COMPARABILITY OF A COMPUTER TOUCH SCREEN VERSUS WRITTEN METHOD FOR COLLECTING ORAL HEALTH RELATED QUALITY OF LIFE (OHRQOL) DATA

MAZNURFARHATUNNISAK BINTI ANOWAR

FACULTY OF DENTISTRY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

COMPARABILITY OF A COMPUTER TOUCH SCREEN VERSUS WRITTEN METHOD FOR COLLECTING ORAL HEALTH RELATED QUALITY OF LIFE (OHRQOL) DATA

MAZNURFARHATUNNISAK BINTI ANOWAR

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF DENTAL SCIENCE

FACULTY OF DENTISTRY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: (I.C/Passport No:)

Matric No:

Name of Degree:

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Field of Study:

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 to 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:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

ABSTRACT

Current advances in technology have enabled the growth of computerized questionnaires that provide many benefits over the paper-based mode of administration, such as automatic data entry, eliminate errors during data handling, storage, time and cost effectiveness and able to provide immediate results automatically. However, before a computerized questionnaire can be implemented in a clinical setting, its equivalence with its original paper-based questionnaires must first be demonstrated.

This study was divided into two parts. The purpose of the study for part 1 was to evaluate a computer touch screen version of Oral Health Impact Profile (OHIP-14) questionnaires (CTSA) against the traditional written paper version (QBA) for the score agreement, respondent's acceptance and preference, and time for completion. Meanwhile, for part 2, the purpose of the study was to investigate the OHIP-14 score agreement obtained using English (EQA) and Malay language version (MQA) via web-based questionnaires.

Fifty-nine students and sixty students were recruited from the Tun Syed Zahiruddin Residential College in University of Malaya for each part of the study consecutively. A randomized crossover design was used for both parts, with respondents randomly assigned into two groups. In part 1, respondents in group A completed CTSA first followed by QBA and group B completed QBA first followed by QBA. As for part 2, group A respondents answered EQA first followed by MQA and group B completed these questionnaires in the reverse order. All the respondents answered both versions of the questionnaire with a time interval of at least 3 hours between questionnaires. Times taken to complete both versions were recorded for both part 1 and 2. Respondents were asked about the preferred mode of administration in part 1.

Intraclass correlation coefficients between CTSA and QBA demonstrated a very good agreement for total score (0.90), while for the subscales the ICCs showed a dual

mixture of high and moderate agreement (0.58-0.90). The weighted kappa coefficients of individual items also demonstrated a substantial to fair agreement (0.25-0.78) between CTSA and QBA. The CTSA was well accepted and 78% of the respondents preferred CTSA over QBA. The time required to complete the CTSA did not differ statistically from the QBA. Meanwhile, for the part 2, the ICCs between EQA and MQA showed an excellent agreement at both the total score (0.92) and subscales level (0.76-0.93). The weighted kappa coefficients for items showed a substantial to fair agreement (0.33-0.72).

The CTSA is equivalent to its original written questionnaires which show a strong score agreement, eliminate the need for data entry and able to provide immediate results. It is well accepted and preferred over the QBA with no significant difference found for the time completion. The level of agreement between English and Malay version of OHIP-14 questionnaires has showed a very high level of agreement. Therefore, both languages can be used for literate Malaysian population.

ABSTRAK

Perkembangan teknologi semasa telah menggalakkan penggunaan soal selidik berkomputer. Ia memberikan banyak kelebihan berbanding soal selidik bertulis seperti kemasukan data secara automatik, mengelakkan kesilapan semasa pengendalian data, memudahkan penyimpanan, efektif dari segi kos dan masa, serta mampu memberikan keputusan secara serta-merta dan automatik. Walau bagaimanapun, sebelum soal selidik berkomputer boleh dilaksanakan dalam persekitaran klinikal, kesetaraan antara soal selidik berkomputer dengan soal selidik bertulis harus dibuktikan terlebih dahulu.

Kajian ini terbahagi kepada dua bahagian. Objektif kajian bagi bahagian pertama adalah untuk menilai kesetaraan antara soal selidik "Oral Health Impact Profile (OHIP-14)" yang diambil menggunakan soal selidik berkomputer dengan soal selidik bertulis dari segi skor, penerimaan dan pilihan responden, dan masa yang diperlukan untuk menjawab soal selidik. Sementara itu, objektif kajian untuk bahagian kedua pula adalah untuk mengkaji kesetaraan antara skor OHIP-14 yang dijawab menggunakan bahasa Inggeris (EQA) dengan Bahasa Malaysia (MQA) melalui soal selidik berasaskan web.

Seramai 59 pelajar dan 60 pelajar dari Kolej Kediaman Tun Syed Zahiruddin di Universiti Malaya mengambil bahagian di dalam setiap bahagian kajian. Reka bentuk "randomized crossover" telah digunakan untuk kedua-dua bahagian kajian. Responden dibahagikan kepada dua kumpulan secara rawak. Responden di dalam kumpulan A akan menjawab CTSA dahulu diikuti dengan QBA dan responden di dalam kumpulan B akan menjawab QBA dahulu diikuti dengan CTSA untuk bahagian pertama kajian. Manakala untuk bahagian kedua, responden di dalam kumpulan A akan menjawab EQA dahulu diikuti oleh MQA dan untuk kumpulan B, responden akan menjawab MQA dahulu diikuti EQA. Semua responden menjawab kedua-dua versi soal selidik dengan selang masa sekurang-kurangnya 3 jam antara soal selidik. Masa yang diambil untuk

melengkapkan kedua-dua versi bagi bahagian 1 dan 2 direkodkan. Untuk bahagian pertama kajian, responden diminta untuk memilih mod pilihan umtuk menjawab soal selidik sama ada berkomputer, bertulis atau tiada pilihan.

Nilai “intraclass correlation (ICC)” yang diperolehi antara CTSA dan QBA menunjukkan kesetaraan yang sangat baik bagi jumlah keseluruhan skor (0.90), manakala nilai ICC bagi subskala menunjukkkan kesetaraan yang sederhana dan tinggi (0.58-0.90). Nilai “weighted kappa” untuk setiap item individu juga menunjukkan nilai yang wajar dan tinggi (0.25-0.78) di antara CTSA dan QBA. CTSA telah diterima dengan baik dan 78 peratus responden memilih CTSA berbanding QBA. Masa yang diperlukan untuk responden melengkapkan CTSA tidak berbeza secara statistik daripada QBA. Sementara itu, bagi bahagian kedua kajian, nilai ICC antara MQA dan EQA menunjukkan kesetaraan yang sangat baik bagi jumlah keseluruhan skor (0.92) dan subskala (0.76-0.93). Nilai “weighted kappa” untuk setiap item pula menunjukkan nilai yang wajar dan tinggi (0.33-0.72).

CTSA terbukti dapat menghasilkan hasil yang setara dengan QBA di mana kedua-duanya menunjukkan kesetaraan skor yang sangat baik. Selain itu, ia juga tidak memerlukan proses kemasukan data secara manual dan mampu memberikan keputusan soal selidik secara serta-merta. Ianya juga diterima dengan baik dan menjadi mod pilihan umtuk menjawab soal selidik selain tiada perbezaan signifikan dengan QBA bagi masa yang diperlukan untuk melengkapkannya. Tahap kesetaraan di antara soal selidik OHIP-14 versi bahasa Inggeris dan bahasa melayu telah terbukti berada di tahap yang sangat tinggi. Oleh itu, kedua-dua versi bahasa boleh digunakan untuk penduduk Malaysia.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my greatest gratitude towards Allah S.W.T., the Most Compassionate and Merciful, for providing me the strength, patience, blessing, and guidance to accomplish this thesis.

I would like to take this opportunity to express millions of appreciations and gratitude to my supervisor, Associate Professor Dr. Roslan Bin Saub for his kindness and attention, being so patience in his guidance and dedication in teaching. His encouragement and constructive comments from beginning till the end of this project is gratefully acknowledged.

My greatest appreciation goes to my co-supervisor Professor Colman McGrath from The University of Hong Kong for all his guidance, knowledge, advice, and constructive critics in the on-going process of this research.

I am really grateful for having both of them as my supervisor as all their cares, teachings, supports, encouragement, and comments inspire me to complete this thesis. All the knowledge and experience gained throughout the whole process were very valuable and directly contributed to the achievement of this study. Special thanks to Dr. Mahmoud Danaee from the Academic Development Centre (ADeC) University of Malaya and Mr. Vignes Gopal Krishna for guidance for data analysis for this study.

I would like to thank Associate Professor Dr. Sabri Bin Musa, the principle of the Tun Syed Zahiruddin Residential College for giving approval and allowing me to carried out the survey at the college. Not to forget, I would also express my gratitude to Mohammad Asnawi Bin Amiruddin, the student representative from the Ninth Residential College for his cooperation and willingness for helping throughout the data collection. I would like to thank all the Ninth Residential College residents for their

cooperation, helpfulness during the progress of this survey, and willingness to participate this survey.

My sincere appreciations and gratitude goes to my late father, Anowar Bin Inun and my late mother, Maznah Binti Johan for their undivided love, support, encouragement, and memories that kept me fighting till the end. Thank you to my beloved brother for his love, understanding, and encouragement towards me during completing this project.

Special thank to my friend, Amirul Nurshuhada Binti Romle for developing an excellent web-based questionnaires system that really helps me in gathering the data efficiently and helping me during the data collection for phase 1. Last but not least, I would like to thank all my beloved lecturers, friends and classmates for all their help. Thank you so much for all the help, suggestions, comments, assistance, and support that accompanied me along this way. From the bottom of my heart, I really appreciate and always remember all the moments we shared together. In addition, my greatest gratitude to those who have contribute whether directly or indirectly throughout the whole process of this study.

Thank you so much and may Allah bless us all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

OHRQoL	:	Oral Health-Related Quality of Life
OHIP-14	:	Short version of the Oral Health Impact Profile
OHIP (M)	:	Malaysian Oral Health Impact Profile
HRQoL	:	Health-Related Quality of Life
GOHAI	:	Geriatric (General) Oral Health Assessment Index
SIDD	:	The Social Impacts of Dental Disease
OHIP	:	Oral Health Impact Profile
DIP	:	The Dental Impact Profile
OIDP	:	Oral Impacts on Daily Performance
WHO	:	World Health Organization's
S-OHIP (M)	:	Short version of the Malaysian Oral Health Impact Profile
QoL	:	Quality of Life
CTSA	:	Computer Touch Screen Questionnaires
QBA	:	Written Questionnaires
SPSS	:	Statistical Packages for Social Sciences
SD	:	Standard Deviation
ICC	:	Intraclass Correlation Coefficient
HTTP	:	Hypertext Transfer Protocol
URL	:	Uniform Resource Locator
IC Number	:	Identification card number
MQA	:	Web-Based Malay version of OHIP-14
EQA	:	Web-Based English version of OHIP-14

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background.

Oral health-related quality of life (OHRQoL) plays an important role in clinical practice in terms of identifying needs, selecting therapies, and monitoring patients progress (Fernandes, Ruta, Ogden, Pitts, & Ogston, 2006). There are several instruments that have been developed to assess OHRQoL. In this study, assessment of OHRQoL will be using a short form version of the Oral Health Impact Profile (OHIP-14). OHIP-14 is a self-completed questionnaire that consists of 14 items subdivided into 7 subscales which are functional limitation, physical pain, psychological discomfort, physical disability, psychological disability, social disability and handicap (Locker, Jokovic, & Clarke, 2004; Ng & Leung, 2006). Responses to the items will be recorded in a simple Likert-type frequency scale.

Traditionally, OHRQoL data have been collected through self-reported questionnaires printed on a paper form and the responses are usually entered into a database manually. These may cause some problems in the process of data entry, especially if dealing with a large number of samples. With the use of electronic questionnaires, some of the problems regarding the process of data entry may be overcome since the result can be compiled automatically in a database (Cella, 1995; G Velikova et al., 1999). The use of computerized questionnaires may avoid missing data and contribute to the reduction of confusion because of only one item presented at a time and it is not possible to move on to the next question without completing the previous one (Finegan & Allen, 1994; Webb, Zimet, Fortenberry, & Blythe, 1999).

With the advancement in the touch screen technology, data gathering can further be improved. In addition, with touch screen questionnaires, the result can be obtained automatically, which makes it suitable to be used in clinical practice. However, before it

can be used in clinical practice, it is important to know whether it gives similar results as in written questionnaires.

Most OHRQoL instruments were originally developed in English. In order for these OHRQoL instruments to be used in non-English-speaking countries as well as in multiethnic societies, these instruments have been translated into other languages (Cheung et al., 2004; Ferrer et al., 1996; Saub, R., Locker, & Allison, 2005; Ward et al., 1999). The OHIP is one of the OHRQoL instruments that has been translated into many languages (Ravaghi, Farrahi-Avval, Locker, & Underwood, 2009; Tubert-Jeannin, Riordan, Morel-Papernot, Porcheray, & Saby-Collet, 2003). It has also been translated into the Malay language, which was designated as the Malaysian Oral Health Impact Profile [OHIP (M)]. The original OHIP was translated from English into the Malay language using a forward–backward translation technique (Saub, R. et al., 2005; Saub, R., Locker, Allison, & Disman, 2007).

However, to date, there has been no report on the agreement of the OHRQoL score obtained using OHIP-14 instrument between English and Malay version. Therefore, it is necessary to investigate whether the Malay version of the OHIP-14 instrument gives equivalent scores with its English translated version before using it for Malaysian population.

This study was divided into two parts; Part 1 and Part 2 which was planned to answer the following questions:

<u>Part 1.</u>

a) Whether the OHRQoL score obtained using a computer touch screen and written questionnaires are the same?

b) What is the respondent's acceptance and preference mode of administration of questionnaires?

c) Which mode of administration of questionnaires has higher feasibility?

Part 2.

a) Whether the OHRQoL score obtained using a computer touch screen for both English and Malay language are the same?

1.2 Research Objectives.

Aim.

To investigate the comparability of OHRQoL obtained via computer touch screen versus those obtained via written questionnaires.

Objectives.

Part 1:

a) To investigate the agreement between the OHRQoL score obtained using a computer touch screen and written questionnaires.

b) To assess the respondent's acceptance and preference mode of administration of questionnaires.

c) To assess the feasibility (the time taken to complete the questionnaires) of computerized versus written questionnaires.

Part 2:

To assess the agreement between the OHRQoL score obtained using web-based questionnaires for different languages.

1.3 Hypothesis of the Study.

Part 1:

a) There is an agreement between OHIP-14 score between computer touch screen and written questionnaire.
b) Respondents find that computer touch screen is the preferred mode of administration with high acceptance compared to the written questionnaire.
c) The time taken to complete the touch screen is shorter than written questionnaire.

Part 2:

a) There is an agreement between OHIP-14 score obtained via computer touch screen questionnaires for different languages.

CHAPTER 2: LITERATURE REVIEW

2.1 Oral Health Related Quality of Life (OHRQoL).

Oral health-related quality of life (OHRQoL) is somewhat new but rapidly growing concept. Based on the literature, it is evident that the concept of OHRQoL appeared only in the early 1980s compared to the general Health-Related Quality of Life (HRQoL) that started to become known in the late 1960s (Bennadi & Reddy, 2013). The poor perception of the impact of oral disease on QoL could be one of the causes for the delay in the development of OHRQoL. Several researchers rejected the idea that oral disease could be related to general health and apart from pain and life-threatening cancers, oral disease is only associated with cosmetic issues but does not have any impact on social life (Davis, 1976; Dunnell & Cartwright, 1972; Gerson, 1972). However, nowadays there is a rising interest in recognizing oral health as a part of the quality of life. Many efforts have been done by dental researchers in exploring the relationship between oral health status and quality of life, in order to assess, improve and maintain it.

OHRQoL is defined as a "self-reported measure specifically pertaining to oral health capturing both the functional, social and psychological impacts of oral disease" (Gift & Atchison, 1995). In later contributions, Locker defined OHRQoL in relatively simple terms as 'the extent to which oral disorders affect functioning and psychological well-being' (Locker, Clarke, & Payne, 2000) and 'the symptoms and functional and psychological impacts that emanate from oral diseases and disorders' (Locker, Matear, Stephens, & Jokovic, 2002).

It is acknowledged and accepted by the dental community that if oral symptoms remain untreated it can be a major source of diminished quality of life. Oral health status can cause considerable pain and suffering, thus may affect people's choice of

food, swallowing and speaking, may lead to lack of sleep, depression, and various psychological problems (de la Fuente Hernández, Díaz, & Vilchis, 2015). Even though oral health problems are rarely life threatening, they remain a major public health concern because of its prevalence and there are significant indications that oral health problems have a social, economic and psychological impact on the quality of life. Some researchers argued that OHRQoL is highly subjective because it is related to daily life and unique to each individual. Even patients with severe conditions can report that their quality of life is good. Therefore, the patients are the best judges of their own quality of life (Cella & Bonomi, 1995).

The OHRQoL is a multidimensional concept that is capturing people's perception and satisfaction about factors that are important in their daily life with respect to their oral health (Bennadi & Reddy, 2013). It is useful in clinical practice because it can help in identifying and prioritizing problems, facilitating communication, screening for hidden problems, facilitating shared clinical decision making, and monitoring changes or responses to treatment (Inglehart & Bagramian, 2002).

There are several instruments used to assess OHRQoL, but the most widely used are multiple item questionnaires that contain functional factors, psychological factors, social factors, and experience of pain or discomfort (Inglehart & Bagramian, 2002). Some of the examples of OHRQoL instruments are Geriatric (General) Oral Health Assessment Index (GOHAI), The Social Impacts of Dental Disease (SIDD), Oral Health Impact Profile (OHIP), The Dental Impact Profile (DIP), and Oral Impacts on Daily Performance (OIDP).

Among all OHRQoL instruments, application of the OHIP seems to be popular compared to other OHRQoL instruments because researchers found that information assessed using OHIP is very useful as a patient-based outcome measure (Armellini,

Heydecke, Witter, & Creugers, 2007; De Oliveira & Sheiham, 2004; Heydecke, Locker, Awad, Lund, & Feine, 2003; Nickenig, Wichmann, Andreas, & Eitner, 2008).

2.2 Oral Health Impact Profile (OHIP).

In this study, Oral Health Impact Profile (OHIP-14) was used as the instrument to measure OHRQoL. OHIP was developed by Slade & Spencer that based on Locker's adaptation of World Health Organization's (WHO) classification on impairments, disabilities, and handicaps (Locker, 1988). It is the most widely used OHRQoL questionnaire. OHIP was developed to measure self-perceived oral health. It is used to assess the 'social impact' of oral disorders, which means in term of the dysfunction, discomfort, and disability caused by these conditions. This instrument focused on documenting social impact among individuals and groups, understanding oral health behaviors, evaluating dental treatment and providing information for advocating for oral health in order to assess priorities of care (Locker & Allen, 2007).

