inventory (M-BPI)** scale which has been validated by Aisyaturridha, Naing, & Nizar (2006).

- The **Malay JFLS** was concurrently validated against the **Malay version of OHIP-14** (Saub, Locker, Allison, & Disman, 2007).

(b) Construct validity: how well a test measures up to its claims. Construct validity was assessed by evaluating the association between;

a) **<u>Graded Chronic Pain Scale (GCPS)</u>**

i. between the **Malay GCPS** and the self-reported **Global Oral Health** ('perceived oral/jaw health status', 'perceived satisfaction level of oral/jaw health' and 'the need for treatment'),

ii. between the **Malay GCPS** and **M-OHIP14** which involved 6 subdomains (physical pain, psychological discomfort, physical disability, psychological disability, social disability, and handicap).

Hypothesis:

i. Subjects with poor oral/jaw health status have higher Malay GCPS scores than those with good oral/jaw health status.

ii. Subjects who were dissatisfied with their level of oral/jaw health have higher Malay GCPS scores than those who were satisfied with their level of oral/jaw health.

iii. Subjects who need oral/jaw treatment(s) have higher Malay GCPS scores than those who do not need any oral/jaw treatment.

iv. Subjects with physical pain have higher Malay GCPS scores than those who do not have physical pain.

v. Subjects with psychological discomfort have higher Malay GCPS scores than those who do not have psychological discomfort.

vi. Subjects with physical disability have higher Malay GCPS scores than those who do not have physical disability.

vii. Subjects with psychological disability have higher Malay GCPS scores than those who do not have psychological disability.

viii. Subjects with social disability have higher Malay GCPS scores than those who do not have social disability.

ix. Subjects who were handicapped have higher Malay GCPS scores than those who were not handicapped.

b) <u>Jaw functional limitation Scale (JFLS)</u>

i. between the **Malay JFLS** and the self-reported **Global Oral Health** ('perceived oral/jaw health status', 'perceived satisfaction level of oral/jaw health' and 'the need for treatment'),

ii. between the **Malay JFLS** and **M-OHIP14** (jaw functional limitation item), and

iii. between the **Malay JFLS and limited mouth opening** question.

Hypothesis:

i. Subjects with poor oral/jaw health status have higher Malay JFLS scores than those with good oral/jaw health status.

ii. Subjects who were dissatisfied with their level of oral/jaw health have higher Malay JFLS scores than those who were satisfied with their level of oral/jaw health.

iii. Subjects who need oral/jaw treatment(s) have higher Malay JFLS scores than those who do not need oral/jaw treatment(s).

iv. Subjects with jaw functional limitation(s) have higher Malay JFLS scores than those who do not have jaw functional limitation(s).

v. Subjects with limited mouth opening have higher Malay JFLS scores than those who do not have any limited mouth opening.

(c) *Discriminant validity: to assess GCPS and JFLS abilities to discriminate patients with TMD and without TMD*. The diagnosis of TMD was done based on the clinical examination supported by the findings using the Pain Screener questionnaire of Axis I in DC/TMD.

3.2.6 Statistical Analysis

Data collected from the instruments including the subjects' demography was keyed into and analysed using the IBM SPSS Statistical Software version 22. The demographic data was analysed using non-parametric descriptive statistics to determine the subject's characteristics such as gender, age, race, marital status, level of education and income.

A) Internal Consistency & Test-Retest Reliability

A reliability test was run on SPSS by calculating the Cronbach's alpha which is the indicator for internal consistency reliability. A coefficient alpha reading of above 0.7 is generally acceptable thus indicating data in this study was reliable and reproducible (Bland and Altman,1997). The Cronbach's Alpha values were determined based on the total score of the Malay GCPS and Malay JFLS. The item-total correlation and inter-item correlation of both domains were also analysed.

The test-retest reliability was established via Interclass Correlation Coefficients (ICC) which were collected from 40 subjects who completed the questionnaire after the 2-weeks interval. The ICC descriptors values (Bartko, 1966) designated;

- < 0.4 (poor to fair reliability)
- 0.41 – 0.61 (moderate reliability)
- 0.61- 0.80 (good reliability)
- >0.80 (excellent reliability)

The determination of ICC for the Malay GCPS was done for all the 3 items which were 'Characteristic Pain Intensity', 'Interference of Function' and 'Days with Interference', while for the Malay JFLS, the assessment was done based on the 3 items of the functional limitation which were 'Mastication', 'Mobility' and 'Verbal & Non-verbal Communication'.

B) <u>Validity</u>

(i) Concurrent validity

The concurrent validity was assessed through the Spearman's rho correlation test. Spearman's rho correlation test was conducted based on the non-parametric data. The Malay GCPS scores was concurrently validated with the Malay version of Brief Pain Inventory (M-BPI) while the Malay JFLS was validated against the Malay-Oral Health Impact Profile -14 items (M-OHIP 14).

The degree of correlation among different variables of the DC/TMD Axis II was evaluated using the Spearman's rank correlations. The correlation values indications (Feise and Menk, 2001) are as below;

- $r < 0.20$ (poor correlation)
- r 0.21- 0.40 (fair correlation)
- r 0.41- 0.60 (good correlation)
- r 0.61-0.80 (very good correlation)
- $r > 0.81$ (excellent correlation)

(ii) Construct validity

The construct validity was established by the Spearman's rho correlation test as all data were non-parametric data. The Malay GCPS and Malay JFLS scores were associated with the mean scores of M-OHIP14 to assess the construct validity. Other than that, the Kruskal-Wallis test was also conducted as the non-parametric data was used to compare more than 2 outcomes in each Global Oral Health questions ('perceived oral/jaw health status', 'perceived satisfaction level of oral/jaw health' and 'the need for treatment').

The association between the Malay JFLS and 'limited mouth opening' question was established via Mann-Whitney U test.

(iii) Discriminant validity

Calculations by comparing the mean of TMD cases with means of controls (which are non-TMD patients) to calculate the Discriminant validity using the Mann-Whitney U test. This test was used to compare if there was a difference *in* the dependent variables (The Malay GCPS and Malay JFLS) for two independent groups which were the TMD and control group.

3.2.7 Ethical Approval and Funding

This study was approved by the ethics committee of The Dental Faculty, University of Malaya, Kuala Lumpur (Ethics Number: DF OS1623/0069P) supported by University Malaya Postgraduate by Coursework Research Fund (Research Grant Number: PPPC/C1-2016/DGJ/05). All subjects obtained for this study were provided with a written consent.

CHAPTER 4: DATA ANALYSIS AND RESULTS

4.1 Validity of Translation

The English version of DC/TMD was translated into the Malay DC/TMD via forward and back-translation technique. Forward translation was done by 3 forward translators. During this process, there was no major issues except for the word 'temple'. At first, the word 'temple" was translated into temporal (first forward translator), side of the head (*bahagian tepi kepala* by the second forward translator) and the side of the forehead - left or right (*tepi dahi sebelah kanan atau kiri* by the third forward translator). Upon discussion among review team members, we decided to choose 'side of the forehead - left or right' (*tepi dahi sebelah kanan atau kiri*) as the best translation for the term 'temple'. Due to the inaccuracy of the translation, the word 'temple' has caused many confusions among subjects during pre-testing phase. Thus, a second review was conducted. A final decision of not to translate the word 'temple' into Malay was taken into account. Instead, the word 'temple' was directly explained personally to the subjects by the examiner during the distribution of the questionnaires. As a result, great response and better understanding among subjects was obtained.

