HEALTH UTILITY OF PATIENTS WITH NON-HEALING DIABETIC FOOT ULCERS

HEALTH UTIILITY OF PATIENTS WITH NON-HEALING DIABETIC FOOT ULCERS

By ADAM EDWARD HAYNES, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2018) Hamilton, Ontario (Health Sciences)

TITLE: Health Utility of Patients with Non-Healing Diabetic Foot Ulcers

AUTHOR: Adam Edward Haynes, B.Sc. (Brock University) SUPERVISOR: Professor Daria O'Reilly NUMBER OF PAGES: xi, 158

LAY ABSTRACT

Diabetic foot ulcers (DFUs) impart a large burden on patients and the healthcare system in Canada. The objective of this thesis was to identify health utility estimates for patients with non-healing DFUs. A systematic review was conducted and included nine studies, for which health utility estimates ranged from 0.44 to 0.89. An exploratory regression analysis of data from an Ontario-based study in patients with non-healing DFUs was conducted. No factors associated with health utility were identified but further research is required. A descriptive regression model, adjusting for several baseline factors, provides a health utility of 0.647 for Canadian patients, but should be interpreted with caution. Guidance is needed on the best methodology to conduct studies to estimate the health utilities for use in economic evaluations and for a tool to critically appraise studies to help select the best estimate of health utility for inclusion in economic evaluations.

ABSTRACT

Diabetic foot ulcers (DFUs) impart a large burden on patients and the healthcare system in Canada. Health utility estimates are an integral part of determining the cost-effectiveness of treatments for DFUs. The objective of this thesis was to identify health utility estimates for patients with non-healing DFUs. A systematic review of studies reporting health utility estimates for non-healing DFUs was conducted and included nine studies. The quality of the studies, as it related to the health utility estimates for non-healing DFUs, was difficult to determine due to a lack of reporting of study and patient characteristics. The health utility estimates ranged from 0.44 to 0.89. None of the studies investigated for factors associated with the health utility of patients with non-healing DFUs.

In addition, an exploratory regression analysis of data from a randomized controlled trial (RCT) of hyperbaric oxygen therapy (HBOT) in patients with chronic, non-healing DFUs was conducted. No factors were identified that were associated with health utility; however, the sample size was small and the analysis exploratory. Further research is required to identify such factors. Finally, a descriptive regression model, including several baseline factors, was created which provided a heath utility estimate of 0.647 for Canadian patients with non-healing DFUs; however, the results should be interpreted with caution, especially as some subgroups had very small numbers of patients (e.g., Wagner Grade of 4; patients with 4 or more wounds).

In summary, guidance is lacking on the best methodology to conduct and analyze studies that provide estimates of the health utility of patients with non-healing DFUs, or any other health state, that are to be used to inform economic evaluations. Additionally, a tool is needed to aid analysts in critically appraising studies so that they can select the best estimate of health utility value to include in economic evaluations.

ACKNOWLEDGEMENTS

This work would not have been possible without the support of several individuals. I owe many thanks to my committee members, Dr. Daria O'Reilly, Dr. Feng Xie, and Dr. Lehana Thabane, for your support, feedback on various drafts, and guidance. Special thanks to my supervisor, Dr. Daria O'Reilly. Without your help and guidance, this would not have been possible. I appreciated your patience throughout this process, especially when it came to deciding on a topic.

Julie Makarski; your willingness to help screen citations and extract data, all for the cost of a cup of tea, will never be forgotten and will always be appreciated. Thank you for your friendship and your willingness to help.

My colleagues at CADTH. Your support, helpful conversations, and friendships helped me more than you may realize.

My amazing children—Nathaniel, Josiah, Ethan, and Elizabeth—you have been so supportive and patient while I took this journey, one that we really took together. I hope that, one day, each of you will take this journey yourself, and I will be there for you as you were for me.