The original 49-item OHIP contains 7 domains that assess the impacts of oral conditions on people's OHRQoL which are the functional limitation, physical pain, psychological discomfort, physical disability, psychological disability, social disability, and handicap. A short version of OHIP which is called as OHIP-14 was later developed. It was derived from a subset of 2 questions for each of the 7 domains (Locker, Matear, Stephens, Lawrence, & Payne, 2001). Since its development, the OHIP-14 has been favored over the OHIP-49 by a number of researchers due to its practicality.

The OHIP-14 responses were recorded on a five-point Likert scale: 0, never; 1,seldom; 2, sometimes; 3, quite often; 4, very often. Three different scoring methods have been reported in studies using the OHIP-14 which are a summary OHIP-14 score, a weighted and standardized summary score, and the total number of problems reported (Allen & Locker, 1997; Slade, 1997). Among these 3 methods, the summary score

method is recommended due to its simplicity (Allen & Locker, 1997; Fernandes et al., 2006).

The OHIP-14 questionnaires have been translated and validated into many languages in different regions of the world such as Chinese, Persian, Sinhalese, Korean, Greek, Romanian, and including Malaysian. All research regarding the translated OHIP-14 questionnaires showed that the translated version proved to be valid, and reliable to be used among its population (Bae et al., 2007; Ekanayake & Perera, 2003; Navabi, Nakhaee, & Mirzadeh, 2010; Papagiannopoulou, Oulis, Papaioannou, Antonogeorgos, & Yfantopoulos, 2012; Ravaghi et al., 2009; Slusanschi et al., 2013; Xin & Ling, 2006).

In Malaysia, the development of a short version of the Malaysian Oral Health Impact Profile designated as the S-OHIP (M) has been conducted by using a forward-backward translation technique followed by a method known as the 'item frequency method'. The study found that the S-OHIP (M) is valid, reliable and appropriate for use in the Malaysian adult population (R Saub et al., 2005; R Saub et al., 2007).

2.3 Quality of Life (QoL) measures Used in Clinical Practice.

In recent years, there has been a growing awareness in the use of quality of life assessments in daily clinical practice as an aid to detect physical or psychological problems that might be overlooked, to monitor disease and treatment, and improve the delivery of care (Jacobsen, Davis, & Cella, 2002; Lohr, 1992; G Velikova, Brown, Smith, & Selby, 2002). QoL measures have several potential uses in aiding clinical practice. Previous research suggested that QoL reports helps in facilitating physician-patient communication, increase physician's awareness of their patients' QoL, helps in identifying and prioritizing problems, and increase the frequency of QoL issues discussed (Bennadi & Reddy, 2013; Detmar, Muller, Schornagel, Wever, & Aaronson,

2002; Higginson & Carr, 2001; Jacobsen et al., 2002; McLachlan et al., 2001; Galina Velikova et al., 2004).

QoL measures are not a replacement for measuring outcomes associated with disease but are an adjunct to them. The fundamental reason for applying QoL measures in clinical practice is to make sure that treatment plans and evaluations focus on the patient, not just the disease (Kleinman, Leidy, Crawley, Bonomi, & Schoenfeld, 2001). One way of ensuring that quality of life assessments influence clinical decision making is to use them as a basis for making choices about treatment.

Despite the acceptance of QoL as outcome measures in clinical research and the fact that Florence Nightingale was one of the earliest clinicians that insisted on measuring the outcome of routine care to evaluate treatment (Rosser, 1985), these measures are rarely used in routine clinical practice. This might be due to some barriers that occur in the process of transferring QoL measures from research to clinical practice. One of the most important barriers is the lack of provider experience in conducting QoL assessment. Some previous research showed that with sufficient exposure to these measures, providers were willing to use QoL measures on a routine basis (Detmar et al., 2002; Jacobsen et al., 2002). Concerns about reliability and validity of the QoL measures are also one of the barriers. In order for a measure to be used in clinical practice, it must be valid, appropriate, and reliable (Higginson & Carr, 2001; Jacobsen et al., 2002).

Most existing measures were developed for use in clinical research where time constraints are different from those in clinical practice. Therefore, problems regarding feasibility and logistics are also barriers that occur in the process of transferring QoL measures to clinical practice. Healthcare providers are less likely to use the assessment unless the assessments are simple, understandable, quick to complete, user-friendly, easy to score and provide immediate results (Bezjak et al., 2001; Higginson & Carr,

2001). Recently, electronic methods of collecting data from patients have allowed real-time QoL measurements and presentation of results to clinicians, making this approach feasible in busy clinical practices.

Many OHRQoL instruments are known to be reliable and valid but trials are required to evaluate their routine use in clinical practice (Keren, Locker, & Grad, 2004). OHIP-14 is one of the OHRQoL instruments that has been proven as a valid and reliable instrument to measures OHRQoL in general dental practice (Fernandes et al., 2006; Saub, R. et al., 2005) but not for Malaysian population as yet.

2.4 Method for Collecting OHRQoL Data.

Gradually OHRQoL has gained recognition in the field of dentistry. The dentist and researchers agree that the patient's point of view is as equally important as clinical parameters (Inglehart & Bagramian, 2002). Over recent years, several methods have been developed in order to collect the OHRQoL questionnaires. Two basic methods that have been suggested to be used for collecting OHRQoL questionnaires are interview and self-administration (Fayers & Machin, 2013; Reissmann, John, & Schierz, 2011).

An interview can either be conducted face-to-face or via the telephone, while for the self-administered questionnaire can be completed either manually on paper or electronically via computer touch screen or web-based (Reissmann et al., 2011). All modes of administrations have their advantages and disadvantages in terms of patient burden, feasibility, response rates and costs. Besides, it is also related to the environment in which they are used because some methods can be applied in one setting, but some may not (Reissmann et al., 2011).

However, it is largely unknown whether the data quality of OHRQoL assessments is affected by the mode of administration. This is an important concern because if the mode of administration does influence OHRQoL assessments, then this issue might

need to be considered when designing future studies. Data quality is a vague concept, therefore there is no agreed gold standard in assessing data quality. It may be evaluated based on the response rates, the precision of responses, lack of bias, and no missing information obtained from respondents (Bowling, 2005).

In general, interview surveys are associated with higher response rates compared to a self-administered questionnaire but lead to social desirability bias (Bowling, 2005; Tsakos, Bernabé, O'Brien, Sheiham, & de Oliveira, 2008). The respondents tend to give more positive responses, resulting in the over-reporting of desirable behaviors and under-reporting of undesirable behaviors (Bowling, 2005). Besides that, interview survey requires higher administration cost due to the need of hiring and training for the interviewers (Tsakos et al., 2008).

On the other hand, the self-administered questionnaire may be affected by respondent bias, for example, due to exclusion of respondents with reading difficulties. However, this mode of administration gives respondents more time to reflect on each answer (Desai, Durham, Wassell, & Preshaw, 2014), more practical, cost effective and reducing administrative burden which makes it more suitable to be used in clinical practice (Tsakos et al., 2008). In addition, self-administered questionnaires permit more confidentiality, therefore increase respondent's willingness to disclose sensitive information that leads to a more accurate report on sensitive topics such as health and behaviors (Bowling, 2005).

Several researchers have conducted studies on the effect of the administration method on OHRQoL assessments. Some studies showed that methods of administration did not influence the OHRQoL assessments (Malter, Hirsch, Reissmann, Schierz, & Bekes, 2015; Reissmann et al., 2011; Rutherford et al., 2016; Tsakos et al., 2008), but

some studies showed little effect of mode of administration on the OHRQoL assessments (Desai et al., 2014; Gundy & Aaronson, 2010).

Sometimes it might be needed to mix multiple modes of administration in the same study, for example when conducting longitudinal research or combining data from various sources. Therefore, it is important to choose the best method of administering an instrument (Hawthorne, 2003) and all users of questionnaires must be aware of the potential effects of methods of administration on their data (Bowling, 2005). The use of a computerized mode of administration can be considered in clinical practice since it gives several benefits such as the ability to provide good quality data with no missing data, time and cost effectiveness, thus facilitating the dentist in daily clinical practice.

2.5 Comparability of A Self-administered Computerized Versus Written Questionnaires.

In this study, the self-administered questionnaire was used as the method for collecting OHRQoL data. Paper and computer touch screen were used for the first phase, meanwhile, for the second phase, a web-based questionnaire was used. Therefore, the discussion will be more focused on these methods.

Traditionally, the self-administered questionnaires are collected using paper-and-pencil form. However, there are several disadvantages associated with the use of these written questionnaires. The needs for manual entering or scanning individual questionnaires into a database when using the written questionnaire is time consuming and may increase error rate (Ryan, Corry, Attewell, & Smithson, 2002; G Velikova et al., 1999). Respondents may either miss or mark an item ambiguously thus leads to missing data (Ryan et al., 2002). Written questionnaire can be costly and labor intensive

due to the need of health professionals to print out, distribute the questionnaire, collect the data after completion, and manually enter the data into a database (Handa et al., 2008; Eun Hyun Lee, 2007; Morris, Perez, & McNoe, 1997).

By means of the emergence of computer technology, computerized questionnaires have become a more practical alternative compared to written questionnaire in collecting surveys (Bennett, Jensen, & Basch, 2012; C.-H. Chang, 2007; Gwaltney, Shields, & Shiffman, 2008). Computerized administered questionnaire has the ability to collect high quality data with no missing data or problematic responses by providing software safeguard that prevents accidental omission of questions by respondents and completely eliminating error during the data entry (Greenwood, Hakim, Carson, & Doyle, 2006; Eun-Hyun Lee et al., 2014; F Salaffi, Gasparini, & Grassi, 2009). Besides that, computerized questionnaires offer the benefits of time saving in terms of administration, scoring, and data entry because it provides automatic data entry, permits immediate calculation and allows immediate access to results (Chen & Li, 2010; Handa et al., 2008). A web-based questionnaire is one of the most recent computerized methods that seem to be promising due to its characteristic that allows data entry from any place that is internet-accessible location and at any time convenient to the respondent. This approach is really useful and suitable to use in a multicenter clinical trial and epidemiological studies (Gwaltney et al., 2008; Handa et al., 2008).

Up till now, there is no study done on comparing the computerized versus written questionnaire in collecting OHRQoL, including OHIP-14. However, many studies have been conducted which compares both modes of administrations in collecting Health-Related Quality of Life (HRQoL) questionnaires and other questionnaires in various fields. Most of the studies aims are to measure the equivalence of both methods, evaluate the feasibility, and respondents' preferences between computerized versus

written questionnaires, and data completeness. Table 2.1 showed the list of related studies and its findings. However, only three aspects that are directly related to our study objectives which are score agreement, feasibility and respondent's preferences will be discussed further in this chapter.

2.5.1 Score Agreement.

Based on the literature, most of the studies found that the scores agreement between written and computerized questionnaire were equivalent and showed moderate to excellent agreement, thus can be used interchangeably (Ashley, Keding, Brown, Velikova, & Wright, 2013; Eun-Hyun Lee et al., 2014; Oliveira, Ferreira, Antunes, & Pimentel, 2010). Only a few studies showed different findings which are related to the order effect of mode of administration. They found that respondents who completed the touch screen version first had higher scores that those completing the touch screen second (Ring et al., 2006). The most common tests used in assessing the score agreement are intraclass correlation coefficient (ICC), weighted kappa, Cronbach's α, and concordance correlation coefficients (CCC).

2.5.2 Respondent's Preferences.

Almost all studies that compare computerized and written questionnaire showed that computerized questionnaire is preferable compared to written questionnaire (Bent et al., 2005; Broering et al., 2014; Ribeiro et al., 2010). Only one study showed same proportion of respondents preferred both modes of administration (E.-H. Lee, 2009b). The majority of the respondents preferred computerized version because of its user-friendly, simple and well-arranged interface which presents only one question at a time (Ring et al., 2006; F Salaffi et al., 2009). Besides that, respondents also rated computerized questionnaire as easy to use and to navigate even by the computer naïve (Y.-J. Chang et al., 2014; Greenwood et al., 2006; Wilde Larsson, 2006). Most of the

studies found that age, computer experience, and education level had no significant impact on the respondent's preference for mode of administration (Gudbergsen et al., 2011; Fausto Salaffi, Gasparini, Ciapetti, Gutierrez, & Grassi, 2013).

2.5.3 Feasibility.

The feasibility was assessed by measuring the time needed for the respondents to complete answering the questionnaires. Various findings were found regarding the feasibility of both modes of administration. Most of the studies found that computerized questionnaire required shorter time completion than written questionnaire (Chen & Li, 2010; Fausto Salaffi et al., 2013; F Salaffi et al., 2009), but some found that written questionnaires were faster (Y.-J. Chang et al., 2014; Goodhart et al., 2005; E.-H. Lee, 2009a). There were also studies that showed no significant difference in time completion for both modes (Greenwood et al., 2006; E.-H. Lee, 2009b). The findings of time completion may vary but if the time needed for the whole transfer process of data administration into database were taken into consideration, the computerized questionnaire is the most feasible mode of administration (Chen & Li, 2010; Goodhart et al., 2005).

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires.

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Chang, Y. J., Chang, C. H., Peng, C. L., Wu, H. C., Lin, H.C., Wang, J. Y., ... & Liang, W. M. (2014)	99 prostate cancer patients (Taiwan)	EORTC QLQ-PR25 (Taiwan Chinese version)	▪ 120 minutes.	▪ Paper-and-pencil. ▪ Touch screen.	▪ Data obtained from both modes were equivalent (ICC showed moderate to excellent agreement). ▪ 67% of patients preferred the touch screen. ▪ Approximately 92% of patients found that touch screen easy to use regardless of age groups and computer-use experience. ▪ Touch screen required approximately 30% more time to complete than the paper. ▪ Touch screen feasible and suitable alternative to paper mode.
Lee, E. H., Lee, Y. W., Lee, K. W., Kim, D. J., Kim, Y. S., & Nam, M. S. (2014)	208 diabetes patients (South Korea)	Diabetes-Specific Quality-of-Life questionnaire (D-QOL)	▪ Approx. 20 to 25 minutes.	▪ Paper-and-pencil. ▪ Touch screen.	▪ The touch screen was equivalent to paper-based (ICC for both total score and subscales were excellent, weighted kappa showed substantial & almost perfect agreement). ▪ No significant difference was found for the time completion. ▪ 82.7% preferred touch screen over paper questionnaire. ▪ 86.9% reported that touch screen was easy to use.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Broering, J. M., Paciorek, A., Carroll, P. R., Wilson, L. S., Litwin, M. S., & Miaskowski, C. (2014).	245 men with prostate cancer (Gudbergsen et al.)	Short Form-36 (SF-36) & University of California Los Angeles Prostate Cancer Index (UCLA-PCI)	▪ 2 to 5 days.	▪ Paper mode. ▪ Web-based mode.	▪ Web-mode and paper mode administration of the SF-36 and UCLA-PCI are equivalent (ICC were high for each item on both instruments). ▪ Patients evaluated both modes as easy to navigate and neither mode was confusing or stressful to complete. ▪ 70% of patients preferred the web-mode. ▪ Patients rated the web-mode as more convenient and faster to complete.
Ashley, L., Keding, A., Brown, J., Velikova, G., & Wright, P. (2013)	111 patients (United Kingdom)	The Social Difficulties Inventory (SDI-21)	▪ Complete both versions within 2 weeks.	▪ Paper mode. ▪ Web-based mode.	▪ Paper and web-based versions can be considered equivalent and used interchangeably. ▪ Score distributions and internal reliabilities for both versions were highly similar (ICC between paper and web-based score were uniformly high and above the standard acceptable level of reliability).