Cross-cultural adaptation of the source instrument also involved foods and common musical instruments. Examples of Western foods such as macaroni and pureed food was replaced by *Kuey Teow* (a type of Malaysian noodle) and bubur kanji (Starch porridge). There was no example of musical instruments in the 'Oral Behaviour Checklist' domain. Thus, we decided to include examples of common musical instruments among Malaysians such as saxophone, trumpet and violin for better understanding in the constructed item.

Back translation was compared with the source version. Most of the concepts and meanings of the translated version was retained. No substantive difficulties were encountered during back translation. In general, the forward and backward translations of this study were done according to the established guidelines without any major issues.

4.2 Demographic Data

This study was conducted from September 2016 to February 2018. A total of 252 subjects were involved in this study which consisted of 171 non-TMD subjects and 81 TMD patients. In total, the response rate was 100%. The demographic data was analysed using non-parametric descriptive statistics. The distribution of the subjects' characteristics are depicted in **Table 4.1** below. Subject's gender, age, race, marital status, level of education and income were tabulated. From the table below, it is clearly shown that the majority of the subjects involved in the current study were females (69.8%) and most of the subjects were aged between 18-30 years old (69.9%). Almost three quarters (72.2%) of the subjects were Malays and 73% of the total subjects were either single or not married. All of the subjects were literate and 77% of them have received the tertiary level of education. In general, the subjects were mostly from the middle-income group.

Table 4.1: Descriptive Statistics of the Demographic Data (N = 252)

Characteristic	Item	n	%
Gender	Male	68	27.0
	Female	176	69.8
Age	18-30 years old	176	69.9
	31-40 years old	55	21.8
	41-50 years old	6	2.4
	51-60 years old	5	2.0
	61 and above	2	0.8
Race	Malay	182	72.2
	Chinese	31	12.3
	Indian	36	14.3
	Kadazan/Iban	1	0.4
	Other Bumiputera	1	0.4
	Others	1	0.4
Marital Status	Married	67	26.6
	Staying together	1	0.4
	Single/Never married	184	73.0
Level of Education	Primary school	3	1.2
	Secondary school	55	21.8
	Diploma / College	109	43.3
	Degree	80	31.7
	Masters/PhD	5	2.0
Income	RM1200 & below	36	14.3
	RM1201 – RM2500	50	19.8
	RM2501 - RM5000	72	28.6
	RM5001 – RM7500	59	23.5
	RM7501 – RM10,000	16	6.3
	RM10,000 & above	19	7.5

*N less than 252 were due to missing data (e.g.: age and gender)

4.3 Psychometric Properties

4.3.1 Internal Consistency of the Malay GCPS and Malay JFLS

The internal consistency of this study was done by calculating the Cronbach's Alpha value among the items in the respective domains as demonstrated in **Table 4.2**. The Cronbach's α coefficients for both domains (the Malay GCPS and Malay JFLS) were found to have high internal consistency with the α values of 0.95 and 0.97 respectively. Values higher than 0.7 is considered ideal and highly consistent (Bland and Altman,1997).

Table 4.2: Internal consistency (Cronbach's Alpha) for 2 domains of the Malay - DC/TMD

Domain	Internal consistency (Cronbach's Alpha)	Cronbach's Alpha Based on Standardized Items	N of itemN of items
Graded Chronic Pain Scale Version 2 (GCPS)	0.95	0.96	7
Jaw Functional Limitation Scale (JFLS)	0.97	0.97	20

For the GCPS domain, **Table 4.3a** and **Table 4.3b** showed the internal consistency values for the item-total and inter-item correlation respectively. The corrected item-total correlation values ranged from 0.56 for 'days with interference' subdomain to 0.95 for the 'average pain' subdomain. All the 7 items of GCPS surpassed the value of 0.20 which is the least recommended correlation value (Vetter T & Schober P, 2018). Any corrected item-total correlations that is lesser than 0.2 should be revised or excluded. In general, the alpha value did not significantly reduce if any of the tested items were removed.

Table 4.3a: Internal Consistency (Corrected Item-total correlations) for the Malay GCPS

Malay GCPS Items (N = 7) Version 2.0	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
'Characteristics Pain Intensity' subdomain		
• **'Pain right now' item (GCPS 2)**	0.85	0.95
• **'Worst pain' item (GCPS 3)**	0.90	0.95
• **'Average pain' item (GCPS4)**	0.95	0.94
'Days with interference' subdomain		
• **'Interference days' item (GCPS 5)**	0.56	0.97
'Interference of function' subdomain		
• **'Daily activities' item (GCPS 6)**	0.93	0.94
• **'Social activities' item (GCPS 7)**	0.92	0.94
• **'Work activities' item (GCPS 8)**	0.90	0.95

The inter-item correlation for the Malay GCPS is presented in **Table 4.3b**. All the Malay GCPS items were inter- correlated except for item 1 (GCPS 1). The lowest correlation value was 0.45 (correlation between GCPS 2 and GCPS 5), while the highest correlation value was 0.93 (correlation between GCPS 6 and GCPS 7).

Table 4.3b: Internal Consistency (Inter-item correlations matrix) for Malay GCPS

Items	GCPS 2	GCPS 3	GCPS 4	GCPS 5	GCPS 6	GCPS 7	GCPS 8
GCPS 2	1.00	0.84	0.86	0.45	0.78	0.79	0.78
GCPS 3	0.84	1.00	0.92	0.54	0.83	0.81	0.81
GCPS 4	0.86	0.93	1.00	0.52	0.87	0.88	0.89
GCPS 5	0.45	0.54	0.52	1.00	0.62	0.55	0.45
GCPS 6	0.78	0.83	0.87	0.62	1.00	0.93	0.88
GCPS 7	0.79	0.81	0.88	0.58	0.93	1.00	0.90
GCPS 8	0.78	0.81	0.89	0.45	0.88	0.90	1.00

The degree of correlation between the different items of the Malay JFLS is illustrated in **Table 4.4a** (corrected item-total correlations) and **Table 4.4b** (inter-item correlations). Based on **Table 4.4a**, the corrected item-total correlations values ranged from 0.70 for 'eating soft food requiring no chewing' item to 0.88 for 'putting on a happy face' item. All 20 items exceeded the minimum recommended correlations value and Cronbach's Alpha did not change if any of the 20 items were deleted.