Lastly, and most importantly, to my amazing and beautiful wife, Suzanne: your support, encouragement, long talks and love have carried me through this journey. I cannot thank you enough. You are truly the Proverbs 31 wife. To my only, forever.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

BMI	body mass index
DFU	diabetic foot ulcer
Emtree	Embase Subject Heading
EQ-5D	EuroQol-5 Dimension
HBOT	hyperbaric oxygen therapy
HRQoL	health-related quality of life
HUI	health utility index
ICER	incremental cost-effectiveness ratio
ICUR	incremental cost-utility ratio
MeSH	medical subject heading (Medline)
NHS	National Health Service (UK)
NICE	National Institute for Health and Care Excellence (UK)
QALY	quality-adjusted life year
RCT	randomized controlled trial
SD	standard deviation
SIGN	Scottish Intercollegiate Guidelines Network
TTO	time trade-off

DECLARATION OF ACADEMIC ACHIEVEMENT

A.E. Haynes and Daria O'Reilly designed the study. A.E. Haynes and Julie Makarski identified studies for inclusion and extracted data. A.E. Haynes collected and analyzed data for the systematic review portion of the study, analyzed previously collected data from the HBOT clinical trial, and drafted the manuscript. Daria O'Reilly assisted in the analysis and interpretation of the data (by providing advice), and provided input into drafting the manuscript.

CHAPTER 1: INTRODUCTION

In 2008-2009, the prevalence of diabetes in Canada was 2.4 million cases, representing 6.8% of the entire population of Canada (1). During the same time period, the incidence of diabetes was 6.3 new cases per 1,000 (1). In 2008, the prevalence of diabetes in Ontario was 857,810 cases or 8.4% of the population of the province, but in 2012, the prevalence had increased to 1.1 million cases or 10.2% of the population (2).

People with diabetes are susceptible to many complications associated with their disease, including peripheral artery disease (leading to reduced blood flow to the lower extremities) and peripheral neuropathy (damage to the nerves of the peripheral nervous system) (3). Both of these complications can result in poor wound healing in these individuals, which can lead to the development of chronic non-healing diabetic ulcers of the lower limb (often referred to as a diabetic foot ulcer or DFU). It is estimated that 15-25% of all patients with diabetes will develop a DFU within their lifetime (4-7). Left untreated, a DFU can result in infections, which may lead to amputation of the affected limb.

In Ontario, treatments for patients with non-healing DFU's are paid through a publicly funded health care system. As in other publicly funded health care systems, resources are limited and the choice of which treatment options to fund must be based on which provide the best value for money (i.e., which are the

most cost-effective). Decision makers often rely on economic evaluations that determine the incremental cost-effectiveness of a treatment option(s) compared with the currently funded treatment option(s) in order to make funding decisions. A specific type of cost-effectiveness analysis is the cost-utility analysis. A cost-utility analysis considers the impact that an intervention may have on survival and the quality of life of patients compared with other interventions. Health-related quality of life is incorporated into these analyses by using health utilities, which is anchored at 0, representing death, and 1, representing full or perfect health. A cost-utility analysis takes into consideration the time spent in a given health state multiplied by the health utility associated with that health state in order to determine the amount of quality-adjusted life years (QALYs) that could be expected for a given treatment. Therefore, a QALY of 0 would be equivalent to being dead, 1 would represent a year of perfect health, and a QALY of 0.5 could represent a full year at half quality of life or half a year at perfect quality of life. A cost-utility analysis, therefore, allows comparisons between treatments that may impact disparate outcomes, by providing a common denominator upon which to make comparisons. However, in order to have reliable estimates of the incremental cost-utility of a treatment compared with another treatment, it is critical to have reliable estimates for the health utilities of the population included in the evaluation. Without reliable health utilities, and hence, accurate incremental cost-utility estimates, decision makers are unable to make appropriate decisions regarding which treatments to fund.

The identification of relevant health states is paramount in any economic evaluation. When estimating the cost-effectiveness of interventions aimed at treating chronic non-healing DFUs, the starting state of an economic model comparing treatments for DFUs would be a non-healing DFU. The time spent in this state is dependent on the effectiveness of the treatments being compared. Having a reliable estimate of the health utility of patients with non-healing DFUs is paramount in order to obtain an accurate estimate of the cost-effectiveness of any treatment.

Patients with diabetes may suffer from multiple co-morbidities, which further complicates the determination of an accurate estimate of health utility for these patients due to the potential for confounding between these co-morbidities, chronic non-healing DFUs, and health utility scores. Understanding the factors that are associated with the health utility of patients with non-healing DFUs would also help to ensure that the full effects of treatments on the health utility of patients with non-healing DFU and other co-morbidities are accounted for in an economic evaluation.