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Salaffi, F., Gasparini, S., Ciapetti, A., Gutierrez, M., & Grassi, W. (2013)	55 patients with axial spondyloarthritis (SpA)	Patient-reported outcome (PRO)	▪ 60 minutes.	▪ Paper-and-pencil. ▪ Touch screen.	▪ The agreement between paper and touch screen format was excellent (ICC range from 0.90 to 0.96). ▪ 83.4% preferred the touch screen format. ▪ Mean time spent completing the touch screen questionnaire was shorter than paper format. ▪ Age, computer experience, and education level had no significant impact on the results.
Gudbergsen, H., Bartels, E. M., Krusager, P., Wæhrens, E. E., Christensen, R., Danneskiold-Samsøe, B., & Bliddal, H. (2011)	20 female patients with knee osteoarthritis (Denmark)	Knee Osteoarthritis Outcome Score (KOOS), VAS pain, function and patient disability, SF-36, Physical Activity Scale, painDETECT, Activity of Daily Living (ADL) Taxonomy.	▪ 5 minutes.	▪ Paper. ▪ Touch screen.	▪ Touch screen questionnaires gave comparable results to answer given on paper (ICC range from 0.86 to 0.99 for all 6 questionnaires used). ▪ 16 out of 20 patients preferred the touch screen version. ▪ The touch screen was reported to be easier to use. ▪ Age, computer experience or education level had no significant impact on the results.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Oliveira, A., Ferreira, P. L., Antunes, B., & Pimentel, F. L.(2010)	200 cancer patients (Portugal)	EORTC QLQ-C30	▪ 1 hour.	▪ Paper-and-pencil. ▪ Touch screen.	▪ The touch screen showed similar results compared to paper. ▪ No statically significant differences between both modes. ▪ High correlations and large agreements between both modes. ▪ Touch screen has higher acceptability among patients.
Ribeiro, C., Moreira, L., Silveira, A., Silva, I., Gestal, J., & Vasconcelos, C. (2010)	50 patients with chronic immune system diseases (Portugal)	Short Form 36version2 (SF-36v2)	▪ Clinical consultation.	▪ Paper-and-pencil. ▪ Touch screen.	▪ Very high correlations in SF-36v2 responses between both modes. ▪ Internal reliability coefficients (Cronbach's α) showed good internal consistency. ▪ 84% preferred the touch screen. ▪ The touch screen provides valid and reliable results, produce comparable results to paper, highly accepted and preferred as the mode of administration.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Chen, T., & Li, L. (2010)	100 medical college freshman (China)	Chinese SF-36	▪ 30 minutes.	▪ Paper-and-pencil. ▪ Computer (Laptop).	▪ SF-36 scores were equivalence for both modes. ▪ 89% showed good or excellent agreement (Kappa coefficients). ▪ Time completion using computer quicker than paper.
Frennered, K., Hägg, O., & Wessberg, P. (2010)	79 patients (Sweden)	Visual Analogue Scale, Euro-Qol 5, SF-36, General Function Score, Zung Depression Scale.	▪ Approx. 3 weeks.	▪ Mailed paper questionnaires. (Mailed 1 to 3 weeks before the visit). ▪ Web-based.	▪ Touch-screen questionnaires for background data, pain, function, quality of life, and depression performed well compared to paper forms. ▪ The SF-36 tended to produce slightly higher values on touch screen recordings in questions concerning mental health. ▪ Touch screen virtually eliminates missing values and show good validity and reliability compared to paper forms.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Lee, E. H. (2009)	261 patients (Korea)	Asthma-Specific Quality of Life (A-QOL)	▪ Appointment with the physician.	▪ Paper-and-pencil. ▪ Touch screen.	▪ The touch screen was equivalent to paper (Weighted kappa coefficients of all items showed almost perfect agreement). ▪ The time required to complete the paper version is faster than the touch screen. ▪ 37.5% preferred the touch screen, 29.9% preferred paper. ▪ Most patients reported that touch screen was "easy" or "very easy" to complete.
Lee, E. H. (2009)	105 patients (South Korea)	Cancer-Specific Quality of Life (C-QOL)	▪ Appointment with the physician.	▪ Paper-and-pencil. ▪ Touch screen.	▪ The touch screen was equivalent to paper (Weighted kappa coefficients of items showed very good to moderate agreement). ▪ Time completion for touch screen did not differ statistically from the paper version. ▪ The same proportion of patients preferred both versions. ▪ 94.8% reported that the touch screen was "easy" or "very easy" to complete.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Salaffi, F., Gasparini, S., & Grassi, W. (2009)	87 rheumatoid arthritis patients (Italy)	Rheumatoid Arthritis (RA) patient-reported outcome (PRO)(Greenwood et al., 2006)	▪ 60 minutes (minimum).	▪ Paper-and-pen. ▪ Touch screen.	▪ Agreement between scores obtained with the two modes was very good, with concordance correlation coefficients (CCCs) from 0.887 to 0.972. ▪ Nearly all found the touch-screen easy to use, although over half the patients had no prior computer experience. ▪ 86% preferred the computer format to the paper format. ▪ The mean time completion for touch screen (7.3 minutes) was smaller than paper (7.9 minutes). Older patients being slower both on modes. ▪ The quality of the data collected with the touch-screen system was good, with no missed responses.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Handa, V. L., Barber, M. D., Young, S. B., Aronson, M. P., Morse, A., & Cundiff, G. W. (2008)	52 women	Pelvic Floor Distress Inventory 20 (PFDI-20) & Pelvic Floor Impact Questionnaire 7 (PFIQ-7)	▪ Varied between 2 days to 6 weeks.	▪ Paper mode. ▪ Web-based.	▪ There were no significant differences in scores or scale scores between the web-based and paper questionnaires. ▪ The web-based format was preferred by 53% participants. ▪ Paper and web-based versions of the PFDI-20 and PFIQ-7 can be used interchangeably.
Ring, L., Lindblad, Å. K., Bendtsen, P., Viklund, E., Jansson, R., & Glimelius, B. (2006)	40 pharmacy students (Sweden)	The Schedule for the Evaluation of Individual Quality of Life (SEIQoL-DW)	▪ 4 to 21 days.	▪ Paper-and-pen. ▪ Touch screen.	▪ Respondents who completed the touch screen version first had higher scores that those completing the touch screen second. ▪ Time completion for touch screen was faster than the paper version. The mean completion time for touch screen was about half that taken to complete the paper version. ▪ 65% preferred touch screen due to it being user-friendly, familiar, simple and well-arranged.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Greenwood, M. C., Hakim, A. J., Carson, E., & Doyle, D. V. (2006)	40 rheumatoid arthritis patients (United Kingdom)	Rheumatoid Arthritis Quality of Life Questionnaire (RAQoL)	▪ Not stated.	▪ Paper-and-pencil. ▪ Touch screen.	▪ Touch screen questionnaires can produce comparable results to paper (ICC were high, and Cronbach's α showed that touch screen retained a high level of internal consistency). ▪ There was no significant difference in time completion for both modes. ▪ 64% preferred touch screen and rated as easy to use even by computer naïve. ▪ Touch screen acceptable and preferable regardless of age and previous experience of computers.
Wilde Larsson, B. (2006)	199 patients answered touch screen, 219 patients answered paper-and-pen (Sweden)	Quality from the patients' perspective (QPP) questionnaire	▪ Not applicable	▪ Paper-and-pen. ▪ Touch screen.	▪ Both methods yielded almost identical results in the quality of care ratings. ▪ The touch-screen method was perceived to be easier to use, take less time to complete and resulted in a complete data set. ▪ Elderly patients and patients with compulsory comprehensive school education only, appeared to find the touch-screen method easier to use ▪ Both methods are acceptable, but the touch-screen method appears to be preferred by patients.

Table 2.1: List of Studies on Comparability of Computerized Versus Written Questionnaires (Continued).

Author	Sample (Country)	Instrument	Time Interval	Mode of Administration	Outcomes
Goodhart, I. M., Ibbotson, V., Doane, A., Roberts, B., Campbell, M. J., & Ross, R. J. M. (2005)	50 pituitary patients (United Kingdom)	Questions on Life Satisfaction-Hypopituitarism (QLS-H)	▪ 30 minutes.	▪ Paper-and-pencil. ▪ Touch screen.	▪ No statistically significant difference between scores obtained by the touch screen and paper questionnaires. ▪ 80% preferred the touch screen. ▪ Patients took less time to answer the paper questionnaire.
Bent, H., Ratzlaff, C. R., Goligher, E. C., Kopec, J. A., & Gillies, J. H. (2005)	50 patients with low back pain (Canada)	Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Functional Index (BASFI), Global Score (BAS-G) & the Quebec Scale.	▪ At least 40 minutes.	▪ Paper. ▪ Touch screen.	▪ Strong agreement between paper and computer-administered versions of the Quebec Scale, the BASDAI, BASFI, and BAS-G. ▪ No statistically significant difference in completion time between the 2 methods of administration. ▪ 84% indicated a preference for the touch screen method.

2.6 Score Agreement Using Different Language Questionnaires.

In part 2 of the study, the English and Malay version of the OHIP-14 questionnaire were used to assess the score agreement between both versions. Therefore, the discussion will be more focused on previous studies that investigate the agreement or equivalence of questionnaires using different language.

Quality of life (QoL) is increasingly being used as an outcome measure in clinical trials and most QoL instruments originally have been developed in English (Guillemin, Bombardier, & Beaton, 1993). The psychometric instruments that have been developed for use in English-speaking countries may not be appropriate for use with multi-cultural populations. Therefore, many QoL instruments have been translated into other languages and culturally adapted for use not only in non-English-speaking countries but also in multinational clinical trials and in a multiethnic population. In multiethnic populations, assessing QoL becomes challenging as culture and language may affect respondents' interpretation of and answers to QoL instruments (Aaronson, 1988; Avis, Assmann, Kravitz, Ganz, & Ory, 2004; Lubetkin, Jia, Franks, & Gold, 2005). Combining the scores of a QoL instrument administered in different languages increases the power and representativeness of such studies. However, the instrument needs to be available in several languages and the scores of these versions must be equivalent.

A model of six key types of equivalence in approaching the adaptation of QoL questionnaires has been suggested by Herdman et al. (Herdman, Fox-Rushby, & Badia, 1998) which includes conceptual equivalence, item equivalence, semantic equivalence, operational equivalence, measurement equivalence, and functional equivalence. Measurement equivalence refers to comparability in reliability, validity, responsiveness, prediction, and other measurement properties. Measurement equivalence also referred to

metric equivalence (Anderson, Aaronson, Leplege, & Wilkin, 1996), differs from conceptual equivalence and psychometric equivalence. It is achieved if a scale generates comparable scores for individuals at the same level of health regardless of the populations they come from (Herdman, Fox-Rushby, & Badia, 1997).

The equivalence of different language versions of a QoL instrument can be studied by assessing the similarity of psychometric properties of the instrument in bilingual (Guillemin et al., 1993) or monolingual (Gandek & Ware, 1998; Ware & Gandek, 1998) subjects. Each approach has its advantages and limitations and therefore complements the other. Using bilingual individuals to assess the equivalence of different language version, we can minimize variability in questionnaires scores due to random variation or known determinants of QoL (e.g. socio-economic status, age), thus improving the power to detect small differences between scale scores (Thumboo et al., 2002). Ideally, a randomized cross-over study design (Thumboo et al., 2002) using bilingual people (Gao et al., 2009) is often the best design for QoL equivalence studies.

Quantitative confirmation of equivalence between different language versions of the same instrument is an important issue in the development and application of QoL measures for use within multiethnic populations (Cheung & Thumboo, 2006). Different language versions of a QoL instrument would demonstrate measurement equivalence if they yielded similar scores at item and scale levels for respondents with identical levels of QoL (Drasgow & Kanfer, 1985).

The study of equivalence of different language versions of the same instrument is often controversial (Herdman et al., 1997). Based on the literature, most of the studies found that the scores agreement between different language questionnaire were equivalent and showed moderate to high agreement, thus either version can be used interchangeably (Chang, A. et al., 1999; Yu, J. et al., 2003; Thumboo, J. et al., 2002;

Cheung, Y. T. et al., 2013). However, few studies showed slightly different findings which found that most of the scales showed excellent equivalence but with exception of several scales that requires further refinement to the questions to strengthen the correlations between the two questionnaires (Cheung, Y. B. et al., 2004; Choy, S. C., Goh, P. S. C., & Liew, Y. L., 2015). Table 3.1 showed the list of related studies and its findings. The most common test used in assessing the score agreement between different language questionnaire are intraclass correlation coefficient (ICC), weighted kappa, Cronbach's α, and predefined equivalence margin.

Malaysia is a multiracial and multicultural country. The Malaysian population consists of 3 major ethnic groups (Malay, Chinese, and Indian). Bahasa Malaysia or the Malay language is the national and official language in the country. It is the medium of instruction and is a compulsory language to be learned in school (Lim, 2008). English is also a compulsory language taught in all Malaysian schools and widely spoken in the larger cities. The languages considered primary in Malaysia and which are used both for intra-group and inter-group communication are Malay, the national language, and English, an international language (Omar, 1987). Given such that Malay language is the official language and the rising need for multilingualism, it is important to translate OHRQoL measures into the Malay language and to assess both their measurement properties and their comparability with the English versions.

The Malaysian Oral Health Impact Profile [OHIP (M)] had been developed by using cross-cultural translation and adaptation, and found to be valid, reliable and appropriate for use in the cross-sectional studies in Malaysian adult populations (Roslan Saub, Locker, & Allison, 2005; R Saub, Locker, Allison, & Disman, 2007). Thus, the only equivalence that remained to be determined for the OHIP (M) would be the measurement equivalence. Our study is the first study that investigated the measurement

equivalence of the English and Malay versions of the OHIP-14 by investigating the score agreement between both versions.

Table 2.2: List of Studies on Score Agreement Using Different Languages.

Author	Sample (Country)	Instrument	Time Interval	Language Used	Outcomes
Choy, S. C., Goh, P. S. C., & Liew, Y. L. (2015).	177 full-time university students (Malaysia).	Learner Awareness Questionnaire (LAQ)	▪ Not stated.	▪ English. ▪ Malay language.	▪ The four factors of the Malay version showed moderate to strong correlations with those of the English versions. ▪ The results suggest that the Malay version of the questionnaire is similar to the English version. ▪ However, further refinement to the questions is needed to strengthen the correlations between the two questionnaires.
Ching, S. M., Yee, A., Ramachandran, V., Lim, S. M. S., Sulaiman, W. A. W., Foo, Y. L., & kee Hoo, F. (2015).	228 medical students from Universiti Putra Malaysia (Malaysia).	Smart Phone Addiction Scale (SAS)	▪ Not stated.	▪ English. ▪ Malay language.	▪ The parallel reliability between the SAS-M and the SAS was high, as demonstrated by an ICC of 0.95 (95% Confidence interval = 0.937–0.962). ▪ To date, this is the first study of its kind related to smart phone addiction, and it shows that the SAS-M is as good as the English version. ▪ This study also provides evidence that the SAS-M is a valid and reliable, self-administered tool to screen for those at risk of having smart phone addiction..

Table 2.2: List of Studies on Score Agreement Using Different Languages (Continued).

Author	Sample (Country)	Instrument	Time Interval	Language Used	Outcomes
Wang, Y., Tan, N. C., Tay, E. G., Thumboo, J., & Luo, N. (2015).	729 patients with type 2 diabetes mellitus (T2DM) (Singapore).	5-level EQ-5D (EQ-5D-5L)	▪ Not applicable.	▪ English. ▪ Chinese. ▪ Malay language.	▪ Equivalence was demonstrated between the Chinese and English versions and between the Malay and English versions of the EQ-5D-5L index scores. ▪ Equivalence was also demonstrated between the Chinese and English versions. ▪ However, equivalence could not be determined between the Malay and Chinese versions of the EQ-5D-5L index score. ▪ No significant difference was found in responses to EQ-5D-5L items between any languages, except that patients who chose to complete the Chinese version were more likely to report "no problems" in mobility compared to those who completed the Malay version of the questionnaire. ▪ This study provided evidence for the measurement equivalence of the different language versions of EQ-5D-5L in Singapore.
Tan, M. L., Wee, H. L., Lee, J., Ma, S., Heng, D., Tai, E. S., & Thumboo, J. (2013).	4,973 ethnic Chinese subjects (Singapore).	Short Form 36 version 2 (SF-36v2) & Short Form 6D (SF-6D) (Singapore).	▪ Not applicable.	▪ English. ▪ Chinese.	▪ All SF-36v2 domains were equivalent after adjusting for known HRQoL. ▪ SF-6D utility/items had the 90% CI either fully or partially overlap their predefined equivalence margin. ▪ The English- and Chinese-language versions of the SF-36v2 and SF-6D demonstrated equivalence. ▪ Compared with English-speaking Chinese, Chinese-speaking Chinese were significantly older (47.6 vs. 55.5 years).

Table 2.2: List of Studies on Score Agreement Using Different Languages (Continued).

Author	Sample (Country)	Instrument	Time Interval	Language Used	Outcomes
Cheung, Y. T., Lim, S. R., Shwe, M., Tan, Y. P., & Chan, A. (2013).	185 English-speaking and 143 Chinese-speaking patients with breast cancer (Singapore)	Functional Assessment of Cancer Therapy Cognitive Function (FACT-Cog).	▪ Not applicable.	▪ English. ▪ Chinese.	▪ Total FACT-Cog scores and cognitive domain scores for English and Chinese versions did not differ statistically, with the exception of the multitasking domain. ▪ Both versions of the FACT-Cog are equivalent, valid, and reliable to a great extent in assessing the perceived cognitive functioning of Asian patients with breast cancer. ▪ The English and Chinese FACT-Cog total scores had strong and moderate correlations. ▪ The English and Chinese versions of the FACT-Cog are valid, reliable, and equivalent for clinical and research use.
Gao, F., Ng, G. Y., Cheung, Y. B., Thumboo, J., Pang, G., Koo, W. H., ... & Goh, C. (2009).	771 patients [409 used the Chinese language, 362 used the English language] (Singapore).	EuroQol Group's 5-domain (EQ-5D).	▪ Not applicable.	▪ Singaporean English. ▪ Chinese.	▪ Based on the predefined equivalence margin of ±10% for binary outcome, ±0.05 for utility index, and ±5 points on the visual analog scale, the two language versions of the EQ-5D gave equivalent mean values at item and scale levels. ▪ They also showed similar characteristics in validity, responsiveness, and reliability. ▪ The Singaporean English and Chinese versions of the EQ-5D were validated in cancer patients and were shown to achieve measurement equivalence.

Table 2.2: List of Studies on Score Agreement Using Different Languages (Continued).

Author	Sample (Country)	Instrument	Time Interval	Language Used	Outcomes
Cheung, Y. B., Thumboo, J., Goh, C., Khoo, K. S., Che, W., & Wee, J. (2004).	1136 ethnic Chinese patients with cancer from the National Cancer Centre of Singapore (Singapore).	FACT-G & EORTC QLQ-C30.	▪ Not applicable.	▪ English. ▪ Chinese.	▪ The English & Chinese versions of the Total, Emotional, and Functional Well Being Scales of the FACT-G and the Physical and Emotional Functioning Scales of the EORTC QLQ-C30 were equivalent. ▪ Scores for the other scales on the 2 questionnaires, at most, had small differences that did not exceed 0.5 SD. ▪ The Chinese translation of the question "I have a lack of energy" in the Physical Well Being Scale of the FACT-G produced results that differed from the English version. (Main source of nonequivalence that led to the failure to confirm equivalence between both version of FACT-G Physical Well Being Scale). ▪ Data collected from English-speaking and Chinese-speaking respondents were capable of being pooled, and either version could be used for bilingual respondents.
Luo, N., Chew, L. H., Fong, K. Y., Koh, D. R., Ng, S. C., Yoon, K. H., ... & Thumboo, J. (2003).	66 subjects completed the English & 48 subjects completed the Chinese (Singapore).	EuroQol Group's 5-domain(EQ-5D).	▪ Not applicable.	▪ Singaporean English. ▪ Chinese.	▪ The 95% CI of the score differences between these versions overlapped with but did not fall completely within pre-defined equivalence margins for 4 EQ-5D items. ▪ The results of this exploratory study provide promising evidence for the measurement equivalence of Singaporean English and Chinese EQ-5D versions.

Table 2.2: List of Studies on Score Agreement Using Different Languages (Continued).

Author	Sample (Country)	Instrument	Time Interval	Language Used	Outcomes
Lim, T. O., Das, A., Rampal, S., Zaki, M., Sahabudin, R. M., Rohan, M. J., & Isaacs, S. (2003).	136 subjects (Malaysia).	International Index of Erectile Function (IIEF)	▪ Physician consultation.	▪ English. ▪ Malay language.	▪ Cronbach's alpha for both measures are uniformly high for all domains that ranges from 0.80 to 0.96. ▪ The instrument had good reliability and discriminant validity. ▪ More work is needed to fine tune the Malay IIEF to achieve equivalence with the English IIEF.
Thumboo, J., Fong, K. Y., Chan, S. P., Machin, D., Feng, P. H., Thio, S.T., & Boey, M. L. (2002).	168 bilingual, ethnic Chinese volunteers in Singapore. (Singapore)	Short Form 36 Health Survey.	▪ 3 to 16 days.	▪ English (UK). ▪ Chinese (HK).	▪ Cronbach's a exceeded the recommended value of 0.7 for all scales except English VT and RE scales, which had values of 0.68 and 0.67, respectively. ▪ Paired t-tests to assess the effect of the language of administration showed no significant difference between English and Chinese mean scale scores except for the VT scale. ▪ Intra-class correlations followed a similar pattern, being moderate to excellent for all scales by standard criteria, and exceeding 0.65 for all scales except the RP, SF and RE scales. ▪ English (UK) and Chinese (HK) SF-36 versions are equivalent in bilingual Singapore Chinese.

Table 2.2: List of Studies on Score Agreement Using Different Languages (Continued).