Table 4.4a: Internal Consistency (Corrected Item-total correlations) for the Malay JFLS

JFLS Items (N = 20) Functional Limitation		Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
'Mastication' subdomain			
• **'Chew tough food' item**	**(JFLS 1)**	0.78	0.97
• **'Chew hard bread' item**	**(JFLS 2)**	0.82	0.97
• **'Chew chicken' item**	**(JFLS 3)**	0.76	0.97
• **'Chew crackers' item**	**(JFLS 4)**	0.80	0.97
• **'Chew soft food' item**	**(JFLS 5)**	0.76	0.97
• **'Eat soft food requiring no chewing' item**	**(JFLS 6)**	0.70	0.97
***'Mobility' subdomain* – open wide enough:**			
• **'To bite from a whole apple' item**	**(JFLS 7)**	0.78	0.97
• **'To bite into a sandwich' item**	**(JFLS 8)**	0.81	0.97
• **'To talk' item**	**(JFLS 9)**	0.86	0.97
• **'To drink from a cup' item**	**(JFLS 10)**	0.83	0.97
'Verbal & Non-verbal communication' subdomain			
• **'Swallow' item**	**(JFLS 11)**	0.80	0.97
• **'Yawn' item**	**(JFLS 12)**	0.78	0.97
• **'Talk' item**	**(JFLS 13)**	0.86	0.97
• **'Sing' item**	**(JFLS 14)**	0.82	0.97
• **'Putting on a happy face' item**	**(JFLS 15)**	0.88	0.97
• **'Putting on an angry face' item**	**(JFLS 16)**	0.83	0.97
• **'Frown' item**	**(JFLS 17)**	0.82	0.97
• **'Kiss' item**	**(JFLS 18)**	0.71	0.97
• **'Smile' item**	**(JFLS 19)**	0.80	0.97
• **'Laugh' item**	**(JFLS 20)**	0.75	0.97

The inter-item correlation for JFLS is shown in **Table 4.4b**. All of the 20 JFLS items were inter-correlated with the lowest value of 0.40 denoted by the correlation between item 7 and item 18, while the highest correlation value was 0.98 for the correlation between item 16 and item 17.

Table 4.4b : Internal Consistency (Inter-item correlations) for Malay JFLS

Items	JFLS1	JFLS2	JFLS3	JFLS4	JFLS5	JFLS6	JFLS7	JFLS8	JFLS9	JFLS10	JFLS11	JFLS12	JFLS13	JFLS14	JFLS15	JFLS16	JFLS17	JFLS18	JFLS19	JFLS20
JFLS1	1.000	0.903	0.895	0.849	0.587	0.500	0.777	0.782	0.613	0.529	0.552	0.648	0.564	0.502	0.538	0.510	0.507	0.424	0.495	0.508
JFLS2	0.903	1.000	0.836	0.842	0.569	0.552	0.781	0.823	0.687	0.629	0.544	0.690	0.587	0.604	0.565	0.594	0.585	0.551	0.484	0.461
JFLS3	0.895	0.836	1.000	0.896	0.608	0.495	0.760	0.757	0.594	0.523	0.543	0.663	0.566	0.437	0.505	0.465	0.461	0.416	0.481	0.546
JFLS4	0.849	0.842	0.896	1.000	0.591	0.500	0.803	0.804	0.624	0.586	0.570	0.685	0.606	0.508	0.549	0.481	0.490	0.526	0.487	0.506
JFLS5	0.587	0.569	0.608	0.591	1.000	0.747	0.466	0.544	0.708	0.649	0.666	0.490	0.670	0.607	0.724	0.765	0.762	0.608	0.672	0.589
JFLS6	0.500	0.552	0.495	0.500	0.747	1.000	0.427	0.506	0.641	0.644	0.605	0.479	0.636	0.700	0.729	0.676	0.660	0.603	0.624	0.498
JFLS7	0.777	0.781	0.760	0.803	0.466	0.427	1.000	0.883	0.585	0.543	0.582	0.731	0.561	0.607	0.601	0.550	0.536	0.404	0.539	0.593
JFLS8	0.782	0.823	0.757	0.804	0.544	0.506	0.883	1.000	0.670	0.592	0.590	0.785	0.589	0.613	0.619	0.569	0.563	0.452	0.553	0.505
JFLS9	0.613	0.687	0.594	0.624	0.708	0.641	0.585	0.670	1.000	0.879	0.719	0.625	0.827	0.776	0.819	0.811	0.805	0.725	0.733	0.660
JFLS10	0.529	0.629	0.523	0.586	0.649	0.644	0.543	0.592	0.879	1.000	0.715	0.675	0.885	0.804	0.824	0.781	0.774	0.804	0.722	0.651
JFLS11	0.552	0.544	0.543	0.570	0.666	0.605	0.582	0.590	0.719	0.715	1.000	0.584	0.800	0.728	0.880	0.771	0.791	0.582	0.804	0.769
JFLS12	0.648	0.690	0.663	0.685	0.490	0.479	0.731	0.785	0.625	0.675	0.584	1.000	0.733	0.665	0.670	0.590	0.592	0.529	0.605	0.594
JFLS13	0.564	0.587	0.566	0.606	0.670	0.636	0.561	0.589	0.827	0.885	0.802	0.733	1.000	0.838	0.906	0.775	0.777	0.738	0.836	0.750
JFLS14	0.502	0.604	0.437	0.508	0.607	0.700	0.607	0.613	0.776	0.804	0.728	0.665	0.838	1.000	0.892	0.857	0.834	0.712	0.745	0.662
JFLS15	0.538	0.565	0.505	0.549	0.724	0.729	0.601	0.619	0.819	0.824	0.880	0.670	0.906	0.892	1.000	0.898	0.885	0.699	0.889	0.805
JFLS16	0.510	0.594	0.465	0.481	0.765	0.676	0.550	0.569	0.811	0.781	0.771	0.590	0.775	0.857	0.898	1.000	0.975	0.728	0.788	0.700
JFLS17	0.507	0.585	0.461	0.490	0.762	0.660	0.536	0.563	0.805	0.774	0.791	0.592	0.777	0.834	0.885	0.975	1.000	0.733	0.775	0.697
JFLS18	0.424	0.551	0.416	0.526	0.608	0.603	0.404	0.452	0.725	0.804	0.582	0.529	0.738	0.712	0.699	0.728	0.733	1.00	0.669	0.560
JFLS19	0.495	0.484	0.481	0.487	0.672	0.624	0.539	0.563	0.733	0.722	0.804	0.605	0.836	0.745	0.889	0.788	0.775	0.669	1.000	0.874
JFLS20	0.508	0.461	0.546	0.506	0.589	0.496	0.593	0.505	0.660	0.651	0.769	0.594	0.750	0.661	0.805	0.707	0.697	0.560	0.874	1.000

4.3.2 Test- retest Reliability of the Malay GCPS and Malay JFLS

For the test-retest reliability, we have randomly selected a total of 40 subjects who consisted of 20 non-TMD subjects and 20 TMD patients. The 40 subjects were asked to answer the same questionnaire 2 weeks after the first attempt. The Intraclass Correlations (ICC/) of the two domains are shown in **Table 4.5** (The Malay GCPS) and **Table 4.6** (The Malay JFLS). The 95% confidence interval of the mean was also calculated.

Table 4.5: Test-retest Reliability of the Malay GCPS (Intraclass Correlations)

GCPS Subdomain Test-retest (ICC) (n = 40)	Single measures	95 % Confidence Interval
• **Total score**	0.98	0.96 – 0.99
'Characteristics pain intensity' subdomain	0.99	0.98 – 0.99
'Interference of function' subdomain	0.99	0.98 - 0.99
'Days with interference' subdomain	0.99	0.98 – 0.99

Based on the table above, the ICC of the total score for the domain of the Malay GCPS was 0.98 with its 95% confidence interval ranging from 0.96 to 0.99 which indicates a superior agreement. For the ICC value, the highest measure was 0.99 for 'Characteristic Pain Intensity' subdomain while the lowest ICC was 0.99 for 'Days with Interference' subdomain. In general, an excellent level of intraclass correlations were demonstrated in the Malay GCPS as all subdomain values were above 0.90.