The primary objective of this thesis was to identify published health utilities for non-healing DFUs from a systematic review of the literature as well as to estimate health utilities from the analysis of primary data from a clinical trial

conducted in Ontario. A secondary objective of this latter analysis was to investigate what factors might affect the health utility value.

CHAPTER 2: SYSTEMATIC REVIEW

2.1 INTRODUCTION

Accurate estimates of the health utility of patients with chronic non-healing DFUs are needed in order to understand the quality of life in these individuals and to also inform economic evaluations of interventions for non-healing diabetic ulcers of the lower limb. In order to identify an estimate of the health utility for patients with non-healing DFUs, a systematic review of studies investigating the health utility of patients with non-healing DFUs was conducted. A systematic review minimizes bias by seeking and identifying all relevant studies that address a specific research question. This is accomplished by using explicit and systematic methods to identify studies that answer the research question and by establishing clear criteria for the inclusion and exclusion of studies in the review. Furthermore, a systematic review also applies systematic methods to the collection of data and critical appraisal of included studies, thus further minimizing potential biases in the reporting of data as well as reducing the likelihood of errors in the collection of data.

The objective the systematic review was to identify studies that that included patients with chronic non-healing DFUs, and collected and reported health utility data for these patients. A secondary objective was to identify studies that investigated factors associated with the health utility of patients with chronic non-healing DFUs.

Selection bias (of studies) is minimized in a systematic review by using clearly defined inclusion and exclusion criteria and by utilizing more than one reviewer when screening citations. As citations can have unclear descriptions of studies, it is possible for a single reviewer to misunderstand the description and exclude a potentially relevant study. Furthermore, a single reviewer may mistakenly exclude a relevant citation due to fatigue. Steps to mitigate bias in the selection of studies are important and the exact methods chosen often depend on the availability of funding and resources. The gold standard approach is duplicate screening, where all citations are reviewed by at least two independent reviewers (8).

The extraction of data from included studies is another important step and needs to be done in an unbiased manner. The gold standard approach is to use data collection forms and duplicate data extraction i.e., two reviewers extract data from each study publication (8). However, the cost of duplicate data extraction can be prohibitive for many researchers. In those cases, alternative measures may be taken to ensure the integrity of the collected data, both in accuracy and in the interpretation of what was reported.

Health utility data are measured in a number of ways, however, there are two general approaches. The first is the direct elicitation of utilities. With this

approach, the researcher will elicit preferences for a given health state(s) using established techniques from a sample of a given population. If a societal perspective is desired, the researcher would sample from the general population. Commonly used preference elicitation techniques include time trade off and standard gamble approach (9). In both cases, the researcher provides a description of the health state of interest to the participant. In the time trade off technique, the participant is asked to state how much time they would be willing to trade off in a better health state rather than living longer in the "lesser" health state described by the researcher (9). The time in the better state is varied until the point at which the participant is indifferent between the two choices, which represents the relative preference for that state. The health utility is then estimated by dividing the time duration where the indifference is reached by the duration in the lesser state. The standard gamble asks the patient to choose between living in a defined health state or to take a gamble between a given probability (p) of living at optimal health or a probability (1-p) of living the worst possible outcome (usually, this is immediate death). The probability, p, is changed until the participant is indifferent between choosing a probability of having optimal health and living in the defined health state, at which point, p, represents the utility of being in the defined health state (9).

The second approach to elicitation of health utilities is an indirect approach, where patients who are experiencing the disease state of interest are asked to

define their health state using a standardized questionnaire; for example, the EuroQoL 5-Dimension (EQ-5D) asks patients whether they have experienced problems with usual activities, self-care, mobility, pain or discomfort, and anxiety or depression (9). The resulting combination of responses defines the patient's health state, to which a tariff or preference weight is applied in order to produce a health utility score. The tariffs are usually based on a sample of a general population and represent a societal preference.

The objective of this systematic review was to identify estimates of health utility that could apply to patients with chronic non-healing DFUs. A secondary objective is to determine what factors (e.g., patient or disease characteristics) are associated with health utility scores for patients with chronic non-healing DFUs.