Author	Sample (Country)	Instrument	Time Interval	Language Used	Outcomes
Yu, J., Coons, S. J., Draugalis, J. R., Ren, X. S., & Hays, R. D. (2003).	309 Chinese international students and scholars from the University of Arizona (United State).	Short Form 36 Health Survey.	▪ Consecutively.	▪ Chinese. ▪ US-English.	▪ Cronbach's alpha coefficients for the English version ranged from 0.60 to 0.88 and for the Chinese version ranged from 0.64 to 0.85. ▪ Product–moment correlations ranged from 0.81 to 0.98 for the English–Chinese order of administration and from 0.82 to 0.95 for the Chinese–English order. ▪ There was no systematic effect of administration of the English and the Chinese version on the reliability estimates.
Chang, A. M., Chau, J. P., & Holroyd, E. (1999).	70 female university students (Hong Kong).	Menstrual Distress Questionnaire.	▪ 1 day.	▪ English. ▪ Chinese.	▪ Cronbach's alpha reliability coefficients of the total MDQ were 0.95 for the Chinese MDQ and 0.93 for the English MDQ. ▪ Eight items had weighted kappa values of less than 0.4, 35 items were between 0.4 and 0.75, with the remaining 4 items being more than 0.75. ▪ The intraclass correlation coefficients between 2 versions were high for the total score and half of the scales and moderate for the scales of Autonomic Reaction (0.71), Water Retention (0.72), Control (0.68) and Arousal (0.63).

CHAPTER 3: METHODOLOGY

This chapter describes the methodology used in this study. This study has two parts: Part 1 and Part 2.

3.1 Part 1: Computer Touch Screen versus Written Questionnaire.

In part 1, the respondents were asked to answer both touch screen and written questionnaire. The OHIP-14 scores for both instruments were analyzed.

3.1.1 Study Site.

One of the University of Malaya residential colleges that are the Ninth Residential College also known as Tun Syed Zahiruddin Residential College was chosen for this part of the study. It is located outside the University of Malaya which consists of 4 blocks for females, 2 blocks for males and 1 block for administrators. It currently houses 900 students pursuing undergraduate studies at the University of Malaya. The university students were chosen because we want to test the mode of administration, not the OHIP-14 instrument. Therefore it is important to test on a literate respondent.

3.1.2 Study Design.

This study used a randomized crossover study design. The participants were randomized into Group A and Group B. Group A started with test followed by control after the washout period of at least 3 hours and vice versa for Group B.

Test: The respondent answered the Oral Health Impact Profile (OHIP-14) questionnaire using the computer touch screen (CTSA).

Control: The respondent answered the OHIP-14 using written questionnaire (QBA).

3.1.3 Subject Recruitment.

An official letter of application for approval to conduct the survey at the college was sent to the principal of the Tun Syed Zahiruddin Residential College (Appendix A). After the application was approved (Appendix A), a meeting with the chairman of the student representative council of the college was held to discuss on the flow of the study.

With the help of the chairman of the student representative, an advertisement regarding the study was posted on the official Facebook group of the Ninth Residential College stating a brief explanation on the survey, requirement, time, date and the venue of the study. The survey was conducted at the Sri Tanjung hall. Students who were residents of the Ninth Residential College and a Malaysian citizen were eligible to participate in this study.

3.1.4 Sample Size.

Statistical power analysis techniques allow the quantitative researcher to decide on how large a sample is needed to enable statistical judgments that are accurate and reliable. It also helps the researcher to identify how likely is the selected statistical test to detect effects of a given size in a particular situation (StatSoft, 2003). The power of a test is the probability of correctly rejecting a false null hypothesis (Lane, 2007), thus avoiding a Type II error (beta) which sometimes considered a false negative. The power of a test equals 1 - β. One of the main functions of power analysis is to assist in determining the sample size before or during a study. A power analysis has the capability to detect impacts of the sample size in a certain situation and increase the reliability of statistical decisions.

G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) is a stand-alone power analysis software for various statistical tests commonly used in the social, behavioral, and biomedical sciences. In this research, the main statistical method was based on bivariate correlation, therefore, G.Power (Version 3.3.1) was used to calculate the minimum sample size based on power analysis. The calculated sample size was 29 people per group and finally, total 60 students were involved in this study which divided into 2 groups.

Exact - Correlation: Bivariate Normal model

Options: Exact distribution

Analysis: A priori: Compute required sample size

Input:	Tail(s)	= Two
	Correlation ρ H1	= 0.5
	α err prob	= 0.05
	Power (1-β err prob)	= 0.80
	Correlation ρ H0	= 0
Output:	Lower critical r	= -0.3672777
	Upper critical r	= 0.3672777
	Total sample size	= 29
	Actual power	= 0.8139420

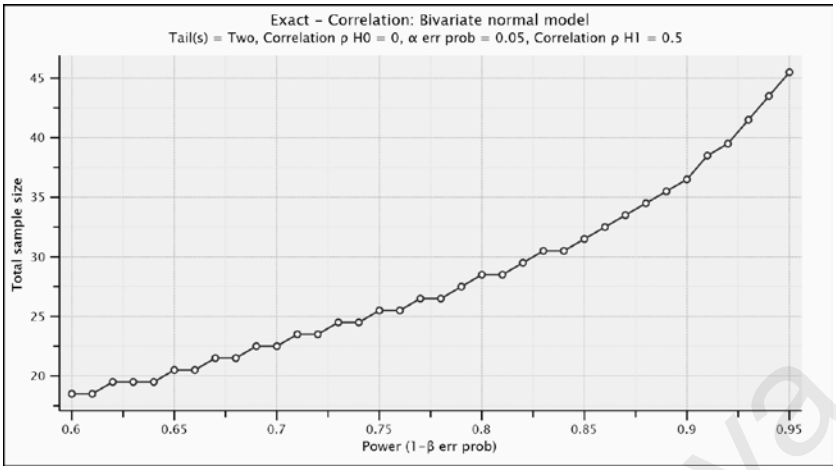

Figure 3.1: X-Y Plot for a Range of Values for Part 1.

3.1.5 Randomization.

Sixty students were involved in part 1 of the study. They were block randomized into the two 'arms' of the crossover study: Group A or B. The students who randomized into Group A answered the touch screen (the test group) followed by written and the students who in the Group B answered the written followed by the touch screen.

While the survey was conducted, the first 4 students that fulfilled the inclusion criteria were block randomized into Group A and Group B (2 students for each group). The pattern continued until the sample size reached 60. Each student was asked to complete one version of the questionnaire in the morning and the second version in the afternoon, with a target interval of at least 3 hours between the two versions. (Figure 3.2)

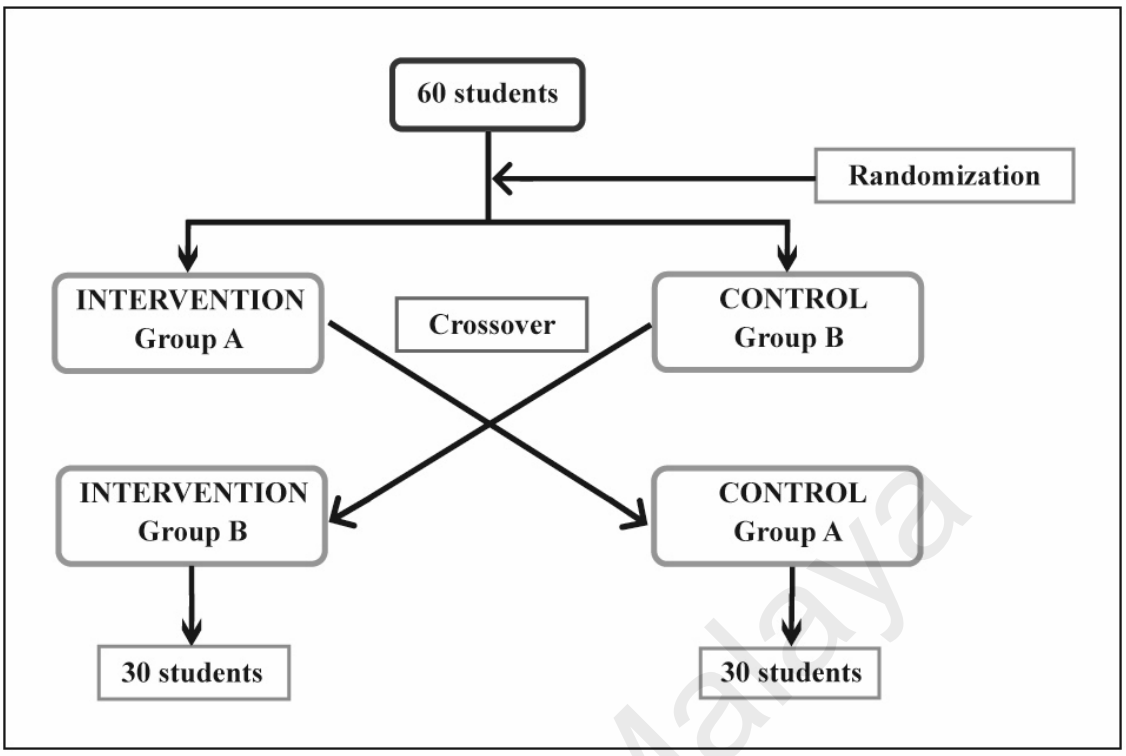

Figure 3.2: Sampling Method for Part 1.

3.1.6 Oral Health Impact Profile (OHIP-14) Questionnaires.

OHIP-14 is one of the OHRQoL instruments. In this study, the Malaysian version of OHIP-14 was used (as described in Chapter 2 the Literature review page). Two types of administration were prepared: Written Questionnaires (QBA) and Computer Touch Screen Questionnaires (CTSA).

Written Questionnaires (QBA): This is the traditional way of collecting data by using pen and paper-based questionnaire. The OHIP-14 questionnaire was printed on a paper form and the responses were entered into a database manually (Appendix B). The orders of questions were jumbled up and were made different from the CTSA to reduce the possibility of remembering the question. After the respondents submitted the responses, the data were entered into the database manually.

Computer Touch Screen Questionnaires (CTSA): The OHIP-14 questionnaire software was developed and installed into 2 units of 7 inch Samsung tablet (Appendix C). The respondents answered the questions one at a time via the tablet because only one item presented at a time and was not possible to move on to the next question without completing the previous one. After the respondents submitted the responses, the result in the form of graphs was generated automatically. The responses were compiled automatically into an Excel file that was used for data analysis directly.

3.1.7 Measurement.

Demographic data.

Respondent's demographic data were collected which includes age, gender, ethnicity, and level of education.

OHIP-14 Score.

The responses for each OHIP-14 items was measured using a simple 5 point Likert-type frequency scale; coded as 0 = never, 1 =hardly ever, 2= occasionally, 3= fairly often, and 4= very often. The score was calculated by adding up all the respondent points. Minimum score was 0 and the maximum score was 56.

Respondent's Preference.

A Respondent Feedback Form was given after the respondent completed both versions of OHIP-14 questionnaires asking the preferred mode of OHIP-14 questionnaires. The proportion of respondents stating either a preference for touch screen, written or no preference was calculated. The Respondent Feedback Form is given in Appendix D.

Respondent's Acceptance.

The touch screen questionnaires were considered as acceptable if more than 60% of the respondents either preferred them to written questionnaires or had no preference (G Velikova et al., 1999).

Feasibility.

Feasibility was measured by examining the time taken for completion of both versions of the OHIP-14 questionnaire.

3.1.8 Data Collection.

Figure 3.2 shows the data collection process. All eligible students, who fulfilled the inclusion criteria, were invited to participate in this study via advertisement posted on the official Facebook group of the Ninth Residential College. The students must be residents of the Ninth Residential College and Malaysian citizens to be eligible to participate in this study.

A total of 60 students were block randomized into 2 groups; Group A: touch screen, followed by written and Group B: written followed by the touch screen. The survey was conducted at the Sri Tanjung hall and started in the morning. During the survey, the first 4 students who came to the Sri Tanjung hall and fulfilled the inclusion criteria were block randomized into Group A and B Group B (2 students for each group). The pattern continued until the sample size reached 60. Each student was asked to complete one version of the questionnaire in the morning and the second version in the afternoon, with a target interval of at least 3 hours between the two versions (washout period).

The eligible respondents were given a Respondent Information Sheet and a brief explanation about the study was given. The respondent that agreed to participate in the

study were asked to sign an informed consent (Appendix E). The respondents answered the questionnaires according to the assigned group. After completing the first version, the respondents were asked to come again in the afternoon to complete the second version. The time taken to complete each version of the questionnaire was recorded for each mode of administration.

On completion of both versions of the questionnaires, a Respondent's Feedback Form was given to record the respondent's preferred mode of administration (touch screen, paper, or no preference). A token of appreciation was given to every respondent (Appendix F) after completing both version of questionnaires and the Respondent's Feedback Form. Data collected using written questionnaires were transferred into a computer database manually whereas data obtained using the touch screen were transferred electronically.

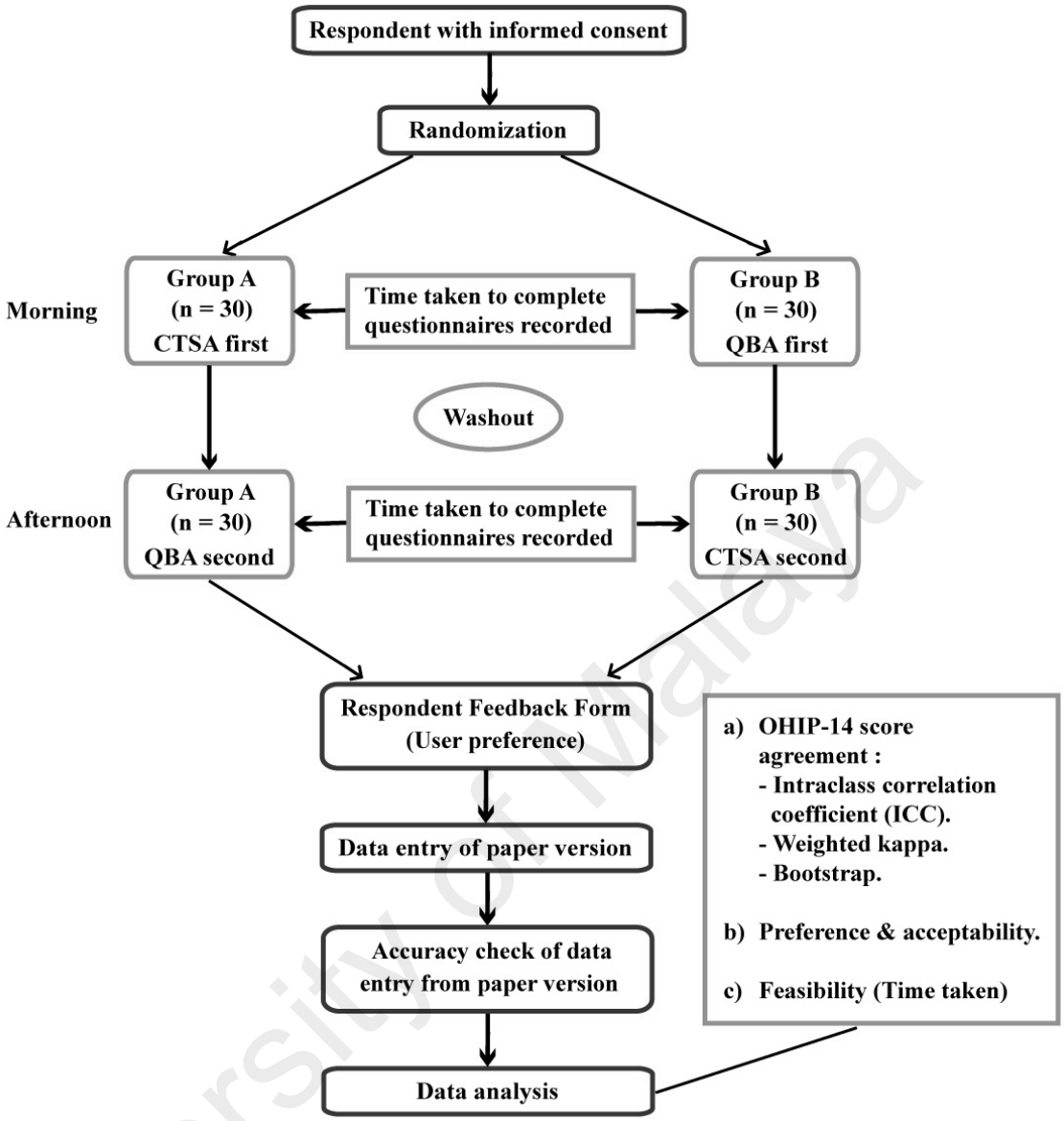

Figure 3.3: Flow Chart of Data Collection Procedure for Part 1.

3.1.9 Data Management and Analysis.

Statistical analysis was carried out using statistical packages for social sciences software (IBM SPSS, version 20, USA). Data cleaning was performed prior to data analysis. Imputation was carried out for missing data and don't know responses using the mean score of overall respondents for each of the 14 questions.

The descriptive analysis of the respondents and normality test were conducted. The normality testing was conducted using Kolmogorov-Smirnov test and Shapiro Wilk test. The descriptive analysis of the demographic profile of the respondents and difference in

OHIP-14 score obtained were assessed using frequency distribution for categorical variables and mean (standard deviation, SD) for continuous variables.

The degree of score agreement between the OHIP-14 score obtained using a computer touch screen and written questionnaires were assessed using intraclass correlation coefficient (ICC), weighted kappa, standard regression, and bootstrap regression test. ICC was utilized in this study to examine the strength of absolute agreement of total score and elements of OHIP-14 subscales between CTSA and QBA. ICC mixed model effects have been chosen because the variables were continuous, ratings were static and respondents were selected randomly. Meanwhile, the weighted kappa was used to test the validity of the ICC test results. The weighted kappa was calculated for each item in the OHIP-14 subscales. The standard and bootstrap regressions tests were conducted for the total score of QBA, age, gender and constant. The bootstrap is useful for accurate prediction, validation, forecasting and etc (Singh & Xie, 2008).

The respondent's preferences were assessed by calculating the proportions of respondents stating either a preference for touch screen, written or no preference. The proportion of respondents stating the preference for touch screen and no preference were calculated to assess the respondent's acceptance. The Wilcoxon Signed Ranks test was used to test the feasibility. The normality distribution for the time completion was checked using Kolmogorov-Smirnov test and Shapiro Wilk test.

3.2 Part 2: Comparing OHIP-14 Score between Malay Version and English.

This section answered the research question in part 2. In this part, only one research question will be discussed which are whether the OHRQoL score obtained using a web-based questionnaires for both English and Malay language were the same.

3.2.1 Study Site.

The second part of this study was also carried out at the Ninth Residential College. The data collection for the part 2 was conducted 2 weeks after the data collection for the first part had finished. Brochures regarding this study were given to the students at the Ninth Residential College 1 day prior to the data collection.

3.2.2 Study Design.

This study used a randomized crossover study design. The participants were randomized into Group A and Group B. Group A started with test followed by control after the washout period of at least 3 hours and vice versa for Group B.

Test: The respondent answered the web-based English version OHIP-14.

Control: The respondent answered the web-based Malay language version of OHIP-14.