Table 4.6: Test-retest Reliability of the Malay JFLS (Intraclass Correlations)

JFLS Subdomain Test-retest (ICC) (n = 40)	Single measures	95 % Confidence Interval (CI)
• **Total score**	0.99	0.98 – 0.99
'Mastication' subdomain	0.99	0.98 – 0.99
'Vertical mobility' subdomain	0.98	0.97 - 0.99
'Verbal & Non-verbal communication' subdomain	0.99	0.99 - 1.00

The pattern of excellent agreement was also shown by the Malay JFLS domain. From the Malay JFLS ICC table **(Table 4.6)**, it was clearly demonstrated that the ICC

of the total score for the Malay JFLS was 0.99 with its 95% confidence interval ranging from 0.98-0.99. The highest ICC value among the items was 'Verbal & Non-verbal Communication' subdomain (0.99, 95% CI = 0.99 – 1.00) while the lowest ICC was 'Vertical Mobility' subdomain which was 0.98 with its 95% CI value ranging from 0.97-0.99.

4.3.3 Concurrent Validity of the Malay GCPS and Malay JFLS

Concurrent validity of the Malay GCPS and M-BPI was done based on the similar subdomain which are the 'Characteristic Pain Intensity' subdomain and 'Interference of Function' subdomain. On the other hand, concurrent validity for the Malay JFLS and M-OHIP14 was based on all subdomains.

Table 4.7: Concurrent Validity (Spearman's rho correlation test) of the Malay GCPS and Malay JFLS

Concurrent Validity	Correlation coefficient (μ)	p-value
GCPS/M-BPI		
• **'Characteristics Pain Intensity' subdomain**	0.56	$p < 0.01$
• **'Interference of Function' subdomain**	0.60	
JFLS/M-OHIP14		
• **JFLS-Malay OHIP14 all subdomains**	0.56	$p < 0.01$

According to **Table 4.7**, both M-DC/TMD domains (the Malay GCPS and Malay JFLS) had good positive correlations coefficient in the Spearman's Rho Correlation test. The Malay GCPS subdomain correlates positively with M-BPI subdomain in measuring 'Characteristic Pain Intensity' ($\mu = 0.56$, $p<0.01$) and 'Interference of Function' ($\mu = 0.60$, $p<0.01$). Both pain subdomains had a statistically significant correlation. A good positive correlation was also obtained between the Malay JFLS and M-OHIP ($\mu = 0.56$, $p<0.001$).

4.3.4 Construct Validity of the Malay GCPS and Malay JFLS

The construct validity is presented in **Table 4.8**. Based on the table below, fair positive associations were seen between the Malay GCPS and OHIP 'physical pain' subdomain (0.37), 'physical disability' subdomain (0.34) and 'psychological disability' subdomain (0.34). However, there were negative associations reported between the Malay GCPS and OHIP 'psychological discomfort' subdomain (-0.14, poor negative association), 'social disability' subdomain (-0.40, moderate negative association) and 'handicap' subdomain (-0.37, fair negative association). Thus, with those 3 negative associations, the correlations hypotheses for these 3 subdomains were rejected.

For the association of the Malay JFLS, a good correlation in the hypothesized direction was observed with OHIP 'functional limitations' subdomain which scored 0.48. In general, all items were statistically, significantly correlated.

Table 4.8: Construct validity –Association between the Malay GCPS & Malay JFLS with Malay OHIP14 (Spearman's rho correlation test)

Association	n	Correlation coefficient (μ)	p-value
Malay GCPS			
• OHIP 'physical pain' subdomain		0.37	
• OHIP 'psychological discomfort' subdomain		-0.14	
• OHIP 'physical disability' subdomain		0.34	p<0.01
• OHIP 'psychological disability' subdomain		0.34	
• OHIP 'social disability' subdomain		-0.40	
• OHIP 'handicap' subdomain		-0.37	
Malay JFLS			
• OHIP 'functional limitations' subdomain	200	0.48	p<0.01

A Kruskal-Wallis test was conducted for the association of the Global Oral Health questions as this question had more than 2 outcomes as shown by **Table 4.9.** Based on the table, the p-value of the test was $p < 0.01$. This provides evidence of construct

validity. The mean scores for both domains (the Malay GCPS and Malay JFLS) increased gradually as the subjects perceived a poor and dissatisfied global oral health rating. Those subjects who perceived need for oral/jaw treatment, had significantly higher Malay GCPS and Malay JFLS scores respectively.

Table 4.9: Construct validity –Association between the Malay GCPS & Malay JFLS with Global Oral Health ratings (Kruskall-Wallis test)

Global Oral Health		Malay GCPS Mean score (S.D)	Malay JFLS Mean score (S.D)	p-value
• **Perceived oral/jaw health status**	Excellent	0.18 (0.39)	0.63 (0.98)	p<0.01
	Good	0.29 (0.49)	0.65 (1.41)	
	Fair	0.77 (0.86)	1.53 (1.82)	
	Poor	1.59 (1.24)	3.20 (2.73)	
	Very poor	3.25 (0.50)	5.70 (1.42)	
• **Perceived satisfaction with oral/jaw health**	Very Satisfied	0.32 (0.53)	0.73 (1.45)	p<0.01
	Satisfied	0.38 (0.61)	0.75 (1.48)	
	Moderate	1.27 (1.10)	2.66(2.31)	
	Dissatisfied	2.00 (0.94)	3.65 (1.83)	
	Very dissatisfied	2.50 (2.12)	4.13 (3.39)	
• **Perceived need for oral/jaw treatment**	Yes	1.36 (1.28)	2.77 (2.57)	p<0.01
	No	0.35 (0.58)	0.85 (1.68)	
	Don't know	0.90 (0.87)	1.49 (1.58)	

Lastly, the association between the Malay JFLS and limited mouth opening was assessed via Man-Whitney U test which is presented in **Table 4.10**. It can be concluded that subjects could not open their mouths wide enough, had a significantly higher Malay JFLS median score (Median=2.22, IQR=3.43, p<0.01) compared with those who could open their mouths wide enough (median=0.11, IQR=1.04, p<0.01). There was evident difference in mean rank for both groups with a statistical significance level ($p < 0.01$).

Table 4.10: Construct validity – Association between the Malay JFLS and Limited Mouth Opening (Mann-Whitney U test)

Association	Can you open wide enough? (n)	Mann-Whitney U	Median (Interquartile Range, IQR)	Mean rank	*p* value
Malay JFLS (n=200)	Yes (n=141)	2632.00	0.11 (1.04)	89.67	$p < 0.01$
	No (n=59)		2.22 (3.43)	126.39	

4.3.5 Discriminant Validity of the Malay GCPS and Malay JFLS

Man-Whitney U test was used for Discriminant Validity to differentiate between non-TMD subjects and TMD patients. According to **Table 4.11**, TMD patients had a significantly higher Malay GCPS median score (Median=3.83, IQR=4.33, p<0.01) compared to non-TMD subjects (median=0.05, IQR=0.17, p<0.01). By the same token, TMD patients had also showed a significantly higher Malay JFLS median score (Median=2.72, IQR=3.52, p<0.01) compared to non-TMD subjects who had a lower median score (Median=0.05, IQR=0.60, p<0.01). There was also evidence of difference in mean rank for both groups with a statistical significance level.