2.2 METHODS

Research Questions

Primary Question

What is the health utility of patients with chronic non-healing diabetic ulcers of the lower limb?

Secondary Question

What factors are associated with the health utility of patients with non-healing DFUs?

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (10).

Literature Search Methods

A systematic literature search of Ovid Medline (1946 to July 21, 2017) and Embase (1974 to July 21, 2017), the Cochrane Database of Systematic Reviews (July 2017), the Cochrane Central Register of Controlled Trials (June 2017), the Database of Abstracts of Reviews of Effects (April 2015), the Health Technology Assessment Database (October 2016), and the UK National Health Service (NHS) Economic Evaluation Database (April 2015) were searched for studies that measured the health utility of patients with chronic non-healing diabetic ulcers of the lower limb.

Search terms included Medline Medical Subject Heading (MeSH) or Embase Subject heading (Emtree) terms as well as keyword searches based on variations of the following terms: diabetes, ulcers of the lower limb, and health utilities. Unfortunately, there are no specific Medical Subject Heading (MeSH) terms in Medline (or analogous controlled vocabulary in other databases) specifically for chronic non-healing diabetic ulcers of the lower limb. As a result of this, the search strategy for chronic non-healing DFUs needed to include text word and keyword searches for the terms chronic and non-healing (and their synonyms), in order to increase the specificity of the search strategy. As the intent of a

systematic review is to identify all relevant studies, the search strategy has to be sufficiently broad in order to capture all of relevant studies. However, as search terms are added in order to ensure that all relevant studies are identified (i.e., sensitivity is increased), the number of irrelevant studies also increases (i.e., specificity is decreased). As resources are finite, there is a need to make the search more efficient by reducing the number of irrelevant citations captured in the literature search (i.e., the specificity needs to increase). The challenge in developing a literature search strategy is that, as one endeavours to identify all relevant studies, the amount of resources required to screen all of the identified citations can become very high. Furthermore, and in a general sense, as sensitivity increases, the specificity decreases, which means that attempts to ensure that all relevant studies are identified can also have the simultaneous effect of increasing the number of irrelevant studies in the search results. In order to ensure that the literature search strategy was complete, did not miss relevant studies, did not contain errors (e.g., missing index terms and subject headings, spelling errors, variations in spelling, appropriate use of truncation, etc), and to improve the efficiency of the search, a librarian with expertise in conducting literature searches in several databases peer-reviewed the search strategy. Appendix 1 includes the full and final literature search strategy.

Selection bias (of studies) is minimized in a systematic review by using clearly defined inclusion and exclusion criteria and by utilizing more than one reviewer

when screening citations. As citations can have unclear descriptions of studies, it is possible for a single reviewer to misunderstand the description and exclude a potentially relevant study. Furthermore, a single reviewer may mistakenly exclude a relevant citation due to fatigue. In order to reduce the risk of bias, all citations were screened by two reviewers, according to the following eligibility criteria:

Inclusion criteria:

- Studies must have included diabetic patients with chronic non-healing ulcers of the lower limb and they must have measured and reported data on the health utility of these patients
- Studies must have had a prospective design (i.e., observational cohort, randomized trial, non-randomized trial), have a cross-sectional design, or be a systematic review.

Exclusion Criteria:

- Non-systematic reviews, economic evaluations, editorials, or letters to the editor
- Case reports
- Case series (where the results of patients are reported individually, rather than combined as a cohort)
- Publications in languages other than English due to lack of resources for translation

The level of agreement between reviewers was calculated using Cohen's kappa (11). Disagreements between reviewers were resolved through discussion or, where the disagreement could not be resolved, through a third reviewer.

Reference lists of studies that were excluded at the full text screening stage were scanned for potentially relevant studies, as were the reference lists of included studies. Furthermore, while economic evaluations were excluded from this study, the reference lists of those studies were scanned to identify additional relevant studies. However, no additional relevant studies were identified in reference lists that were not already captured in the literature search.

The following data from included studies was abstracted: study population (including baseline characteristics such as ethnicity, age, gender, geographic region, severity of foot infection [Wagner classification], type of diabetes), study design, statistical methods (including whether utility estimates were adjusted for confounders), the tool/questionnaire used to measure health utility (e.g., EQ-5D, HUI, etc), the scoring algorithm used (e.g., UK, US, etc), and health utility estimates for individuals with diabetes who had a chronic non-healing ulcer of the lower limb.