3.2.3 Subject Recruitment.

Two weeks after the data collection for part 1 has finished, the data collection for the part 2 was conducted. Brochures regarding this study were given to the students at the Ninth Residential College 1 day prior to the data collection. Besides that, an advertisement regarding the study was also posted on the official Facebook group of the Ninth Residential College stating a brief explanation on the survey, requirement, date, time, and procedure on how to participate in the study. The survey was conducted via a web-based survey. Students who are residents of the Ninth Residential College, a Malaysian citizen, understands both English and Malay languages, and were not the

participants of part 1 of the study were eligible to participate in this study. The university students were chosen because we want to test the language used for administration, not the OHIP-14 instrument. Therefore it is important to test on a respondent that literate in both English and Malay language.

3.2.4 Sample Size.

Statistical power analysis techniques allow the quantitative researcher to decide on how large a sample is needed to enable statistical judgments that are accurate and reliable. It also helps the researcher to identify how likely is the selected statistical test to detect effects of a given size in a particular situation (StatSoft, 2003). The power of a test is the probability of correctly rejecting a false null hypothesis (Lane, 2007), thus avoiding a Type II error (beta) which sometimes considered a false negative. The power of a test equals 1 - β. One of the main functions of power analysis is to assist in determining the sample size before or during a study. A power analysis has the capability to detect impacts of the sample size in a certain situation and increase the reliability of statistical decisions.

G*Power (Faul et al., 2007) is a stand-alone power analysis software for various statistical tests commonly used in the social, behavioral, and biomedical sciences. In this research, the main statistical method was based on bivariate correlation, therefore, G.Power (Version 3.3.1) was used to calculate the minimum sample size based on power analysis. The calculated sample size was 29 people per group and finally, total 60 students were involved in this study which divided into 2 groups.

Exact - Correlation: Bivariate Normal model

Options: Exact distribution

Analysis: A priori: Compute required sample size

Input: Tail(s) = Two

Correlation ρ H1 = 0.5

α err prob = 0.05

Power (1-β err prob) = 0.80

Correlation ρ H0 = 0

Output: Lower critical r = -0.3672777

Upper critical r = 0.3672777

Total sample size = 29

Actual power = 0.8139420

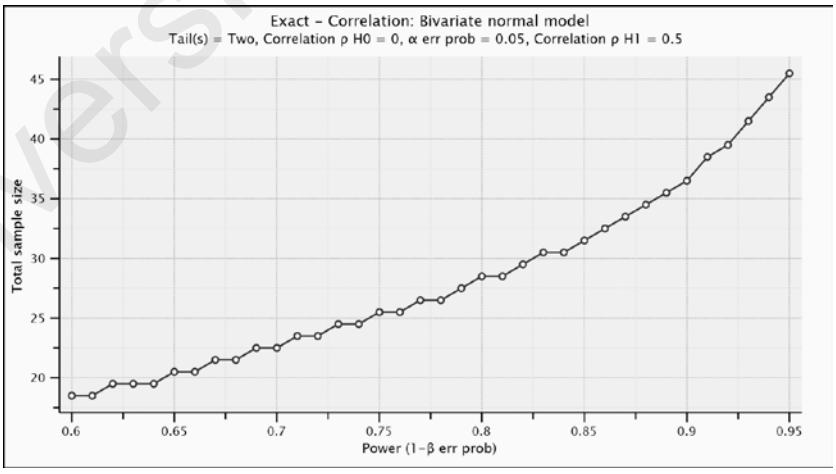

Figure 3.4: X-Y Plot for a Range of Values for Part 2.

3.2.5 Randomization.

Similar to the part 1 of the study, 60 students were randomized into 2 ‘arms’: Group A and B. The students, who in the Group A answered the web-based English version of OHIP-14 (the test group) followed by web-based Malay language version of OHIP-14 and the students who in the Group B answered the web-based Malay language version followed by web-based English version.

The web-based survey was opened according to the time and date stated in the brochures. Once the survey has opened, the first 30 students that logged into the system and fulfilled the inclusion criteria were assigned to Group A, while the next 30 students were assigned to Group B. Each student was asked to complete both versions of questionnaires with a time interval of at least 3 hours between questionnaires The second version of questionnaires was programmed to be made available to the respondents at least 3 hours after they answered the first version (Figure 3.5).

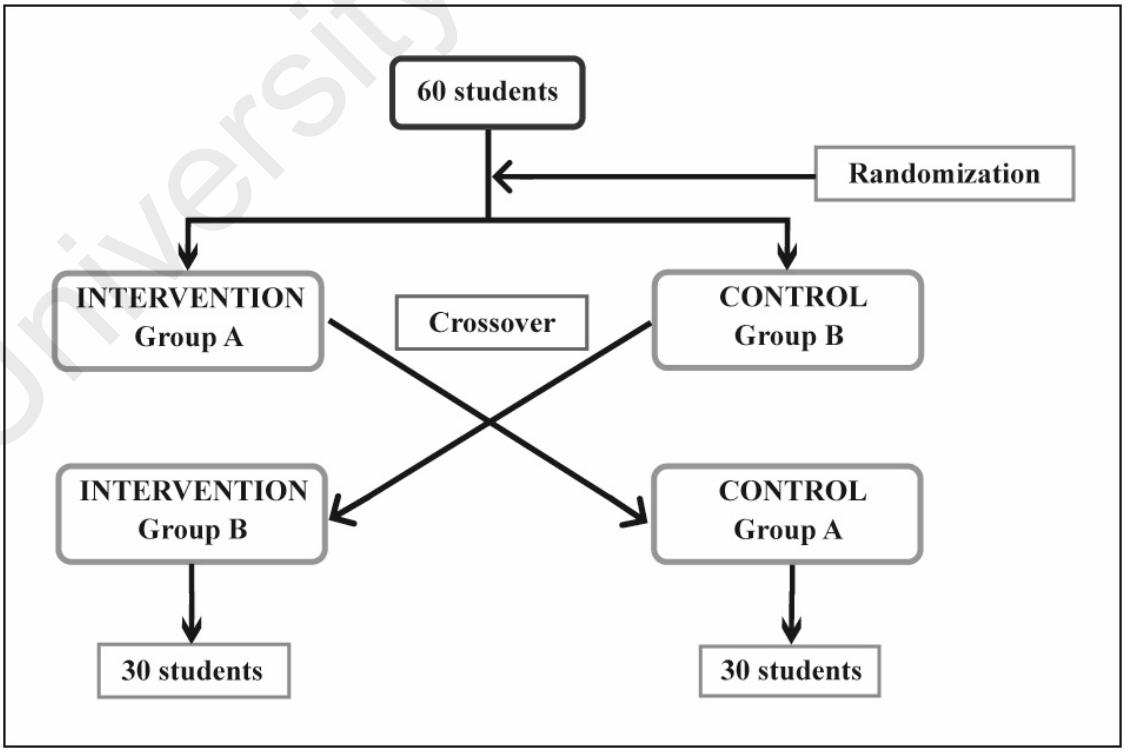

Figure 3.5: Sampling Method for Part 2.

3.2.6 Web-Based Oral Health Impact Profile (OHIP-14) Questionnaires.

OHRQoL instrument used in part 2 of this research was also the Malaysian version of the short-form Oral Health Impact Profile (OHIP-14). Two versions, English and Malay OHIP-14 were prepared. Both versions were administered using a web-based.

The web-based questionnaire or also known as an online survey is an application that is accessible only with an active internet connection and that uses HTTP as its primary communications protocol. It is also called web application. The respondents answered the questionnaire by opening the OHIP-14 questionnaire's URL (Uniform Resource Locator, also known as a web address) which is **http://ohip14.com** (Appendix G). The respondents need to log in to the **ohip14.com** website by key in their identification card number (IC number) and follow the instructions to answer the questionnaire. The respondents that have participated in part 1 of the study were excluded. They will not be able to log in and answered the questionnaires as their IC numbers that were obtained from the first part of the study has already in the database. The web-based questionnaire has been programmed to deny access to the respondents that had participated in part 1 of the study. The application recognized the respondents by their IC number.

All questions were presented on one web page and the respondent cannot submit their responses if they did not answer all the questions. All the rules, regulation and instruction were available on the website. The time taken to complete the questionnaire were taken and recorded automatically. The responses and time completion were exported into an Excel file that was used for data analysis.

The web-Based English version of OHIP-14 (EQA): The OHIP-14 questionnaire was presented on a website using the English language. All the questions are the same with the computer touch screen OHIP-14 questionnaire (CTSA) used in part 1 of the study. However, the orders of questions were jumbled up so were different from the MQA to reduce the possibility of remembering the question.

Web-Based Malay version of OHIP-14 (MQA): The OHIP-14 questionnaire was presented on a website using the Malay language. All the questions and the order of questions are the same with the computer touch screen OHIP-14 questionnaire (CTSA) used in part 1 of the study. As opposed to the CTSA, all the questions were presented in one web page.

3.2.7 Measurement.

Demographic data.

Respondent's demographic data were collected which includes age, gender, ethnicity, and level of education.

OHIP-14 Score.

The responses for each OHIP-14 items was measured using a simple 5 point Likert-type frequency scale; coded as 0 = never, 1 =hardly ever, 2= occasionally, 3= fairly often, and 4= very often. The score was calculated by adding up all the respondent points. Minimum score was 0 and the maximum score was 56

3.2.8 Data Collection.

All eligible students, who fulfilled the inclusion criteria, were invited to participate in this study via the brochures given and an advertisement posted on the official Facebook group of the Ninth Residential College. The brochures regarding this study were given to the students at the Ninth Residential College one day before the website is open to the public for online survey. Website URL, rules, and regulations were included in the brochures (Appendix H). The students must be residents of the Ninth Residential College and a Malaysian citizen to be eligible to participate in this study.

Only the first 60 respondents were allowed to log in and participate in this study. The first 30 respondents who log in were assigned into the Group A and answered the English version of OHIP-14 followed by Malay version of OHIP-14. The remaining 30 respondents who log in were assigned into the Group B and answered the Malay version followed by the English version. Each respondent completed each version in 2 separate sessions with a time interval of at least 3hours between the two questionnaires (washout period).

After successfully logged in, a Respondent Information Sheet was displayed. If the respondent agreed to participate in this study, they were needed to fill in a consent form before answering the questionnaire. After completing the first version of the questionnaire, the respondents need to wait for at least 3 hours to answer the second version. They were not able to answer the second version if they log in before the time limit.

After the survey has completed, a unique code will be generated automatically. The respondents must show the unique code during the honorarium redemption to redeem the token of appreciation. The respondents were asked to sign an informed consent document during honorarium redemption for hardcopy purposes. The responses and

time completion were exported into an Excel file that can be used for data analysis. Figure 3.6 showed the flow of the data collection for Part 2.

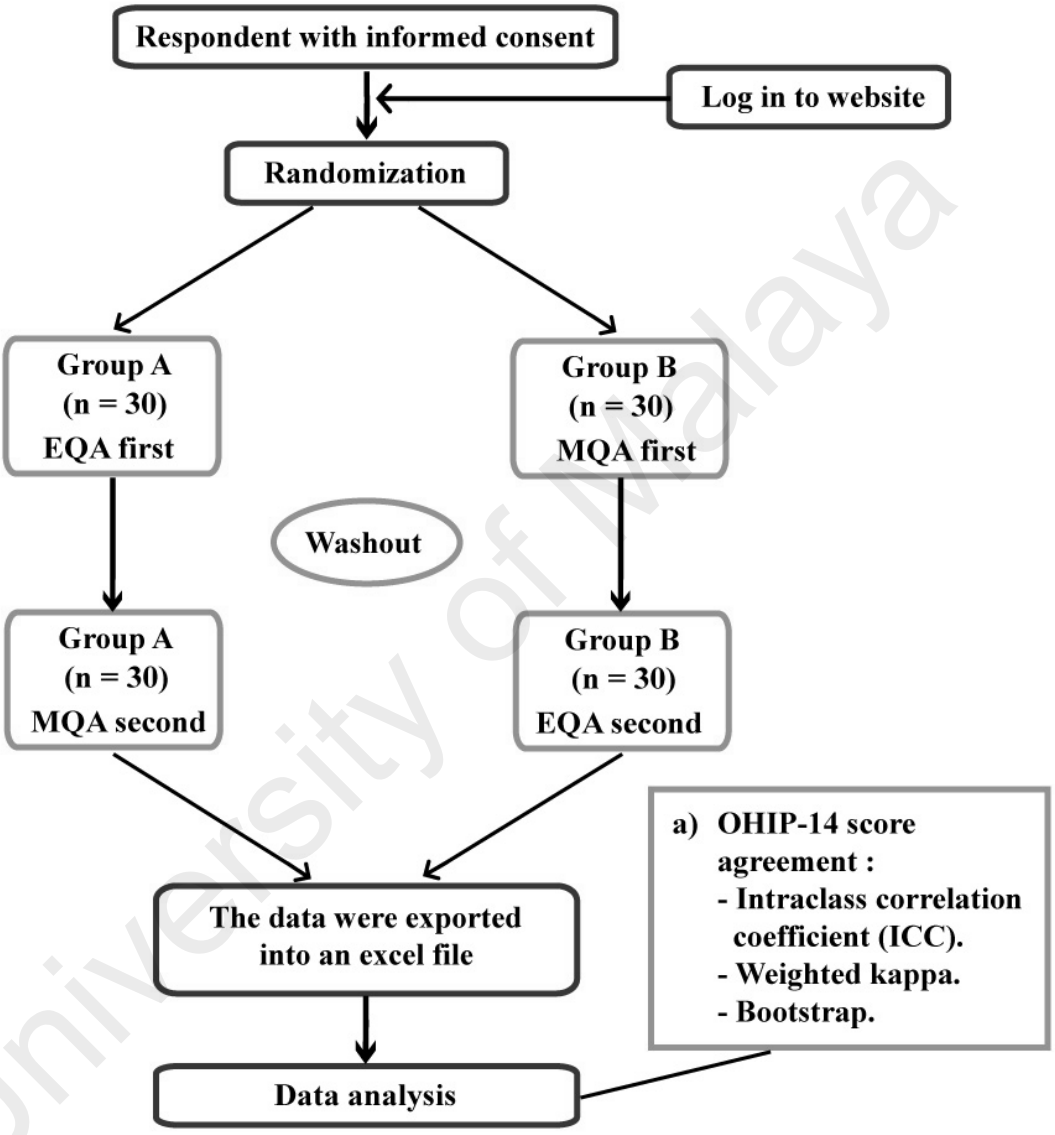

Figure 3.6: Flow Chart of Data Collection Procedure for Part 2.

3.2.9 Data Management and Analysis.

Statistical analysis was carried out using statistical packages for social sciences software (IBM SPSS, version 20, USA). Data cleaning was performed prior to data analysis. Imputation was carried out for missing data and don't know responses using the mean score of overall respondents for each of the 14 questions.

The descriptive analysis of the respondents and normality test were conducted. The normality testing was conducted using Kolmogorov-Smirnov test and Shapiro Wilk test. The descriptive analysis of the demographic profile of the respondents and difference in OHIP-14 score obtained were assessed using frequency distribution for categorical variables and mean (SD) for continuous variables. The Wilcoxon Signed Ranks test was also conducted to evaluate the difference in OHIP-14 score obtained via both languages.

The degree of score agreement between the OHIP-14 score obtained using the Malay language and the English version was assessed using intraclass correlation coefficient (ICC), weighted kappa, standard regression, and bootstrap regression test. ICC was utilized in this study to examine the strength of absolute agreement of total score and elements of OHIP-14 subscales between EQA and MQA. ICC mixed model effects have been chosen because the variables were continuous, ratings were static and respondents were selected randomly. Meanwhile, the weighted kappa was used to test the validity of the ICC test results. The weighted kappa was calculated for each item in the OHIP-14 subscales. The standard and bootstrap regressions tests were conducted for the total score of the English version, age, gender and constant. The bootstrap is useful for accurate prediction, validation, forecasting and etc.

3.3 Pre-test Stage.

The pre-test was carried out at the Postgraduate Clinic, Faculty of Dentistry, University Malaya. This stage is aimed to i) to assess the feasibility of using the computerized intervention instrument, ii) to assess the patient's preference mode of administration of questionnaires.

The methodology of the Pre-Test Stage.

This pre-test was carried out at the Postgraduate Clinic, Faculty of Dentistry, University Malaya. Patients aged 18 and older who visited the Postgraduate Clinic were eligible to participate in this pre-test.

Data Collection.

1. Twenty patients who visited the clinic were randomly selected to participate in this pre-test.
2. Each patient was asked to complete the Malaysian version of the short-form Oral Health Impact Profile (S-OHIP [M]) questionnaires on the touch screen tablet.
3. On completion of the questionnaires, the patient was given a Feedback form which consists of 7 questions that used to assess the patient's preference and opinion on the computerized S-OHIP [M].
4. Data collected using the written Feedback form were transferred into a computer database manually and analyzed using SPSS.

CHAPTER 4: RESULTS

4.1 Introduction.

This chapter presents the results of this study, which are presented in two separate sections. The first section focused on the results for part 1 of the study which is comparability of OHIP-14 score obtained using touch screen versus written questionnaire. Meanwhile, the results for part 2 of the study which is comparability of OHIP-14 score obtained using English versus the Malay language presented in the second section.

4.2 Comparison of the Touch Screen and Written OHIP Questionnaires.

The aim of part 1 was to investigate the comparability of OHIP-14 obtained via computer touch screen versus those obtained via written questionnaires. The investigation involved the evaluation of score difference, the degree of score agreement, acceptance and preference mode of administration and feasibility of both modes.

4.2.1 Demographic Profile of Respondents.

Table 4.1 shows the demographic profile of the respondents participated in part 1 of the study. There were 60 students eligible to participate during the study period. However, one respondent was eliminated because the respondent did not answer 5 out of 14 questions in the written questionnaire. Therefore, the total number of respondents was 59 students (98.3%) which comprised of 29 male students and 30 female students. The mean age of the respondents was 21.36 (SD 1.31) years and it was not normally distributed. The total numbers of respondents were almost equal for both gender and most of them were Malay.

Table 4.1: Demographic Profile of Respondents for Part 1.

Demographic profile	Frequency (%)	Mean ± SD
Age	-	21.36 ± 1.31
No. of respondents	59 (98.3)	-
Gender		
Male	29 (49.2)	-
Female	30 (50.8)	-
Ethnicity		
Malay	41 (69.5)	-
Chinese	12 (20.3)	-
Indian	2 (3.4)	-
Other	4 (6.8)	-

4.2.2 Evaluation of Difference in OHIP-14 Score.

Table 4.2 shows the descriptive analysis of OHIP-14 score obtained via touch screen (CTSA) and written questionnaire (QBA). The mean and SD of the total OHIP-14 score for both CTSA and QBA showed almost consistent results. On the other hand, the mean and SD of the score for OHIP-14 subscales showed some variation. Some subscales showed consistency, but some have a slight contradiction. The subscales *Physical disability* and *Psychological disability* showed a slight contradiction based on the mean and SD of both CTSA and QBA.

Table 4.2: Descriptive Analysis of OHIP-14 Score between CTSA and QBA.