Table 4.11: Discriminant Validity of Malay-GCPS and Malay-JFLS (Mann-Whitney U test)

Domain (n)	Group (n)	Mann - Whitney U	Median (Interquartile Range)	Mean rank	*p* value
Malay GCPS (n=248)	**Non TMD** (n=162)	1983.00	0.05 (0.17)	93.74	$p <$ 0.01
	TMD (n=86)		3.83 (4.33)	181.67	
Malay JFLS (n=200)	**Non TMD** (n=134)	1466.00	0.05 (0.60) (0.60) (3.53)	78.44	$p <$ 0.01
	TMD (n=66)		2.72 (3.52)	145.29	

CHAPTER 5: DISCUSSION

5.1 Translation of the English DC/TMD

The aim of the study is to develop a Malay version of DC/TMD through a formal cross-cultural adaptation process. The translation of the English version of DC/TMD to the Malay language was done according to the 'Translation and Adaptation of the DC/TMD Protocol' (International RDC/TMD Consortium et al.,2014) and the 'Guidelines for Establishing Cultural Equivalency of Instruments' by Ohrbach et al. (2013). This accepted and standard guidelines are based on the DC/TMD consortium protocol. In this study, the translation and cross-cultural adaptation of DC/TMD was done for the usage among Malaysians who are a Malay-speaking population (Khoo et al., 2008). Along the translation process, some items were rephrased or added from the source instrument in order to make the content culturally sensitive to the Malaysian population (Saub et al.,2005).

Most of the forward and back-translations of DC/TMD from the source language to the Malay version was straightforward and simple to translate due to their unambiguous meanings. However, there were some terms that faced difficulties and confusion among translators. One of the terms was 'temple'. The word 'temple' had some predicaments and was ambivalent among translators as there is no clear word in Malay to describe the 'temple' region. Initially, the word 'temple' was translated as 'side of the forehead – left or right' (*tepi dahi sebelah kanan atau kiri*). Nevertheless, during the pre-test session of the translated version, it was still causing confusion among the respondents. Thus, the final consensus was made and it was decided not to translate the word 'temple'. Alternatively, an asterisk mark (*) was placed next to the word 'temple' in the Malay version to denote that particular term will be explained directly by

the examiner to the respondents prior to answering the questionnaire. The consequence from this agreement, a better response and outcome was obtained.

Other than that, a cross-cultural adaptation of the original DC/TMD was also done based on the local foods. For instance, in the Jaw Functional Limitations Scale (JFLS) domain, there were a few items that used western foods as their example in the questionnaire such as macaroni which was replaced with *Kuey Teow* (a type of Malaysian noodle), while pureed food was replaced with *bubur kanji* (starch porridge) which has the same consistency and structure as the original example. The item 'prepared in oven' was replaced with a more precise example (roasted chicken) for a deeper and better understanding among respondents. Besides foods, changes also have been made for the example of musical instruments that involved either mouth or jaw in the 'Oral Behaviour Checklist' (OBC) domain. The examples of musical instruments were adapted based on common instruments among Malaysians. The woodwind instrument was replaced with saxophone; brass instrument was replaced with trumpet while violin was chosen as an example of a string instrument.

Based on Ohrbach et al., (2013), a preliminary testing of initial translated instruments can be administrated to a small group of subjects to access the comprehensibility, practicability and recognition of the instrument's item. A clinical sample testing was done that involved 10 bilingual subjects. Following answering the questionnaires of the translated instrument, an interview session was also conducted. The purpose of this interview session was to highlight ambiguities in item construction, response options and other possible choices of translation in any items that might cause confusion among respondents. Following pre-test, any irrelevant and errors of the items should be identified and further revised by the translators so that the item objectives are met.

The adaptation of the population and demographic items such as socioeconomic status, education level, ethnicity as well as personal monthly income were carried out according to the Department of Statistics Malaysia as well as the National Oral Health Survey of Adults (NOHSA 2010). It was reported by John et al. in 2006, that a German group had recruited their sociodemographic characteristics from the German national dental and medical surveys for their study. In general, the process of translation went on smoothly without any major problems.

5.2 Psychometric Properties

A culturally adapted instrument could modify the reliability and validity of the instrument (Khoo et al., 2008). Thus, Beaten DT, Bombardier, Guillemin, & MB (2000), recommended that the new translated version from source instrument that has undergone translation and adaptation processes should establish the same measurement properties needed for the designed utilization which can be achieved via testing the internal consistency, reliability and validity. In this study; the internal consistency, reliability and test-retest reliability were measured via Cronbach's alpha and Interclass Correlation Coefficient (ICC) while the validity test was done by means of concurrent, construct and discriminant validity (Khoo et al.,2008; John et al., 2006; He & Wang, 2015). To test the psychometric properties, a total number of 252 subjects were involved. Based on Terwee et al., (2007), a minimum of 100 subjects were required to ensure the stability of the items.

5.2.1 Internal Consistency Reliability

Reliability is a substantial and valuable property in patient-related Quality of Life outcome measure. There are two key features in measuring reliability – internal consistency and test-retest reliability (He & Wang, 2015). In this study, two domains of DC/TMD were assessed for its reliability and validity of the translated version which are GCPS and JFLS. The internal consistency is evaluated by Cronbach's Alpha test, in which the alpha value displays the degree of correlation of an item with a scale and an individual item with itself in the same domain (Lucena et al., 2006). According to Bland & Altman (1997), the lowest universal alpha value that is considered relevant to clinical studies is $\alpha = 0.90$, however, for scales items, α value as low as 0.70 can be accepted.

In this study, the Cronbach's alpha value for both domains were more than 0.90. The Malay GCPS α value was 0.95 while JFLS scored 0.97. With these values, it is clearly shown that both domains demonstrated high internal consistencies comparable to the older version of the Malay-RDC/TMD (Khoo et al.,2008) and the German-RDC/TMD (John et al., 2006) which α value ranged between 0.72 to 0.88. All item-total correlation values for both domains scored alpha value more than 0.70. However, there was only one item-total correlation that measured less than 0.70 which was the 'interference days' subdomain ($\alpha = 0.56$). Streiner & Streiner (2003), stated that some researchers acknowledged value less than 0.70 but close to 0.60 as satisfactory. The low value for this subdomain is due to the response option. The response to this question was not scored by scale compared to other subdomains. Rather, the subjects' responses were only interpreted based on the number of days that caused pain for the past 1 month. The response might produce bias answers as subjects will not respond in accurate numbers of days with pain.