It was planned to pool the health utility data and meta-analyze the results from the studies where possible and if this was not possible, a narrative synthesis would be conducted.

Assessment of Study Quality

Given that health utilities can be obtained in different ways, using a variety of methodologies, a critical appraisal of the methods used to estimate health utility values is important in order to determine the validity of the results and the applicability to a Canadian population. There exist many tools for the critical appraisal of the various study designs (e.g., randomized controlled trials, observational cohorts, cross-sectional) that could be identified in this review (12-14). The selection and planned use of one critical appraisal tool in a systematic review is necessary as it reduces bias in the assessment of the validity and the overall quality of the included studies by utilizing a common, systematic approach to the appraisal of each included study. Notably, there is a large number of checklists, tools and scales to aid in the assessment of the quality of RCTs, observational studies and cross-sectional studies (15,16), making a priori choice of one appraisal tool difficult. However, the method of collection of health utility data for patients with chronic, non-healing DFU within each study is what is most relevant to this systematic review. As the objective of this systematic review is to identify and summarize the available health utility data for patients with chronic non-healing DFU's, an assessment of the quality of any comparative elements of

studies was not considered germane to this review. For instance, the results of comparisons of various treatments are not relevant to this review, therefore an assessment of the potential risk of bias in those estimates was not needed. Furthermore, the results of comparisons of the health utilities of patients with chronic non-healing DFU's with that of patients with healed DFU's or resulting amputations were also considered irrelevant. As the health utility data would have been collected at a single point in time, through a questionnaire, survey form (e.g., an EQ-5D questionnaire) or interview, the health utility results of the included studies can be assessed as if they were obtained in a cross-sectional study. Of note, the authors are not aware of any quality assessment tool designed specifically to assess the quality of studies that measure health utilities. As a result, it was decided to use the Appraisal tool for Cross-Sectional Studies (AXIS) (17) to assess the quality of all of the included studies, not just those with a cross-sectional design. For studies with a randomized controlled trial design, the Cochrane Risk of Bias tool (13) was also used.

Duplicate Data Extraction

The extraction of data from included studies is a very important step and needs to be done in an unbiased manner. The gold standard approach is to use data collection forms and duplicate data extraction i.e., two reviewers independently extract data from each study publication (8). Due to limitations in the availability of funding, a second reviewer was not used to extract data for, and assess the

quality of, all included studies. However, in order to limit the potential for bias in the extraction of data and the assessment of study quality, a second reviewer extracted data from and critically appraised three randomly selected included studies. If the extracted data and critical appraisal were in agreement between the two reviewers (i.e., the two reviewers extracted the same values for 80% of the study data and reached the same conclusions for the AXIS items), the remaining studies would not undergo duplicate data extraction and critical appraisal.

2.3 RESULTS

Literature Search Results

A total of 2,393 citations were identified through the literature search (see Figure 1 for a PRISMA flow diagram of identified citations). Of those, 412 were duplicate citations, leaving 1,981 unique citations to be screened by two reviewers.

After the title and abstract screening phase both reviewers agreed that 9 citations met the criteria for the review, 70 citations were considered “maybe” and would require full text review to make the final determination, and 1,776 citations were considered to not meet the criteria for the review (i.e., “no”). However, Reviewer One identified 66 citations as “maybe” while Reviewer Two considered those citations to not meet the criteria. On the other hand, Reviewer Two identified 60

citations as “maybe” while Reviewer One considered those citations to not meet the criteria. The agreement between the two reviewers after the title and abstract screening was moderate (Cohen’s kappa=0.52).

The two reviewers discussed the disagreements and agreed that the full texts for the 60 and 66 citations where they were in disagreement would be retrieved, in addition to the nine relevant citations and 70 possibly relevant papers. This resulted in a total of 205 papers for full text review. After full text review, 195 papers were excluded (Figure 1). Therefore, a total of 10 publications of nine unique studies met the eligibility criteria and were included in this systematic review (Table 1). (18-27).