Domain	CTSA	QBA
	Mean ± SD	**Mean ± SD**
Total OHIP-14 Score.	12.97± 7.77	13.20 ± 7.53
Subscales.		
F*unctional limitation.*	2.37 ± 1.46	2.15 ± 1.75
Physical pain.	2.56 ± 1.50	2.56 ± 1.50
Psychological discomfort.	2.93 ± 1.75	2.59 ± 1.23
Physical disability.	1.41 ± 1.51	1.76 ± 1.55
Psychological disability.	1.47 ± 1.83	1.73 ± 1.79
Social disability.	0.92 ± 1.37	1.00 ± 1.19
Handicap.	1.31 ± 1.38	1.41 ± 1.52

4.2.3 Evaluation of Score Agreement.

The degree of score agreement was assessed using intraclass correlation coefficient (ICC), weighted kappa, standard regression, and bootstrap test. All the test results are presented below.

4.2.3.1 Intraclass Correlation Coefficient between Total and Subscales Score.

Table 4.3 shows the values of intraclass correlation coefficients (ICC) between CTSA and QBA. ICC was utilized in this study to examine the strength of absolute agreement of total score and the OHIP-14 subscales between CTSA and QBA. It has reflected a higher level of agreement between the subscales on the written questionnaire and touch screen by the selected respondents. The ICC for the total score between

CTSA and QBA was 0.90 (95% CI = 0.83-0.94). This implies a strong agreement when 0.70 is considered to be a minimum standard for the ICC (de Vet, Terwee, Knol, & Bouter, 2006) The ICCs for the subscales captured the dual mixture of moderate and high values of agreements between the mode of administrations. Based on Table 4.3, it was observed that *Physical pain*, *Physical disability*, *Psychological disability*, *Social disability*, and *Handicap* showed a strong agreement, while for other subscales which are the *Functional limitation* and *Psychological discomfort* showed a moderate agreement.

Table 4.3: Intraclass Correlations (ICC) of OHIP-14 Score between CTSA and QBA.

Score	ICC	(95% CI)
Total Score.	0.90**	(0.83-0.94)
Subscales.		
F*unctional limitation.*	0.64**	(0.40-0.79)
Physical pain.	0.86**	(0.76-0.92)
Psychological discomfort.	0.58**	(0.30-0.75)
Physical disability.	0.77**	(0.61-0.86)
Psychological disability.	0.90**	(0.83-0.94)
Social disability.	0.85**	(0.75-0.91)
Handicap.	0.84**	(0.72-0.90)

** Significant agreement at 5%.

4.2.3.2 Weighted Kappa.

Table 4.4 presents the agreement of OHIP-14 between CTSA and QBA for each item in the subscales.. The kappa-statistic measure of agreement is scaled to be 0 (the amount of agreement is expected to be observed by chance) and 1 (perfect agreement). The weighted kappa for all items, except item "*Difficulty chewing any foods", "Had bad breath caused by the dental problem", and "Felt discomfort due to food stuck", range from 0.41 to 0.78,* which indicates a moderate to substantial agreement according to the guidelines of Landis and Koch (Landis & Koch, 1977). The item "*Difficulty chewing any foods", "Had bad breath caused by the dental problem",* and "*Felt discomfort due to food stuck"* has a fair agreement which were range from 0.25 to 0.39 (Landis & Koch, 1977).

Both items in the *Functional limitations* subscale showed a fair agreement between CTSA and QBA. Meanwhile, for the subscales of *Psychological discomfort*, the responses on both items showed a mixture of a fair and moderate agreement. A balanced momentum of moderate agreement between the responses can be seen on both items in the *Physical disability* and *Social disability.* A mixture of moderate and substantial agreement between both modes can be seen in the *Physical* pain, *Psychological disability*, and *Handicap* subscale.

Table 4.4: Agreement of OHIP-14 Score between CTSA and QBA for Each Item.

Item	Weighted Kappa
F*unctional limitation.*	
Difficulty chewing any foods.	0.39**
Had bad breath caused by the dental problem.	0.38**
Physical pain.	
Discomfort eating any foods.	0.44**
Had ulcer in the mouth.	0.72**
Psychological discomfort.	
Felt discomfort due to food stuck.	0.25**
Felt shy.	0.42**
Physical disability.	
Avoided eating certain foods.	0.41**
Avoided smiling.	0.56**
Psychological disability.	
Your sleep been disturbed.	0.67**
Your concentration been disturbed.	0.57**
Social disability.	
Avoided going out.	0.59**
Experienced problems carrying out daily activities.	0.46**
Handicap	
Had to spend a lot of money.	0.78**
Felt less confident of yourself.	0.45**

**Significant agreement at 5%.

4.2.3.3 Standard and Bootstrap Regressions.

Table 4.5 shows the results of standard and bootstrap versions of regressions. It shows that the current mode of sample size is valid and useful for generalization. It is proven that total score QBA has a very strong and significant influence in predicting the variations within total score CTSA. It confirms the effective mode of absolute agreement between the total scores via touch screen and written versions of the questionnaire. It has the explanatory power of 65%.

The standard and bootstrap regression were used to see whether by excluding the probable effect of age and gender as 2 confounding variables showed any correlation between total score of both QBA and CTSA. Based on the p-value for both age and gender, it showed that these 2 variables are not significantly contributing or confounding variable. It means that the agreement between QBA and CTSA is not affected by the age and gender and proved that the agreement between QBA and CTSA is high regardless of gender and age.

Table 4.5: Results of Standard and Bootstrap Versions of Regressions for Part 1.

	Standard Regression			**Bootstrap Regression (S.R=100)**		
	SE	**P-value[a]**	**CI (95%)**	**SE**	**P-value[b]**	**CI (95%)**
Total Score QBA	0.08	< 0.01**	0.68-1.00	0.06	< 0.01	0.72-0.97
Age	0.53	0.85	-2.12	0.6	0.87	-2.37
Gender	1.37	0.79	-5.49	1.48	0.8	-5.79
Constant	10.46	0.98	-41.93	11.82	0.98	-46.34

[a] Obtained from t-test.
[b] Obtained from z-test.
** Significant at 5%. The dependent variable is generated total score of responses from touch screen.

4.2.4 Acceptability and Preferred Mode of Administration.

Figure 4.1 shows that among the respondents, 48 respondents (78%) preferred the touch screen version, 10 (16.9%) respondents preferred the written version, and 3 (5.1%) respondents had no preference. Overall, 83.1 % of the respondents found that the touch screen questionnaire acceptable (either preferred or as acceptable as the written questionnaire). It shows that majority respondents have favored the significant role of technology in highlighting the level of convenience and comfortableness in dealing with the OHIP-14. This can be seen through the fact that 78 percent of respondents have selected touch screen as the significant preferred mode of administration rather than written version.

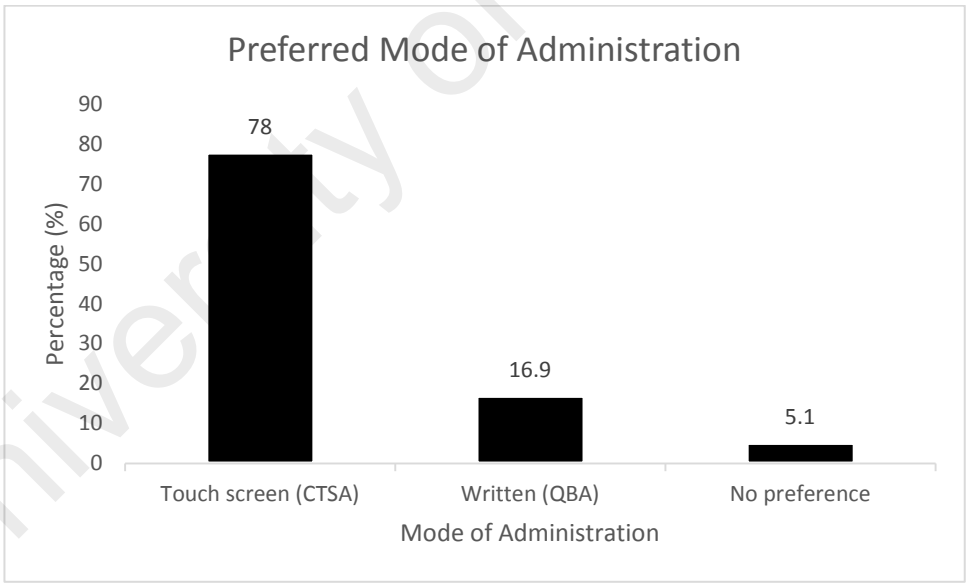

Figure 4.1: Preferred Mode of Administration.

4.2.5 Feasibility of Computerized Versus Written Questionnaires.

Table 4.6 presents the results on the mean differences between the time completion of CTSA and QBA and the normality testing. Wilcoxon Signed Ranks test was used to test the feasibility. The mean (SD) time spent completing the CTSA was 172.19 seconds (38.90) and on QBA was 163.46 seconds (60.79). Respondents took slightly less time to answer the written questionnaire compared to CTSA but the difference is not statistically significant. From Table 4.6, it shows that the catch-up effects within touch screen and written questionnaire are similar via time taken for completion. It has indicated the significant interest in the dual interactions between humans and technology and the attentive nature of the participants. There were some contradictions in the results of normality testing for completion time for both CTSA and QBA. The normality testing results for the time completion showed that they were normally distributed. However, when the results of both tests were validated through Normal Q-Q plot, it showed that they were not normally distributed.

Table 4.6: Results on Normality Testing and the Mean Differences between the Time Completion of CTSA and QBA.

	Mean (seconds)	**SD (seconds)**	**P-value [a]**	**Normality Testing**	
				Kolmogorov-Smirnov	**Shapiro Wilk**
CTSA	172.19	38.9		0.08(0.20)	0.97(0.14)
QBA	163.46	60.79	0.09	0.09(0.20)	0.94(0.01)**

[a] Obtained from Wilcoxon Signed Ranks test.
** Variables are not normally distributed at 5%.

4.3 Comparing OHIP-14 English Version versus Malay Version.

The objective of part 2 was to investigate the agreement between the OHIP-14 score obtained using a web-based questionnaire for different languages. In this part, only the score agreement between English and Malay language questionnaire administered using web-based were assessed. The respondent's acceptance, preference mode of administration, and feasibility testing were not conducted in this part.

4.3.1 Demographic Profile of Respondents.

The demographic profile of the respondents participated in part 2 of the study are listed in Table 4.7. The total respondents that were eligible to participate during the study period were 60 students (100%) and no respondent was eliminated. There was no missing data or incomplete responses because the web-based questionnaire has been programmed that the respondents cannot skip any question. The respondent cannot submit their responses if they were not answering all the questions. The mean age of the respondents was 22.0 (SD 1.18) years and it was not normally distributed. There were 48 (80%) female and 12 (20%) male participated in this part 2 of the study. In general, it can be inferred that females have dominated the game of participation due to their nature of willingness and attentive attitude in giving responses. There were significant different in the ethnicity in the context of active participation. Malay has a greater initiative in participating in this study (70.0%), followed by Chinese (18.3%), Indian (6.7%), and other ethnicity (5.5%).

Table 4.7: Demographic Profile of Respondents for Part 2.

Demographic profile	Frequency (%)	Mean ± SD
Age	-	22.00 (1.18)
No. of respondents	60 (100)	-
Gender		
Male	12 (20)	-
Female	48 (80)	-
Ethnicity		
Malay	42 (70.0)	-
Chinese	11 (18.3)	-
Indian	4 (6.7)	-
Other	3 (5.5)	-

4.3.2 Evaluation of Difference in OHIP-14 Score Using Different Languages.

Table 4.8 presents the descriptive analysis of OHIP-14 score between English and Malay language. Based on the mean and SD of total score and score for OHIP-14 subscales, the selection of responses between English and Malay language questionnaire were almost consistent with each other. The mean (SD) of the total OHIP-14 score for the English language was 16.02 (8.57), meanwhile, for the Malay language, the mean (SD) was 14.65 (8.64). This result shows that the mean (SD) of the total OHIP-14 score between English and Malay language questionnaire were almost consistent. On the other hand, the mean and SD of the score for OHIP-14 subscales showed some variation with several subscales showing consistency and some having a slight inconsistency.

Table 4.8: Descriptive Analysis of OHIP-14 Score between English and Malay Language.

Domain	English	Malay	Z value*	P value*
	Mean ± SD	Mean ± SD		
Total OHIP-14 Score.	16.02 ± 8.57	14.65 ± 8.64	-1.951	0.051
Subscales.				
Functional limitation.	2.40 ± 1.62	2.33 ± 1.70	-0.139	0.889
Physical pain.	2.95 ± 1.44	2.78 ± 1.46	-1.146	0.252
Psychological discomfort.	2.95 ± 1.64	3.07 ± 1.66	-0.625	0.532
Physical disability.	2.27 ± 1.61	1.97 ± 1.81	-2.133	0.033
Psychological disability.	1.98 ± 1.46	1.73 ± 1.47	-1.508	0.131
Social disability.	1.50 ± 1.43	1.10 ± 1.35	-2.642	0.008
Handicap.	1.97 ± 1.67	1.67 ± 1.72	-2.472	0.013

* Obtained from Wilcoxon Signed Ranks test.

4.3.3 Evaluation of Score Agreement Using Different Languagess.

The score agreement for part 2 of the study was assessed using the same statistical tests as in part 1 which are intraclass correlation coefficient (ICC), weighted kappa, standard regression, and bootstrap test. All the test results are presented below.

4.3.3.1 Intraclass Correlation Coefficient between Total and Subscales Score.

Table 4.9 presents the values of intraclass correlation coefficients (ICC) between English and Malay language questionnaire. ICC was utilized in this study to examine the strength of absolute agreement of total score and the OHIP-14 subscales between both languages. It has reflected a higher level of agreement between the subscales on the questionnaire in both languages by the selected respondents. The ICC value for the

total score between English and Malay language was 0.92 (95% CI = 0.87-0.96) showed that questionnaire in both languages has a very strong agreement. The ICCs for the OHIP-14 subscales captured high values of agreements between English and Malay language. Based on Table 4.9, it was observed that *Functional limitation*, *Physical pain*, *Psychological discomfort*, *Psychological disability*, and *Social disability* showed a strong agreement, while for other subscales which are *Physical disability* and *Handicap* showed a very strong agreement (de Vet, Terwee, Knol, & Bouter, 2006).

Table 4.9: Intraclass Correlations Coefficient (ICC) of OHIP-14 Score between English and Malay Language.

Score	ICC	(95% CI)
Total Score.	0.92(0.00)**	(0.87-0.96)
Subscales.		
F*unctional limitation.*	0.76(0.00)**	(0.60-0.86)
Physical pain.	0.80(0.00)**	(0.66-0.88)
Psychological discomfort.	0.80(0.00)**	(0.67-0.88)
Physical disability.	0.89(0.00)**	(0.82-0.94)
Psychological disability.	0.78(0.00)**	(0.63-0.87)
Social disability.	0.80(0.00)**	(0.65-0.88)
Handicap.	0.93(0.00)**	(0.87-0.96)

4.3.3.2 Weighted Kappa.

Table 4.10 shows the agreement of OHIP-14 score between English and Malay language for each item in the subscales that were tested using the weighted kappa test. The weighted kappa for all items, except item "*Discomfort eating any foods*", "*Avoided eating certain foods*", "*Your concentration been disturbed*" and "*Experienced problems*

carrying out daily activities", range from 0.41 to 0.72, which indicates a moderate to substantial agreement according to the guidelines of Landis and Koch . The item "*Discomfort eating any foods*", "*Avoided eating certain foods*", "*Your concentration been disturbed*" and "*Experienced problems carrying out daily activities*" has a fair agreement which were range from 0.33 to 0.39.

Based on Table 4.10, both items in the *Functional limitations* and *Psychological discomfort* reflected the significant mode of moderate agreement and it satisfied the reliability level of an absolute agreement through ICC. The item "*Had bad breath caused by the dental problem*" in the *Functional limitations* and the item "*Felt shy*" have added up values to the increment of reliability to a substantial level. Meanwhile, for the subscales of *Physical pain* and *Physical disability*, an imbalance momentum of agreement within items can be seen which comprise a mixture of a fair and substantial agreement. The item "*Had ulcer in the mouth*" in the *Physical pain* and the item "*Avoided smiling*" in the *Physical disability* have added up the values to the consistency of ICC that lead to the higher level of absolute agreement.

An imbalanced momentum of agreement can also be seen within the items in the *Psychological disability* and *Social disability,* which involved the mixture of a fair and moderate agreement. The items "*Your sleep been disturbed*" and "*Avoided going out*" have a moderate agreement that leads to the substantial level of absolute agreement. As opposed to other subscales, the *Handicap* showed a balanced momentum of agreement within the items that have a substantial agreement for both items which lead to the higher of absolute agreement.

Table 4.10: Agreement of OHIP-14 Score between English and Malay Language for Each Item.

Item	Weighted Kappa
F*unctional limitation.*	
Difficulty chewing any foods.	0.44**
Had bad breath caused by the dental problem.	0.51**
Physical pain.	
Discomfort eating any foods.	0.33**
Had ulcer in the mouth.	0.64**
Psychological discomfort.	
Felt discomfort due to food stuck.	0.41**
Felt shy.	0.50**
Physical disability.	
Avoided eating certain foods.	0.36**
Avoided smiling.	0.72**
Psychological disability.	
Your sleep been disturbed.	0.56**
Your concentration been disturbed.	0.38**
Social disability.	
Avoided going out.	0.49**
Experienced problems carrying out daily activities.	0.39**
Handicap	
Had to spend a lot of money.	0.71**
Felt less confident of yourself.	0.66**

**Significant agreement at 5%.

4.3.3.3 Standard and Bootstrap Regressions.

Table 4.11 shows the results of standard and bootstrap versions of regressions. It shows that the current mode of sample size is valid and useful for generalization. It is proven that total score English has a very strong and significant influence in predicting the variations within total score Malay. It confirms the effective mode of absolute agreement between the total scores in Malay and English. It has the explanatory power of 74%.

The standard and bootstrap regression were used to see whether by excluding the probable effect of age and gender as 2 confounding variables showed any correlation between total score of both English and Malay language. Based on the p-values for both age and gender, it shows that these 2 variables do not significantly contribute or confounding variable. It means that the agreement between English and Malay language version is not affected by the age and gender. Therefore, it is proven that the agreement between English and Malay language version is high regardless of gender and age.

Table 4.11: Results of Standard and Bootstrap Versions of Regressions for Part 2.