5.2.2 Test - Retest Reliability

The test-retest reliability was presented by ICC values which indicated the agreement of continuous data (Vetter & Schober,2018). Based on this publication, the test-retest reliability can be measured when the subjects of a particular study repeatedly answer the same instrument in a period of time. The ICC can be used to measure the intra-rate reliability when the scale measures at two different periods of time which in this case, was 2 weeks. The gap of 2 weeks was planned to reduce remembrance recall bias as result from the first test (Khoo et al., 2008). However, the perfect time for retesting is still questionable as there is still a probability that subjects might recall the current item's feedback if the test-retest gap is too close (Aisyaturridha et al., 2006).

The ICC value will be close to 1 when the different measures of a quantity is equal and comparable to each item, while the ICC is expected to be lower and possibly approach zero when there is little agreement between the items (Vetter & Schober, 2018). In this Malay-DC/TMD version, the ICC values of the total score of GCPS and JFLS were close to 1 (0.98 and 0.99 respectively). The obtained ICC value for GCPS was slightly higher than the German version of RDC/TMD (John et al., 2008) which was 0.97. These scores indicate an excellent mark of the test-retest reliability. In general, all items for the Malay GCPS and Malay JFLS established an ICC value of more than 0.90. A positive rating for reliability is given when the ICC value is at least 0.70 in a sample size population consisting of at least 50 subjects (Terwee et al.,2007).

5.2.3 Validity

Validity is another crucial psychometric property of a questionnaire (He & Wang, 2015). To evaluate the validity of the psychometric property in the Malay-DC/TMD, 3 types of validity tests have been conducted which were concurrent validity, construct validity and discriminant validity.

Concurrent validity refers to scores on a distinct instrument that correlate to a regular approved instrument which refers to a gold standard tool (Terwee et al.,2007). Commonly, the concurrent validity might be a challenge for the researchers as this type of validity requires a 'gold standard' tool which is not easily found in all knowledge fields (Souza et al.,2017). In this study, the concurrent validity is tested between the Malay GCPS and Malay version of Brief Pain Inventory that measures pain, while the Malay JFLS was concurrently validated against the Malay-OHIP14 in which both tools measure the jaw function and oral health. The association of the related domains were assessed by the Spearman's rho correlation test. Correlation can be described when an alteration in the degree of a variable is related with a change of a degree of another variable (Vetter & Schober,2018), either in the same direction (positive correlation) or in the contra direction (negative correlation). A correlation that is not normally distributed in a continuous data or in an ordinal data, a Spearman rank correlation test can be used to measure the association (Vetter & Schober,2018).

In the translation of the German-RDC/TMD by John et al., (2006), the evidence of the constructed dysfunctional chronic pain item was associated with the German version of Multidimensional Pain Inventory (MPI-D). Based on this documentation, we have decided to use the Malay version of The Brief Pain Inventory (M-BPI) (Aisyaturridha et al., 2006) to be correlated with the Malay GCPS via concurrent

validity. The BPI was established to gain estimation of pain prevalence and to determine the severity as well as pain function interference. There were a few studies that used BPI to assess pain among the TMD patients (Park, Kim, Kim, & Kim, 2015). In addition to the pain severity, the 'BPI Pain Interference' was calculated which includes seven items that evaluate the impact of pain on patient's general activity, mood, walking ability, normal work, relations with others, sleep and enjoyment of life. Literally, the walking ability should be replaced with chewing ability since TMD is not associated with any walking inference (Park et al., 2015). However, in this study, there was no change made in the Malay-BPI item for walking ability.

For the Malay JFLS domain, we have decided to validate the domain with Malay-OHIP14. John et al (2006) in his publication documented that the Jaw Disability List (JDL) of RDC/TMD was associated with the German version of Oral Health Impact Profile (OHIP-G). The Malay-OHIP14 was developed by Saub et al. (2007) which consists of 14 domains with 7 items i.e. functional limitation, physical pain, psychological discomfort, physical disability, psychological disability, social disability and handicap. The Malay-OHIP14 was calculated by adding up the scores of each item which ranges from 0 to 56. Higher scores indicate poorer OHRQoL. Thus, with all of these evidences of related association, a Spearman's Rho correlation test was conducted for concurrent validity. The psychometric properties of Malay GCPS and Malay JFLS demonstrated a statistically significant positive good correlation. This evidence supports the concurrent validity of the GCPS and JFLS for the Malay version of DC/TMD by showing a moderate to good concurrent validity. A positive rating correlation is given to a correlation if the correlation test with 'gold standard' is at least 0.70 (Terwee et al, 2007).

The next validity test is construct validity which is a type of validity that usually associates the hypothesis testing (Souza et al.,2017). Based on Terwee et al (2007), construct validity property refers to a state when scores on a specific questionnaire persistently correlate with the constructed hypothesis theoretically pertaining to the measured concepts. In this study, a total of fourteen specific hypotheses were formulated for the Malay GCPS (9 hypotheses) and Malay JFLS (5 hypotheses).

Based on the obtained results, 11 of the constructed hypotheses were in accordance with the established results except for 3 hypotheses. The 3 associations that were found negative were the association of the Malay GCPS with other 3 subdomains of M-OHIP which were 'psychological discomfort', 'social disability' and 'handicap'. Those 3 hypotheses were rejected due to their negative association.

For the 'psychological discomfort" subdomain, one of the items (item no 5) was based on the discomfort due to foods stuck in between teeth/ dentures. This unspecific question that was not related to any jaw discomfort could have contributed to the negative association. Other than that, negative associations were also seen for the 'social disability' and the 'handicap' subdomains. These conditions could be due to the subjects' responses which were more related to the teeth and mouth rather than their jaws/ TMJ status as the questions were generalized to teeth, mouth and jaw in a single item. The responses could be different and more focused to TMJ if the OHIP-TMD were used in this validity. However, up to this moment, there is no Malay OHIP-TMD that has been developed and validated for a more accurate response among subjects.

Based on Terwee et al., (2007), a positive rating of construct validity was considered if the formulated hypotheses were specific and a minimum of 75% of the results are in equivalence with the specified hypotheses. In this study, out of 14 definite

hypotheses, 11 formulated hypotheses were in accordance with the results obtained which concluded that 79% of the results were in correspondence with the itemized hypotheses. These findings provide evidence to support construct validity of the Malay DC/TMD.

Lastly, the validity of the Malay version of DC/TMD was assessed via discriminant validity. The aim of this validity is to differentiate the TMD patients from asymptomatic pain-free individuals which were the control group. The discriminant validity of the Malay-DC/TMD was assessed based on the existence of the TMD symptoms which were estimated by its capability to discriminate between symptomatic and asymptomatic subjects. In this study, TMD patients manifested significantly higher Malay GCPS and Malay JFLS values. This higher value of GCPS in TMD patients were in concurrence and comparable with the previous study (Khoo et al., 2008) which was also determined via discriminant validity. Overall, the Malay DC/TMD was reliable and offered to be a valid instrument for the TMD assessment.

5.3 Limitations

Questionnaires survey is one of the cheapest and easiest methods to gather quantitative data. In general, there are multiple weaknesses of questionnaires. One of the biggest weaknesses is differences in interpretation and perception among respondents since this is a self-administered questionnaire. Without someone who is an expert to the questionnaire, respondents may have difficulties in understanding the constructed questions that may seem clear to the creators or investigators. This misleading communication and understanding can lead to skewed results. A face-to-face interview session should have been done in order to yield better and honest responses.