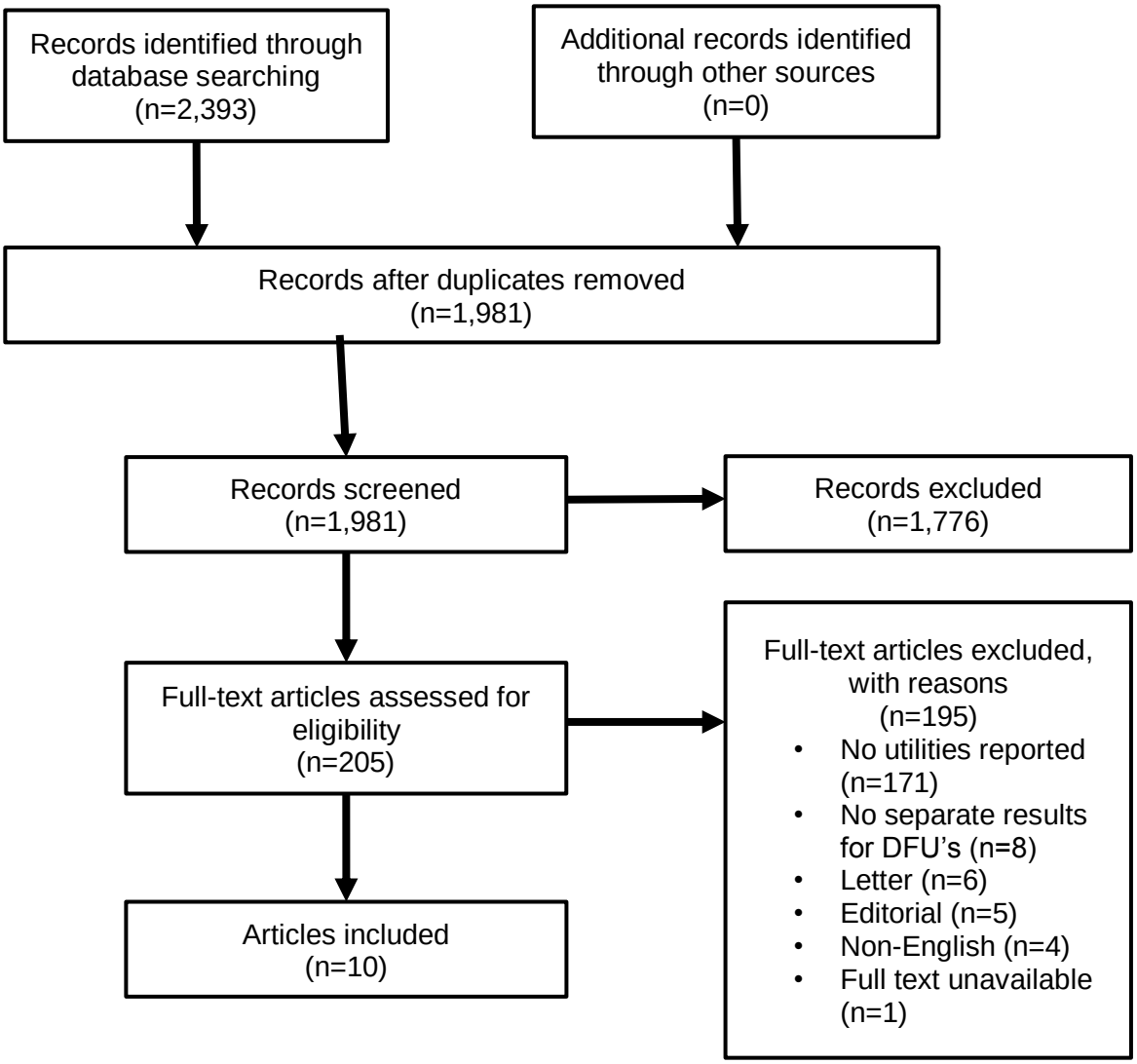

Figure 1. PRISMA flow diagram of identified citations.

Table 1. List of included studies and associated publications.