	Standard Regression			**Bootstrap Regression (S.R=100)**		
	SE	**P-value[a]**	**CI (95%)**	**SE**	**P-value[b]**	**CI (95%)**
Total Score English	0.07	< 0.01**	0.74-1.01	0.07	< 0.01**	0.74-1.01
Age	0.53	0.36	-0.58-1.56	0.69	0.47	-0.85-1.84
Gender	1.57	0.54	-2.17-4.14	1.57	0.54	-2.11-4.04
Constant	13.29	0.37	-38.57-14.66	17.21	0.49	-45.69-21.78

** Significant at 5%. The dependent variable is generated total score of responses from touch screen.

CHAPTER 5: DISCUSSION

5.1 Introduction.

This chapter discussed the findings of this study. The discussion was divided into two sections. The first section discussed the findings from the part 1 of the study which is comparability of OHIP-14 score obtained using touch screen versus written questionnaire. Meanwhile, the second section discussed the findings from the part 2 of the study which was the comparability of OHIP-14 score obtained using English versus Malay languages. The limitation of the study for both parts will also be discussed in this chapter.

5.2 Study Design and Sampling for Part 1 and 2.

The present study used a crossover study design, where each respondent received 2 different questionnaires which were consecutively administered during different time periods (Wellek & Blettner, 2012). The reason crossover design is chosen because it could yield a more efficient comparison of questionnaires than a parallel design since each respondent serves as their own matched control thus eliminates the order effect. Furthermore, crossover study is suitable to be used during early stages of questionnaire development and in the attempt to improve the quality of life (Bland, 2013). In addition, this study design permits opportunities to each respondent and respondents receiving multiple tests can express preferences for or against a particular mode of administration.

However, the crossover design has a disadvantage of carry-over effects. This issue is overcome by implementing at least 3 hours of washout period between questionnaires and the questions in the QBA were jumbled up to reduced learning effect. Past studies showed that human short-term memory has a limited capacity and limited duration which makes the storage very fragile and information can be lost with distraction or passage of time (Cowan, 2008; Peterson & Peterson, 1959). The human brain can hold

only about 7 pieces of information or less at a time in an active and readily-available state for 15 to 30 seconds (Atkinson & Shiffrin, 1971; Miller, 1994).

My present study allocated respondents to group A and B by using randomization. . Randomization is one of the 'gold standards' for making a comparison and has several advantages (Cochran, 1953). This technique is able to prevent the selection bias. Besides that, randomization helps balances the groups with respect to many known and unknown confounding or prognostic variables (Suresh, 2011). As a result, randomization can provide a sample that is highly representative of the population being studied. Therefore, it allows us to make generalizations from the sample to the population. In this study, block randomization was used in part 1. The respondents were block randomized for every 4 respondents into 2 respondents for each group. Only 2 units of touch screen tablet were used in this study, therefore only 2 respondents able to answer the touch screen questionnaire at one time. Block randomization was used in order to ensure that the numbers of respondents assigned to each group is equally distributed (Suresh, K. P., 2011). This sampling method is a commonly used technique in clinical trial design to reduce opportunity for bias and confounding especially when the sample size is small (Efird, J., 2010). We planned to do randomization for part 2. However, due to some difficulties during developing the coding for the process of programming the questionnaire into the website, we changed it to first come first serve basis. The first 30 respondents that logged in to the system and fulfilled the inclusion criteria were assigned to Group A, while the next 30 respondents were assigned to Group B.

5.3 Comparability of OHIP-14 Score Obtained Using Touch Screen Versus Written Questionnaires.

This section answered the research question in part 1 which are whether the OHRQoL score obtained using a computer touch screen and written questionnaires give the same score, what is the respondent's acceptance and preference mode of administration of questionnaires, and which mode of administration has higher feasibility.

5.3.1 Score Agreement between Touch Screen and Written OHIP-14 Questionnaires.

In the last two decades, the main issues related to OHRQoL measurement has been the development of their psychometric properties. A newly rising issue is the transformation of traditional paper-based OHRQoL into computerized instruments. The present study was the first to evaluate this measurement equivalence in the field of OHRQoL using the OHIP-14 instruments. There were not many studies on the touch screen or computerized questionnaires were done for Malaysian population. Most studies were focus on developing the computer application and assessing the acceptance among the healthcare provider (Ghani, A. et al., 2016; Mustaza, T. A. et al., 2016; Safina, F. et al., 2006). Among the healthcare provider point of view, the computerized questionnaire is highly accepted.

Overall the findings of the present study demonstrate a very good level of agreement between the CTSA with its traditional written version. The ICC for total score showed a strong agreement, while for the ICCs of the subscales demonstrate a dual mixture of high and moderate agreement. The ICC for the subscale of *Psychological Discomfort* showed the lowest value compares to other subscales. This finding supported by the weighted kappa value for both items in the *Psychological Discomfort* subscales which

are *"Felt discomfort due to food stuck"* and *"Felt shy"*. Both items showed a fair agreement and among the lowest weighted kappa value compared to other items.

The weighted kappa of individual items also demonstrated a substantial to fair agreement between both CTSA and QBA. Similarly, good level of agreement between the touch screen and paper versions of other established questionnaires has also been demonstrated in the previous similar research (Y.-J. Chang et al., 2014; Chen & Li, 2010; E.-H. Lee, 2009a; F. Salaffi et al., 2013).

In addition to the finding that computerized version of questionnaires appears to be equivalent to traditional paper version, it is also showed to be more effective in terms of data collection and more practical. As this study illustrates, a major advantage of CTSA over its paper-based counterpart is the ability to collect high quality data with no missing, ambiguous or problematic responses. One respondent had to be excluded in this study because the respondent did not answer 5 out of 14 questions in the written questionnaire. The CTSA questionnaire has been programmed to allow only complete responses, thus overcome the problem of missing data. The respondents could alter a response on the CTSA by returning to the previous question, but they cannot skip any question. This feature was incorporated in CTSA because we wanted the computerized version to be as close as possible to the original written questionnaire. The second advantage of CTSA is that the data is automatically entered into the database immediately after the respondents answered the questionnaire, thus eliminating manual entry or scanning of individual questionnaire into the database which avoids the data errors that can occur during manual data entry.

5.3.2 Preference and Acceptability.

In this study, 78% of the respondents preferred to use the CTSA over the QBA. The CTSA questionnaires are well accepted by the respondents, with 83.1% either preferring it to the written questionnaires or having no preference. It shows that majority of respondents have favored the significant role of technology in highlighting the level of convenience and comfortableness in dealing with the OHIP-14 questionnaires. This finding is consistent with other studies that found that the respondents preferred a computerized mode over a paper-based mode of administration (Bent et al., 2005; E.-H. Lee et al., 2014; Ribeiro et al., 2010).

Based on previous research, the respondent's acceptability and preference for mode of administration were not related to gender, age or computer experience (Y.-J. Chang et al., 2014; Gudbergsen et al., 2011; F. Salaffi et al., 2013). The preference towards computerized questionnaire has been reported as highly acceptable and it rated significantly higher for ease of use even if the respondents have rarely or never used a personal computer (Greenwood et al., 2006; Pouwer et al., 1997; F Salaffi et al., 2009). Even elderly patients preferred computerized questionnaire and find this mode of administration as easier to use (W. Larsson, 2006) and well accepted (Ali, N. M. et al., 2012).

5.3.3 Feasibility.

Based on previous research, the results of completion time for computerized and paper-based mode of administration have been controversial, with some found that the time completion for the computerized version was shorter and others discovered the opposite result (Goodhart et al., 2005; E.-H. Lee, 2009b; Ring et al., 2006). Before conducting this study, it was expected that the time taken to complete the computerized version would be shorter than the written version. However, the respondents took

slightly less time to answer the QBA questionnaire compared to CTSA but the difference is not significantly different. This finding has similar results with previous research that showed there was no significant difference in time completion for both mode of administration (Bent et al., 2005; Greenwood et al., 2006; E.-H. Lee et al., 2014).

This might be due to the use of a different presentation. In the QBA version, a total of 14 items are presented on four pages, whereas in CTSA each question is presented on a separate screen. The time completion may be influenced by the different layout and presentation of the questionnaire. In the CTSA, each respondent had to touch a navigation button at the bottom of the screen at least 14 times to move the screen between every question. This factor may have added a movement requirement that could have negated a putative time saving related to using a computer to complete the questionnaire. Another reason that may be contributing to this finding is the unfamiliarity with a computerized questionnaire that caused respondents tend to read items more carefully when using a computerized questionnaire than with the written questionnaire (E.-H. Lee, 2009a).

However, if the aspect of data entry and final results were taken into account, the CTSA required a shorter period of time to accomplished compare to the QBA version. With the use of computerized questionnaires, the time cost of data entry, editing and verifying data were eliminated and the data can be transferred directly to the final computer database, thus allowing immediate printing out and use of the results.

5.4 Comparing OHIP-14 English Version versus Malay Version.

This section answered the research question in part 2. In this part, only one research question will be discussed which are whether the OHRQoL score obtained using a web-based questionnaires for both English and Malay language were the same.

5.4.1 Score Agreement.

The results of this part 2 of the study clearly indicated a high level of agreement between English and Malay version of OHIP-14 questionnaires. Both questionnaires were collected using web-based survey. The ICC and weighted kappa values between both versions were generally good. The ICCs demonstrated an excellent agreement between the English and Malay version at both the total score and subscale levels. This finding is similar with one of the study involving Malaysia population which showed a high level of agreement for all scales for both English and Malay language questionnaires (Lim, T. O. et al., 2003). Meanwhile, for individual items, the weighted kappa showed a substantial to fair agreement between both English and Malay version of OHIP-14 questionnaires. Based on the previous research, most of the studies found that the scores agreement between different language questionnaires were equivalent and showed moderate to high agreement (Yu, J. et al., 2003; Thumboo, J. et al., 2002; Cheung, Y. T. et al., 2013).

However, the descriptive analysis of the OHIP-14 score between both versions showed that the mean for total score was almost consistent but some variation of the mean can be seen in several subscales. Several subscales showed consistency and some have a slight inconsistency. Even though the agreement is high but the mean level of the subscales estimated higher when the respondent used the English version. This may lead to over estimation. This condition may have happened as 70% of the respondents were Malay and the Malay language is their mother tongue. Hence, they understand the

Malay language better than English. Based on the descriptive analysis, 3 out of 7 subscales in the mean scores of the Malay version are significantly lower than the English version. It seems that this level is much closer to the real condition due to the fact that Malay understand the Malay language better than English since the Malay language is the national official language in the Malaysia and a compulsory language to be learned in school (Lim, 2008), thus mostly used as the medium of instruction (Omar, 1987).

5.5 Limitations of Study.

There are several limitations of this study for both part 1 and 2 of the study. These limitations can be taken into account in order to improve future research.

5.5.1 Limitations of the Study for Part 1.

The limitation of the study for part 1 is the 3 hours washout period that can be considered as short between administration of the CTSA and QBA version. This may have allowed a memory effect to contribute to the agreement between both modes of administration. However, respondents did not have access to their scores from the first questionnaires when answering the second questionnaires and the questions in the written questionnaires had been jumbled up to reduce the memory bias.

5.5.2 Limitations of the Study For fart 2.

The limitations of the study for part 2 are listed below:

a) The 3 hours washout period that can be considered as short between administration of the English and Malay language version. This may have allowed a memory effect to contribute to the agreement between both modes of administration. However, respondents did not have access to their scores from the first questionnaires when answering the second questionnaires and the question in the Malay language questionnaires had been jumbled up to reduce the memory bias.

b) The availability of internet connection may affect the survey since in this part of the study, the online survey was used. Some students that do not have internet connection may not able to participate in this study.

CHAPTER 6: CONCLUSION & RECOMMENDATIONS

6.1 Introduction.

This chapter provides the conclusion and recommendation based on the findings of this study.

6.2 Conclusion.

Based on the results of this study, the following conclusions are drawn:

a) The computer touch screen questionnaires (CTSA) provide data that are equivalent to its original written questionnaires (QBA). The results collected from both methods showed a very strong agreement.
b) The CTSA version is well accepted and has been chosen as the preferred mode of administration by the respondents over the written questionnaire.
c) There is no significant difference found for the time completion for both CTSA and QBA.
d) The level of agreement between English and Malay version of OHIP-14 questionnaires showed a very high level of agreement.

6.3 Recommendations for Future Research.

Based on the findings, several recommendations can be proposed for future research such as follows:

a) Conduct an acceptability study on dentists in using the CTSA in dental practice.

b) The additional analysis regarding the cost effectiveness of both modes of administration should be added. The cost associated with using the CTSA and QBA were not evaluated or compared in this present study. It is expected that the establishment of the CTSA could save money and time because it would eliminate the need for manual data entry and eliminate the manpower required to administer, collecting data, entering data and analyzed an OHRQoL questionnaire.

REFERENCES

Aaronson, N. K. (1988). Quantitative issues in health-related quality of life assessment. *Health Policy, 10*(3), 217-230.

Ali, N. M., Shahar, S., Kee, Y. L., Norizan, A. R., & Noah, S. A. M.(2012). Design of an interactive digital nutritional education package for elderly people. Informatics for health and social care, 37(4), 217-229.

Anderson, R. T., Aaronson, N., Leplege, A., & Wilkin, D. (1996). International use and application of generic health-related quality of life instruments. *Quality of life and pharmacoeconomics in clinical trials. 2nd edn. Philadelphia: Lippincott-Raven*, 613-632.

Allen, P. F., & Locker, D. (1997). Do item weights matter? An assessment using the oral health impact profile. *Community dental health, 14*(3), 133-138

Armellini, D. B., Heydecke, G., Witter, D. J., & Creugers, N. (2007). Effect of removable partial dentures on oral health-related quality of life in subjects with shortened dental arches: a 2-center cross-sectional study. *The International journal of prosthodontics, 21*(6), 524-530.

Ashley, L., Keding, A., Brown, J., Velikova, G., & Wright, P. (2013). Score equivalence of electronic and paper versions of the Social Difficulties Inventory (SDI-21): a randomised crossover trial in cancer patients. *Quality of Life Research, 22*(6), 1435-1440.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control processes of short-term memory. Scientific American, 225 (1971), pp. 82–90.

Avis, N. E., Assmann, S., Kravitz, H., Ganz, P., & Ory, M. (2004). Quality of life in diverse groups of midlife women: assessing the influence of menopause, health status and psychosocial and demographic factors. *Quality of Life Research, 13*(5), 933-946.

Bae, K. H., Kim, H. D., Jung, S. H., Park, D. Y., Kim, J. B., Paik, D. I., & Chung, S. C. (2007). Validation of the Korean version of the oral health impact profile among the Korean elderly. *Community Dentistry and Oral Epidemiology, 35*(1), 73-79.

Bennadi, D., & Reddy, C. (2013). Oral health related quality of life. *Journal of International Society of Preventive & Community Dentistry, 3*(1), 1.

Bennett, A. V., Jensen, R. E., & Basch, E. (2012). Electronic patient-reported outcome systems in oncology clinical practice. *CA: a cancer journal for clinicians, 62*(5), 336-347.

Bent, H., Ratzlaff, C. R., Goligher, E. C., Kopec, J. A., & Gillies, J. H. (2005). Computer-administered bath ankylosing spondylitis and Quebec Scale outcome questionnaires for low back pain: agreement with traditional paper format. *The Journal of rheumatology, 32*(4), 669-672.

Bezjak, A., Ng, P., Skeel, R., Depetrillo, A., Comis, R., & Taylor, K. (2001). Oncologists' use of quality of life information: results of a survey of Eastern Cooperative Oncology Group physicians. *Quality of Life Research, 10*(1), 1-14.

Bland, M. (2013). Cross-over trials. Retrieved from https://www-users.york.ac.uk/~mb55/msc/trials/cross.htm

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of public health, 27*(3), 281-291.

Broering, J. M., Paciorek, A., Carroll, P. R., Wilson, L. S., Litwin, M. S., & Miaskowski, C. (2014). Measurement equivalence using a mixed-mode approach to administer health-related quality of life instruments. *Quality of Life Research, 23*(2), 495-508.

Cella, D. F. (1995). Methods and problems in measuring quality of life. *Supportive Care in Cancer, 3*(1), 11-22.

Cella, D. F., & Bonomi, A. E. (1995). Measuring quality of life: 1995 update. *Oncology (Williston Park, NY), 9*(11 Suppl), 47-60.

Chang, A. M., Chau, J. P., & Holroyd, E. (1999). Translation of questionnaires and issues of equivalence. *Journal of advanced nursing, 29*(2), 316-322.

Chang, C.-H. (2007). Patient-reported outcomes measurement and management with innovative methodologies and technologies. *Quality of Life Research, 16*(1), 157-166.

Chang, Y.-J., Chang, C.-H., Peng, C.-L., Wu, H.-C., Lin, H.-C., Wang, J.-Y., . . . Liang, W.-M. (2014). Measurement equivalence and feasibility of the EORTC QLQ-PR25: paper-and-pencil versus touch-screen administration. *Health Qual Life Out, 12*, 1-10.

Chen, T., & Li, L. (2010). Pilot study on equivalence between the electronic and paper version of the Chinese SF-36. *Journal of Happiness Studies, 11*(2), 151-161.

Cheung, Y. B., Thumboo, J., Goh, C., Khoo, K. S., Che, W., & Wee, J. (2004). The equivalence and difference between the English and Chinese versions of two major, cancer-specific, health-related quality-of-life questionnaires. *Cancer, 101*(12), 2874-2880.

Cheung, Y. B., & Thumboo, J. (2006). Developing health-related quality-of-life instruments for use in Asia. *Pharmacoeconomics, 24*(7), 643-650.

Cheung, Y. T., Lim, S. R., Shwe, M., Tan, Y. P., & Chan, A. (2013). Psychometric properties and measurement equivalence of the English and Chinese versions of the functional assessment of cancer therapy-cognitive in Asian patients with breast cancer. *Value in Health, 16*(6), 1001-1013.

Choy, S. C., Goh, P. S. C., & Liew, Y. L. (2015). Learner Awareness Levels Questionnaire: Development and Preliminary Validation of the English and Malay Versions to Measure How and Why Students Learn. *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 9*(10), 3232-3236.

Cochran, W. G. (2007). *Sampling techniques*. John Wiley & Sons.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in brain research, 169*, 323-338.

Davis, P. (1976). Compliance structures and the delivery of health care: the case of dentistry. *Social Science & Medicine (1967), 10*(6), 329-337.

De Oliveira, C., & Sheiham, A. (2004). Orthodontic treatment and its impact on oral health-related quality of life in Brazilian adolescents. *Journal of Orthodontics, 31*(1), 20-27.

de Vet, H. C., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of clinical epidemiology, 59*(10), 1033-1039.

Desai, R., Durham, J., Wassell, R. W., & Preshaw, P. M. (2014). Does the mode of administration of the Oral Health Impact Profile-49 affect the outcome score? *Journal of dentistry, 42*(1), 84-89.

Detmar, S. B., Muller, M. J., Schornagel, J. H., Wever, L. D., & Aaronson, N. K. (2002). Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. *Jama, 288*(23), 3027-3034.

Dunnell, K., & Cartwright, A. (1972). *Medicine takers, prescribers, and hoarders*: Routledge/Thoemms Press.

Drasgow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology, 70*(4), 662.

Efird, J. (2010). Blocked randomization with randomly selected block sizes. International journal of environmental research and public health, 8(1), 15-20.