Other than that, this study should have been carried out in a longer time lapse so that more samples could have been recruited/collected to include more than the 3 main races in Malaysia namely the Malays, Chinese and Indians. Besides, the involvement of Malaysians from East Malaysia such as the people from Sabah and Sarawak should also be included and recruited in this study to portray more a genuine and true Malaysian population who is majority a Malay-speaking population. Further studies should consider a more general sample of population in order to ratify the generalizability of the data.

In terms of data collection for the construct validity test, a specific tool should have been used to validate the Malay JFLS. A specific tool for the TMD measurement which is the OHIP-TMD should be used instead of the general oral health OHIP. However, up to this moment, there is no validated version of the Malay OHIP-TMD. A good and strong association is expected for the Malay JFLS if OHIP-TMD were to be used for construct validity of the Malay DC/TMD.

5.4 Future Study

Here are a few recommendations for future studies:

1. To obtain more variants in gender, age, level of education and race groups especially the population in Sabah and Sarawak.
2. To use a validated Malay-Oral Health Impact Profile for Temporomandibular Disorders, M-OHIP TMD (if any in the future) which is a specific tool for assessing TMD that can be used for evaluating psychometric properties of M-DC/TMD for a better validation test especially in the construct validity.
3. For more reliable clinical trials results, it is recommended, that a series of TMD treatments are given to patients in order to evaluate the outcome measurement so that the items will demonstrate changes over time in subjects' responses as such change should be mirrored in the sensitivity of pre and post treatment value.
4. A more complete psychometric evaluation should have been conducted to include other DC/TMD domains such as TMD Pain Screener, Symptom Questionnaire, Pain Drawing, Patient Health Questionnaire (PHQ), Generalized Anxiety Disorder (GAD), Oral Behaviors Checklist (OBC).
5. To conduct a multinational trial in the future and to compare the study results.
6. To validate other domains of the Malay DC/TMD.
7. Lastly, to validate the cross-cultural English version of the DC/TMD for the usage among Malaysians who are also an English-speaking population.

CHAPTER 6: CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

Based on the methods conducted and results obtained, we can conclude that;

i. The English DC/TMD has been successfully translated into the Malay language and culturally adapted among Malaysians.

ii. The series of actions taken to establish the validity and reliability of the Malay version of the DC/TMD were based on the methodology recommended in the literatures and the established guidelines. The Malay DC/TMD is a reliable instrument to evaluate TMD as it has good internal consistency and excellent test-retest reliability. It is also a valid instrument as it demonstrated good concurrent and construct validity.

Overall, this study provides empirical evidence for the Malay version of Diagnostic Criteria Temporomandibular Disorders (M-DC/TMD) as a valid and reliable instrument for the evaluation of temporomandibular disorders in the Malaysian setting.

6.2 Recommendations

The application of M-DC/TMD is straightforward and brief to complete for assessing the physical and psychosocial status of TMD. This measurement will be useful for the implementation in clinical and research settings by providing methods to the researcher for valid phenotyping of their subjects especially for pain-related TMD. M-DC/TMD will also provide a common language for all clinicians and researchers regarding TMD in Malaysia.

The evidence-based data collected from this validated measurement could produce a full range of diagnostic actions from screening to definitive assessments and diagnoses which are very important for better patient care and systematic management of TMD in the future. The M-DC/TMD could also be used for justification of treatment success.

REFERENCES

Abrahamsson, C., Henrikson, T., Nilner, M., Sunzel, B., Bondemark, L., & Ekberg, E. (2013). TMD before and after correction of dentofacial deformities by orthodontic and orthognathic treatment. *International Journal of Oral and Maxillofacial Surgery*, *42*(6), 752–758.

Aisyaturridha, A., Naing, L., & Nizar, A. J. (2006). Validation of the malay brief pain inventory questionnaire to measure cancer pain. *Journal of Pain and Symptom Management*, *31*(1), 13–21.

Bartko, J.J.The Intraclass Correlation Coefficient As A Measure Of Reliability Reports, 1966, 19, 3-11.

Beaten DT, Bombardier, C., Guillemin, F., & MB, F. (2000). Guidelines for the process of Cross Cultural adaptation of Self Report mesures. *Spine*, *25*(24), 3186–3191.

Bland, J. M., & Altman, D. G. (1997). Cronbach's alpha. *BMJ (Clinical Research Ed.)*.

Bolarinwa, O. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, *22*(4), 195.

Bölenius, K., Brulin, C., Grankvist, K., Lindkvist, M., Söderberg, J. (2012) A content validated questionnaire for assessment of self-reported venous blood sampling practices. BMC Res Notes;5:39.

Brislin, R.W., Lonner, W.J., Thorndike, R.M. (1973) Questionnaire wording and translation. In: Cross-cultural Reaeareh Methods. New York: John Wiley: 32-58.

Bullinger, M., Anderson, R., Cella, D., & Aaronson, N. (1993). Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Quality of Life Research*, *2*(6), 451–459..

Cleeland, C.S.The Brief Pain Inventory Manual.User Guide© 2009

Dubner, R., Ohrbach, R., & Dworkin, S. F. (2016). The Evolution of TMD Diagnosis: : Past, Present, Future. *Journal of Dental Research*, *95*(10), 1093–1101.

Dworkin, S., & LeResche, L. (1992). *Research Diagnostic Criteria For Temporomandibular Disorders: Review, Criteria, Examinations And Specifications, Critique*. *Journal of craniomandibular disorders : facial & oral pain* (Vol. 6).

Feinstein, A.R. Clinical judgment. Baltimore: Williams & Wilkins, 1967:414.

Feise, R.J., Menke J.M. Functional Rating Index: A New Valid and Reliable Instrument to Measure the Magnitude of Clinical Change in Spinal Conditions Spine. 26(1):78-87, JAN 2001

Fillingim, R. B., Ohrbach, R., Greenspan, J. D., Knott, C., Bair, E., Baraian, C., … Maixner, W. (2012). NIH Public Access, *12*, 1–25..

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286–299.

González, Y., Miranda-Rivera Y, Espinosa I. Cross-cultural adaptation of research diagnostic criteria for temporomandibular disorders (RDC/TMD). Rev Fac Odontol Univ Antioq 2013; 25(1): 11-25.

Gonzalez, Y., Schiffman, E.L., Gordon, S.M., et al. Development of a brief and effective TMD-pain screening questionnaire: Reliability and validity. J Am Dent Assoc 2011;24:1183–1191.

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J.Clin.Epidemiol.*, *46*(12), 1417–1432.

He, S., & Wang, J. (2015). Reliability and validity of a Chinese version of the Oral Health Impact Profile for edentulous subjects. *Quality of Life Research*, *24*(4), 1011–1016.

He, S. L., Wang, J. H., Wang, M. H., & Deng, Y. M. (2012). Validation of the Chinese version of the Halitosis Associated Life-quality Test (HALT) questionnaire. *Oral Diseases*, *18*(7), 707–712.

Jang, J.-Y., Kwon, J.-S., Lee, D. H., Bae, J.-H., & Kim, S. T. (2016). Clinical signs and subjective symptoms of temporOmandibular disorders in instrumentalists. *Yonsei Medical Journal*, *57*(6), 1500–1507.