Study	**Primary publication (Author, year)**	**Type of publication**	**Secondary publications (Author, year) and details**
Tennvall et al	Tennvall GR, 2000 (18)	Full	
Coffey et al	Coffey JT, 2002 (19)	Full	
Redekop et al	Redekop WK, 2004 (20)	Full	
Morgan et al	Morgan CL, 2006 (21)	Full	
Javanbakht et al	Javanbakht M, 2012 (22)	Full	
Siersma et al	Siersma V, 2013 (23)	Full	Siersma V, 2017 (24) – includes same participants plus data for additional participants Prompers L, 2007 (28) – includes information about the design and methods of the Eurodiale study.
Sobol et al	Sobol E, 2013 (25)	Abstract	
Li et al	Li G, 2017 (26)	Full	Fedorko L, 2016 (29) – primary trial publication. Includes additional information about the trial methods.
Sothornwit et al	Sothornwit J, 2017 (27)	Abstract	

Duplicate Data Extraction

Three papers were randomly selected for duplicate data extraction were Tennvall et al (18), Coffey et al (19), and Redekop et al (20).

For the Tennvall et al (18), the two reviewers had perfect agreement on the data extracted for each variable and the information. For the Redekop et al study (20), the two reviewers extracted different proportions of male and female participants, but the remainder of the extracted data were in agreement. After discussion, it was noted that one reviewer extracted an incorrect value. For the Coffey et al study (19), the two reviewers extracted the same values for all variables, with the exception of the adjusted health utility estimates. Of note, after discussion, it was determined that one reviewer extracted two incorrect values, which were then corrected.

Given that the extracted values and information were almost identical between the two reviewers for all three studies, further duplicate extraction was deemed unnecessary.

Study Design and Characteristics

The designs and characteristics of the nine included studies varied, as they each had different objectives, although some differed only slightly (Table 2). With respect to the ascertainment of HRQoL in patients with chronic, non-healing

diabetic foot ulcers, all of the studies utilized a cross-sectional approach. That is, the participants' HRQoL was determined at a single time point during the study. The objectives of the studies were generally broader than those of this study. Specifically, the current study aimed to estimate the HRQoL of patients with chronic non-healing DFUs; however, the included studies aimed 1) to estimate the HRQoL of patients with diabetes (and included patients with foot ulcers) (19,21,22); 2) to estimate the HRQoL of patients with DFU and/or amputations (18,20,27); 3) to investigate the factors related to HRQoL in patients with DFU (23,24); 4) to investigate the impact of severity of DFU on HRQoL (25); and, 5) to investigate the effect of hyperbaric oxygen therapy (HBOT) on HRQoL in patients with DFU (26) .

Of note, all of the studies, with the exception of the RCT reported by Li et al (26), were reported as having cross-sectional designs. Li et al (26) used data from a randomized controlled trial to compare treatment with HBOT versus placebo in patients with DFU. While patients in the trial were asked to complete an EQ-5D-3L questionnaire at baseline and at several follow-up points, only the baseline health utility estimates of both trial arms were used for the current study. Furthermore, the results from each study arm were pooled in the current study in order to decrease the uncertainty in the estimate (i.e., to provide a better estimate of the health utility than either arm could provide alone). It is important to note that the HRQoL data used to inform the present study was collected at a single

point in time (i.e., the current study did not consider data collected longitudinally), therefore, those data were considered to be collected in a manner similar to a cross-sectional approach. Of the remaining eight studies, six recruited consecutive patients, identified from hospital records, and who met specific criteria (18,19,21,23-25,27). Those identified were contacted and asked to complete the EQ-5D-3L in four studies (18,21,23-25), the EQ-5D-5L in one study (27), or the Quality of Well-Being Index in one study (19) in order to estimate health utility (Table 2). Javanbakht et al (22) used a multi-stage cluster sampling method to identify patients with Type 2 diabetes mellitus and used the EQ-5D-3L to estimate health utilities (Table 2).

Table 2. Included studies: design details, methods and select quality characteristics.

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Tennvall et al (18)	To investigate HRQoL in patients with diabetes and foot problems and to analyze whether EQ-5D is an acceptable instrument for differentiating patients with current foot ulcers (i.e., non-healed DFU) from patients with primary healed ulcers or patients who have undergone minor or major amputations	Cross-sectional survey	Sweden. Postal survey (with EQ-5D) sent to 457 consecutive patients with diabetes (Type 1 and 2) who had received treatment for DFUs at a multidisciplinary foot care clinic at Lund University Hospital from January 1995 to December 1998.	HRQoL as measured by EQ-5D-3L	EQ-5D-3L (Swedish version) and UK tariffs (Dolan, 1995(30))	Conducted multiple linear regression to determine independent factors related to EQ-5D derived health utilities. Stepwise exclusion of statistically insignificant factors. Variables initially included in the regression model were age, sex, duration of diabetes, other diabetes-related complications, type of ulcer, foot ulcer and amputation status, and living with a healthy partner).