Ekanayake, L., & Perera, I. (2003). Validation of a Sinhalese translation of the Oral Health Impact Profile-14 for use with older adults. *Gerodontology, 20*(2), 95-99.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191.

Fayers, P., & Machin, D. (2013). *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes*: John Wiley & Sons.

Fernandes, M. J., Ruta, D. A., Ogden, G. R., Pitts, N. B., & Ogston, S. A. (2006). Assessing oral health-related quality of life in general dental practice in Scotland: validation of the OHIP-14. *Community Dentistry and Oral Epidemiology, 34*(1), 53-62.

Ferrer, M., Alonso, J., Prieto, L., Plaza, V., Monsó, E., Marrades, R., . . . Antó, J. (1996). Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *European Respiratory Journal, 9*(6), 1160-1166.

Finegan, J. E., & Allen, N. J. (1994). Computerized and written questionnaires: Are they equivalent? *Computers in Human Behavior, 10*(4), 483-496.

Gandek, B., & Ware, J. E. (1998). Methods for validating and norming translations of health status questionnaires: the IQOLA project approach. *Journal of clinical epidemiology, 51*(11), 953-959.

Gao, F., Ng, G.-Y., Cheung, Y.-B., Thumboo, J., Pang, G., Koo, W.-H., . . . Goh, C. (2009). The Singaporean English and Chinese versions of the EQ-5D achieved

measurement equivalence in cancer patients. *Journal of clinical epidemiology, 62*(2), 206-213.

Gerson, L. (1972). Expectations of" sick role" exemptions for dental problems. *Journal of the Canadian Dental Association, 38*(10), 370-372.

Ghani, A., Mariam, S., Lim, T. W., Mustaza, T. A., Hassan, A., & Ibrahim, M. (2016). Engaging the millennial generation student with e-learning in dentistry/Siti Mariam Ab Ghani…[et al.]. International Journal on E-Learning and Higher Education, 4, 114-128.

Gift, H. C., & Atchison, K. A. (1995). Oral health, health, and health-related quality of life. *Medical care*, NS57-NS77.

Goodhart, I., Ibbotson, V., Doane, A., Roberts, B., Campbell, M., & Ross, R. (2005). Hypopituitary patients prefer a touch-screen to paper quality of life questionnaire. *Growth hormone & IGF research, 15*(6), 384-387.

Greenwood, M., Hakim, A., Carson, E., & Doyle, D. (2006). Touch-screen computer systems in the rheumatology clinic offer a reliable and user-friendly means of collecting quality-of-life and outcome data from patients with rheumatoid arthritis. *Rheumatology, 45*(1), 66-71.

Gudbergsen, H., Bartels, E. M., Krusager, P., Wæhrens, E. E., Christensen, R., Danneskiold-Samsøe, B., & Bliddal, H. (2011). Test-retest of computerized health status questionnaires frequently used in the monitoring of knee osteoarthritis: a randomized crossover trial. *BMC musculoskeletal disorders, 12*(1), 190.

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

Gundy, C. M., & Aaronson, N. K. (2010). Effects of mode of administration (MOA) on the measurement properties of the EORTC QLQ-C30: a randomized study. *Health Qual Life Outcomes, 8*, 35.

Gwaltney, C. J., Shields, A. L., & Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value in Health, 11*(2), 322-333.

Handa, V. L., Barber, M. D., Young, S. B., Aronson, M. P., Morse, A., & Cundiff, G. W. (2008). Paper versus web-based administration of the pelvic floor distress

inventory 20 and pelvic floor impact questionnaire 7. *International Urogynecology Journal, 19*(10), 1331-1335.

Hawthorne, G. (2003). The effect of different methods of collecting data: mail, telephone and filter data collection issues in utility measurement. *Quality of Life Research, 12*(8), 1081-1088.

Herdman, M., Fox-Rushby, J., & Badia, X. (1997). 'Equivalence'and the translation and adaptation of health-related quality of life questionnaires. *Quality of Life Research, 6*(3).

Herdman, M., Fox-Rushby, J., & Badia, X. (1998). A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Quality of Life Research, 7*(4), 323-335.

Heydecke, G., Locker, D., Awad, M. A., Lund, J. P., & Feine, J. S. (2003). Oral and general health-related quality of life with conventional and implant dentures. *Community Dentistry and Oral Epidemiology, 31*(3), 161-168.

Higginson, I. J., & Carr, A. J. (2001). Using quality of life measures in the clinical setting. *Bmj, 322*(7297), 1297-1300.

Inglehart, M. R., & Bagramian, R. (2002). *Oral health-related quality of life*: Quintessence Pub.

Jacobsen, P. B., Davis, K., & Cella, D. (2002). Assessing quality of life in research and clinical practice. *Oncology, 16*(10), 133-139.

Kleinman, L., Leidy, N. K., Crawley, J., Bonomi, A., & Schoenfeld, P. (2001). A comparative trial of paper-and-pencil versus computer administration of the Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire. *Medical care, 39*(2), 181-189.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Lane, D. M. (2007). *Hyperstat online statistics textbook*.

Lee, E.-H. (2009a). Computerized measurement for asthma-specific quality of life: Comparison with a conventional paper-and-pencil questionnaire. *Journal of Korean Academy of Nursing, 39*(6), 781-787.

Lee, E.-H. (2009b). Touch-screen computerized quality-of-life assessment for patients with cancer. *Asian nursing research, 3*(1), 41-48.

Lee, E. H. (2007). [Development and psychometric evaluation of a quality of life scale for Korean patients with cancer (C-QOL)]. *Taehan Kanho Hakhoe chi, 37*(3), 324-333.

Lee, E. H., Lee, Y. W., Lee, K. W., Kim, D. J., Kim, Y. S., & Nam, M. S. (2014). Measurement equivalence of touch-screen computerized and paper-based diabetes-specific quality-of-life questionnaires. *International journal of nursing practice, 20*(4), 382-389.

Lim, C. C. (2008). *Language choices of Malaysian youth: A case study* (Doctoral dissertation, Faculty of Languages and Linguistics University of Malaya Kuala Lumpur).

Locker, D. (1988). Measuring oral health: a conceptual framework. *Community dental health, 5*(1), 3-18.

Locker, D., & Allen, F. (2007). What do measures of 'oral health-related quality of life'measure? *Community Dentistry and Oral Epidemiology, 35*(6), 401-411.

Locker, D., Clarke, M., & Payne, B. (2000). Self-perceived oral health status, psychological well-being, and life satisfaction in an older adult population. *Journal of dental research, 79*(4), 970-975.

Locker, D., Jokovic, A., & Clarke, M. (2004). Assessing the responsiveness of measures of oral health-related quality of life. *Community Dentistry and Oral Epidemiology, 32*(1), 10-18.

Locker, D., Matear, D., Stephens, M., & Jokovic, A. (2002). Oral health-related quality of life of a population of medically compromised elderly people. *Community dental health, 19*(2), 90-97.

Locker, D., Matear, D., Stephens, M., Lawrence, H., & Payne, B. (2001). Comparison of the GOHAI and OHIP-14 as measures of the oral health-related quality of life of the elderly. *Community Dentistry and Oral Epidemiology, 29*(5), 373-381.

Lohr, K. N. (1992). Applications of health status assessment measures in clinical practice: overview of the Third Conference on Advances in Health Status Assessment. *Medical care, 30*(5), MS1-MS14.

Lubetkin, E. I., Jia, H., Franks, P., & Gold, M. R. (2005). Relationship among sociodemographic factors, clinical conditions, and health-related quality of life: examining the EQ-5D in the US general population. *Quality of Life Research, 14*(10), 2187-2196.

Malter, S., Hirsch, C., Reissmann, D. R., Schierz, O., & Bekes, K. (2015). Effects of method of administration on oral health-related quality of life assessment using the Child Perceptions Questionnaire (CPQ-G11–14). *Clinical oral investigations*, 1-7.

McLachlan, S.-A., Allenby, A., Matthews, J., Wirth, A., Kissane, D., Bishop, M., . . . Zalcberg, J. (2001). Randomized trial of coordinated psychosocial interventions based on patient self-assessments versus standard care to improve the psychosocial functioning of patients with cancer. *Journal of Clinical Oncology, 19*(21), 4117-4125.

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review, 101*(2), 343.

Morris, J., Perez, D., & McNoe, B. (1997). The use of quality of life data in clinical practice. *Quality of Life Research, 7*(1), 85-91.

Mustaza, T. A., Lim, T. W., Ghani, A., & Mariam, S. (2016). Mobile applications at the dentist. International Journal on E-Learning and Higher Education, 4, 16-28.

Navabi, N., Nakhaee, N., & Mirzadeh, A. (2010). Validation of a Persian version of the oral health impact profile (OHIP-14). *Iranian journal of public health, 39*(4), 135.

Ng, S. K., & Leung, W. K. (2006). Oral health-related quality of life and periodontal status. *Community Dentistry and Oral Epidemiology, 34*(2), 114-122.

Nickenig, H.-J., Wichmann, M., Andreas, S. K., & Eitner, S. (2008). Oral health–related quality of life in partially edentulous patients: Assessments before and after implant therapy. *Journal of Cranio-Maxillofacial Surgery, 36*(8), 477-480.

Oliveira, A., Ferreira, P. L., Antunes, B., & Pimentel, F. L. (2010). Quality of life in oncology: electronic device to collect data. *Acta Médica Portuguesa, 23*(6), 1017-1024.

Omar, A. H. (1987). *Malay in its sociocultural context*: Dewan Bahasa dan Pustaka, Ministry of Education.

Papagiannopoulou, V., Oulis, C. J., Papaioannou, W., Antonogeorgos, G., & Yfantopoulos, J. (2012). Validation of a Greek version of the oral health impact profile (OHIP-14) for use among adults. *Health and quality of life outcomes, 10*(1), 1.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of experimental psychology, 58*(3), 193.

Pouwer, F., Snoek, F., Van der Ploeg, H., Heine, R., & Brand, A. (1997). A comparison of the standard and the computerized versions of the Well-being Questionnaire (WBQ) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ). *Quality of Life Research, 7*(1), 33-38.

Ravaghi, V., Farrahi-Avval, N., Locker, D., & Underwood, M. (2009). Validation of the Persian short version of the Oral Health Impact Profile (OHIP-14). *Oral health & preventive dentistry, 8*(3), 229-235.

Reissmann, D. R., John, M. T., & Schierz, O. (2011). Influence of administration method on oral health-related quality of life assessment using the Oral Health Impact Profile. *European journal of oral sciences, 119*(1), 73-78.

Ribeiro, C., Moreira, L., Silveira, A., Silva, I., Gestal, J., & Vasconcelos, C. (2010). Development and use of touch-screen computer-assisted self interviewing in Portuguese patients with chronic immune diseases: evaluation of an electronic version of sf-36v2. *Acta reumatologica portuguesa*, (35), 208-214.

Ring, L., Lindblad, Å. K., Bendtsen, P., Viklund, E., Jansson, R., & Glimelius, B. (2006). Feasibility and validity of a computer administered version of SEIQoL-DW. *Quality of Life Research, 15*(7), 1173-1177.

Rosser, R. (1985). A history of the development of health indices. *Measuring the social benefits of medicine. London: Office of Health Economics*.

Rutherford, C., Costa, D., Mercieca-Bebber, R., Rice, H., Gabb, L., & King, M. (2016). Mode of administration does not cause bias in patient-reported outcome results: a meta-analysis. *Quality of Life Research, 25*(3), 559-574.

Ryan, J. M., Corry, J. R., Attewell, R., & Smithson, M. J. (2002). A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper version. *Quality of Life Research, 11*(1), 19-26.

Safina, F., BPharmSc, M., Awang, R., Bakar, S. A., & Hamdan, Z (2006). Evaluation of a Web-Based Smoking Cessation Information System by Malaysian Pharmacists Trained for a Smoking Cessation Program.

Salaffi, F., Gasparini, S., Ciapetti, A., Gutierrez, M., Carotti, M., & Grassi, W. (2013). AB0537 Usability of an innovative and interactive electronic system for collection of patient-reported data in axial spondyloarthritis: comparison with the traditional paper-administered format. Annals of the Rheumatic Diseases, 72(Suppl 3), A953-A953.

Salaffi, F., Gasparini, S., & Grassi, W. (2009). The use of computer touch-screen technology for the collection of patient-reported outcome data in rheumatoid arthritis: comparison with standardized paper questionnaires. *Clinical & Experimental Rheumatology, 27*(3), 459.

Saub, R., Locker, D., & Allison, P. (2005). Derivation and validation of the short version of the Malaysian Oral Health Impact Profile. *Community Dentistry and Oral Epidemiology, 33*(5), 378-383.

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, 24*(3), 166.

Singh, K., & Xie, M. (2008). Bootstrap: a statistical method. *Unpublished manuscript, Rutgers University, USA. Retrieved from http://www. stat. rutgers. edu/home/mxie/RCPapers/bootstrap. pdf.*

Suresh, KP (2011). An overview of randomization techniques: an unbiased assessment of outcome in clinical research. Journal of human reproductive sciences, 4(1), 8

Slade, G. D. (1997). Derivation and validation of a short-form oral health impact profile. *Community Dentistry and Oral Epidemiology, 25*(4), 284-290.

Slusanschi, O., Moraru, R., Garneata, U., Mircescu, G., Cuculescu, M., & Preoteasa, E. (2013). Validation of a Romanian version of the short form of the oral health impact profile (OHIP-14) for use in an urban adult population. *Oral health & preventive dentistry, 11*(3).

StatSoft. (2003). Power Analysis. Retrieved from http://www.statsoft.com/Textbook/Power-Analysis.

Thumboo, J., Fong, K.-Y., Chan, S.-P., Machin, D., Feng, P.-H., Thio, S.-t., & Boey, M.-L. (2002). The equivalence of English and Chinese SF-36 versions in bilingual Singapore Chinese. *Quality of Life Research, 11*(5), 495-503.

Tsakos, G., Bernabé, E., O'Brien, K., Sheiham, A., & de Oliveira, C. (2008). Comparison of the self-administered and interviewer-administered modes of the child-OIDP. *Health Qual Life Outcomes, 6*, 40.

Tubert-Jeannin, S., Riordan, P., Morel-Papernot, A., Porcheray, S., & Saby-Collet, S. (2003). Validation of an oral health quality of life index (GOHAI) in France. *Community Dentistry and Oral Epidemiology, 31*(4), 275-284.

Thompson, S. K. (2012). Simple random sampling. Sampling, Third Edition, 9-37.

Velikova, G., Booth, L., Smith, A. B., Brown, P. M., Lynch, P., Brown, J. M., & Selby, P. J. (2004). Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *Journal of Clinical Oncology, 22*(4), 714-724.

Velikova, G., Brown, J., Smith, A., & Selby, P. (2002). Computer-based quality of life questionnaires may contribute to doctorâ€"patient interactions in oncology. *British Journal of Cancer, 86*(1), 51-59.

Velikova, G., Wright, E., Smith, A., Cull, A., Gould, A., Forman, D., . . . Selby, P. (1999). Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. *Journal of Clinical Oncology, 17*(3), 998-998.

Ward, W. L., Hahn, E. A., Mo, F., Hernandez, L., Tulsky, D. S., & Cella, D. (1999). Reliability and validity of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) quality of life instrument. *Quality of Life Research, 8*(3), 181-195.

Ware, J. E., & Gandek, B. (1998). Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. *Journal of clinical epidemiology, 51*(11), 945-952.

Webb, P. M., Zimet, G. D., Fortenberry, J. D., & Blythe, M. J. (1999). Comparability of a computer-assisted versus written method for collecting health behavior information from adolescent patients. *Journal of Adolescent Health, 24*(6), 383-388.

Wellek, S., & Blettner, M. (2012). On the proper use of the crossover design in clinical trials. *Dtsch Arztebl Int, 109*(15), 276-281.

Wilde Larsson, B. (2006). Touch-screen versus paper-and-pen questionnaires: effects on patients' evaluations of quality of care. *International Journal of Health Care Quality Assurance, 19*(4), 328-338.

Xin, W., & Ling, J. (2006). [Validation of a Chinese version of the oral health impact profile-14]. *Zhonghua kou qiang yi xue za zhi= Zhonghua kouqiang yixue zazhi= Chinese journal of stomatology, 41*(4), 242-245.

Yu, J., Coons, S. J., Draugalis, J. R., Ren, X. S., & Hays, R. D. (2003). Equivalence of Chinese and US–English versions of the SF-36 Health Survey. Quality of Life Research, 12(4), 449-457.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

I have presented a poster presentation on Patients' Preferences in Oral Health Related Quality of Life Assessments at the 28th Annual Scientific Meeting of International Association for Dental Research South East Asian (IADR SEA Division) that was held in Kuching, Sarawak from 11-14 August 2014. Below is the abstract of the study:

Patients' Preferences in Oral Health Related Quality of Life Assessments

Maznur Anowar [1], Colman McGrath [2], Roslan Saub [1,3]

[1] Faculty of Dentistry, University of Malaya, Kuala Lumpur

[2] Faculty of Dentistry, The University of Hong Kong, Hong Kong

[3] Community Oral Health Research Group, Faculty of Dentistry, University of Malaya, Kuala Lumpur.

Objectives: To determine patients' preference and acceptance of computer touch screen-based assessments (CTSA) of oral health related quality of life (OHRQoL) compared with conventional questionnaire based assessments (QBA). In addition, to investigate and compare time required to complete CTSA and QBA.

Methods: An experimental randomized cross-over trial involving 49 subjects. Participants were block randomized into two groups: Group A completed CTSA first followed by QBA; Group B completed QBA first followed by CTSA of their OHRQoL, employing the 14-item Malaysian short-form version of the Oral Health Impact Profile [S-OHIP(M)]. On completion of tasks participations reported their preferred mode of assessment (touch screen, paper, or no preference). Time taken to complete assessments was recorded by a stopwatch. Preference for assessment method was determined, and variations in S-OHIP(M) scores, *Cronbach's alpha* values (internal reliability) and time taken to complete assessments were compared for CTSA and QBA of OHRQoL.

Results: There was no significant difference in S-OHIP(M) scores derived by CTSA versus QBA: 17.4 (sd 8.5) versus 17.6 (sd 8.9), $P>0.05$. In addition, there was no significant difference in reliability (*Cronbach's alpha* values) of CTSA versus QBA ($P>0.05$). Most (67.3%, 33) preferred the CTSA approach. Time taken to complete CTSA compared to QBA was longer (137.8 (sd 41.2) *sec* Vs 115.7 (sd 42.8) *sec*, $P<0.01$.

Conclusion: In this experimental randomized cross-over trial a CTSA of OHRQoL was preferred to a conventional QBA approach. OHRQoL values and reliability of assessments were similar for both CTSA and QBA of OHRQoL. Time taken to complete CTSA was somewhat longer than the QBA of OHRQoL. These findings have implications in implementing OHRQoL in research and practice.

Acknowledgement: This study was supported by the Exploratory Research Grant Scheme (Gudbergsen et al.), Ministry of Higher Education, Malaysia, Grant No: ER020-2012A.