John, M. T., Hirsch, C., Reiber, T., & Dworkin, S. F. (2006). Translating the Research Diagnostic Criteria for Temporomandibular Disorders into German: Evaluation of content and process, *20*(1), 43–52.

Jung, H.-D., Kim, S. Y., Park, H.-S., & Jung, Y.-S. (2015). Orthognathic surgery and temporomandibular joint symptoms. *Maxillofacial Plastic and Reconstructive Surgery*, *37*(1), 14. .

Khoo, S., U Jin Yap, A., Chan, Y., & Bulgiba, A. (2008). *Translating the Research Diagnostic Criteria for Temporomandibular Disorders into Malay: evaluation of content and process*. *Journal of orofacial pain* (Vol. 22).

Kim, Y. K., Kim, S. G., Im, J. H., & Yun, P. Y. (2012). Clinical survey of the patients with temporomandibular joint disorders, using Research Diagnostic Criteria (Axis II) for TMD: Preliminary study. *Journal of Cranio-Maxillofacial Surgery*, *40*(4), 366–372.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, *65*(23), 2276–2284.

Kroenke, K., Spitzer, R. L., Janet, B. W., Williams, D. S. W., & Lö, B. (2009). An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ– 4. *Psychosomatics*, *50*, 613–621.

Last, J.M. A Dictionary of Epidemiology, New York : Oxford University Press. 1995

Lei, J., Liu, M.-Q., Yap, A. U. J., & Fu, K.-Y. (2015). Sleep Disturbance and Psychologic Distress: Prevalence and Risk Indicators for Temporomandibular Disorders in a Chinese Population. *Journal of Oral & Facial Pain and Headache*, *29*(1), 24–30.

Look, J.O., John, M.T., Tai, F., et al. The Research Diagnostic Criteria for Temporomandibular Disorders. II: Reliability of Axis I diagnoses and selected clinical measures. J Orofac Pain 2010; 24:25–34.

Lucena, L. B. S. de, Kosminsky, M., Costa, L. J. da, & Góes, P. S. A. de. (2006). Validation of the Portuguese version of the RDC/TMD Axis II questionnaire. *Brazilian Oral Research*, *20*(4), 312–317.

Manfredini, D., Guarda-Nardini, L., Winocur, E., Piccotti, F., Ahlberg, J., & Lobbezoo, F. (2011). Research diagnostic criteria for temporomandibular disorders: A systematic review of axis i epidemiologic findings. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology*, *112*(4), 453–462.

National Institute of Dental and Craniofacial Research. Facial Pain. http://www.nidcr.nih.gov/DataStatistics/FindDataByTopic/ FacialPain/ (accessed 7/28/2013).

Ohrbach, R. (2008). The Jaw Functional Limitation Scale: development, reliability, and validity of 8- item and 20-item versions. *Journal of Orofacial Pain*, *22*(3), 219–230.

Ohrbach, R., Bjorner, J., Metric, Q., Jezewski, M., John, M. T., & Lobbezoo, F. (2013). Guidelines for Estabilishing Cultural Equivalency of Instruments. *University at Buffalo*.

Ohrbach, R., Schiffman, E., Ohrbach, R., Truelove, E., Look, J., … Dworkin, S. F. (2014). Diagnostic Criteria for Temporomandibular Disorders Clinical Protocol and Assessment Instruments DC / TMD : Axis I Screening. *Journal of Oral & Facial Pain and Headache*, *28*(version 2), 1–122.

Okeson, J. P. (1997). Current Diagnostic Classification Schema And Assessment Of Patients With Temporomandibular Disorders Current terminology and diagnostic classification schemes. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*, *83*, 61–64.

Park, J-H., Kim, H-K., Kim, K-S., Kim, M-E.Pain Catastrophizing for Patients with Temporomandibular Disorders.J Oral Med Pain 2015;40(2):47-54.

Sangoseni, O., Hellman, M., Hill, C. Development and validation of a questionnaire to assess the effect of online learning on behaviors, attitude and clinical practices of physical therapists in United States regarding of evidence-based practice. Internet J Allied Health Sci Pract 2013;11:1-12.

Saub, R., Locker, D., Allison, P., & Disman, M. (2007). Cross-cultural adaptation of the Oral Health Impact Profile (OHIP) for the Malaysian adult population. *Community Dental Health*.

Schellhas, K. P., Pollei, S. R., & Wilkes, C. H. (1993). Pediatric internal derangements of the temporomandibular joint: Effect on facial development. *American Journal of Orthodontics and Dentofacial Orthopedics*, *104*(1), 51–59..

Schiffman, E., Ohrbach, R., Truelove, E., Look, J., Anderson, G., Goulet, J.-P., … Dworkin, S. F. (2014). Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: Recommendations of the International RDC/TMD Consortium Network* and Orofacial Pain Special Interest Group†. *Journal of Oral & Facial Pain and Headache*, *28*(1), 6–27.

Smith, M. T., Wickwire, E. M., Grace, E. G., Edwards, R. R., Buenaver, L. F., Peterson, S., … Haythornthwaite, J. A. (2009). Sleep disorders and their association with laboratory pain sensitivity in temporomandibular joint disorder. *Sleep*, *32*(6), 779–790.

Souza, A. C. de, Alexandre, N. M. C., Guirardello, E. de B., Souza, A. C. de, Alexandre, N. M. C., & Guirardello, E. de B. (2017). Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. *Epidemiologia e Serviços de Saúde*, *26*(3), 649–659.

Streiner, D.L., Norman G.R. Health measurement scales. A practical guide to their development and use. New York: Oxford University Press; 2003.

Streiner, D. L., & Streiner, D. L. (2003). Starting at the Beginning : An Introduction to Coefficient Alpha and Internal Consistency Starting at the Beginning : An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, *80*(1), 99–103.

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., … de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42.

U J Yap, A., Dworkin, S., Chua, E. K., List, T., Tan, K., & Tan, H. H. (2003). *Prevalence of temporomandibular disorder subtypes, psychologic distress, and psychosocial dysfunction in Asian patients*. *Journal of orofacial pain* (Vol. 17).

U J Yap, A., K Chua, E., Dworkin, S., Tan, H. H., & Tan, K. (2002). *Multiple Pains and Psychosocial Functioning/Psychologic Distress in TMD Patients*. *The International journal of prosthodontics* (Vol. 15).

Vetter, T. R., & Schober, P. (2018). Agreement Analysis: What He Said, She Said Versus You Said. *Anesthesia and Analgesia*, *XXX*(Xxx), 1–6.

Von Korff, M., Ormel, J., Keefe, F. J., & Dworkin, S. F. (1992). Grading the severity of chronic pain. *Pain*, *50*(1092), 133–149.

Wadhwa, S., Kapila, S. TMJ disorders: future innovations in diagnostics and therapeutics. J Dent Educ 2008;72:930-47.

WHO, Headache Classification Subcommittee of the International Headache Society. The International Classification of Headache Disorders. Cephalalgia (2nd edn). 2004; 24(Suppl 1):9–160.

Wright, E.F. Manual of Temporomandibular Disorders. USA: Blackwell Musnksgaard, 60e71, 2005