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Coffey et al (19)	To describe the health utilities associated with diabetes and its treatments, complications and comorbidities	Cross-sectional survey	Michigan, US. 2,048 patients who attended endocrinology, diabetes, and ophthalmology clinics at the University of Michigan Health System between June 29, 1998 and March 15, 2001, were asked to complete the survey	HRQoL and health utilities	Quality of Well Being Index (QWB); preference weights from a sample drawn from San Diego (US) (31)	Multiple linear regression using demographics (i.e., age, sex, race, age at onset of diabetes, duration of diabetes, BMI) and disease variables (i.e., treatments, retinopathy, nephropathy, neuropathy [including DFU[A]], stroke, cardiovascular disease, hypertension, hypercholesterolemia).

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Redekop et al (20)	To derive health utility values for health states involving diabetic foot ulcers and amputations	Cross-sectional TTO	Rotterdam, Netherlands. Quota-stratified sampling (age and gender) to ensure sample was representative of general Dutch population. In total, 107 subjects were included/participated in the study (of the 107, 11 results were considered invalid, thus, analysis was conducted using results for n=96 participants).	Health utility as measured by TTO	TTO – Participants were interviewed in groups and provided a vignette describing disease state, with participants asked to indicate how many years of life they would be willing to give up to avoid the health state (life expectancy set at 85 years).	Linear regression using to check for associations between utility values assigned to health states and patients age, gender and previous amputation.

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Morgan et al (21)	To derive and compare health utilities for patients with diabetes with single and multiple morbidities	Cross-sectional survey	Cardiff and the Vale of Glamorgan, UK. Postal questionnaire sent to patients six weeks after discharge from hospital. Patients discharged between January 2002 and July 2005 and age >18 years were included. 50, 258 responses received out of 150,113 questionnaires sent.	Health utility using EQ-5D-3L	EQ-5D-3L, preference weights NR (for the purposes of this paper, assumed that UK tariffs (32) were used)	Multiple linear regression, controlling for gender, age, and BMI, and assessing impact of complications as binary variables (presence/history of complication or not). Complete list of explored complications NR; however, diabetic foot ulcer was one complication and results were reported for that subgroup.

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Javanbakht et al (22)	To measure HRQoL in patients with Type 2 diabetes mellitus and determine which socio-demographic and diabetes-related clinical characteristics are associated with better HRQoL	Cross-sectional	Iran. 3,918 patients with type 2 diabetes mellitus identified using multi-stage cluster sampling method. Eligible participants met WHO criteria (fasting plasma glucose ≥7.0 mmol/L or with a glucose tolerance test, two hours after the oral dose a plasma glucose ≥11.1 mmol/L).	Health utility using EQ-5D-3L	EQ-5D-3L (Farsi version) and UK VAS tariff (33)	Tobit regression model used to identify factors that affected EQ-5D score. Factors not reported.

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Siersma et al (23,24)	To investigate factors related to HRQoL in patients with DFU	Cross-sectional (Eurodiale was a prospective observational study; however, health utility for patients with chronic, non-healing DFUs were collected in a cross-sectional sub-study from baseline data for all patients in the main Eurodiale study)	Europe (Belgium, Czech Republic, Denmark, Germany, Italy, Slovenia, Spain, Sweden, the Netherlands, UK). Between September 1, 2003 and October 1, 2004, 1,232 patients with a new DFU at one of 14 centres in 10 countries were identified and asked to participate.	Health utility using EQ-5D-3L	EQ-5D-3L (version NR), preference weights NR	Multiple linear regression using patient characteristics (sex, age, centre, current employment, partner involved in care, current smoker, chronic alcohol use, BMI), disease characteristics (duration of diabetes, insulin treatment, HbA_{1c}, serum creatinine and C-reactive protein concentrations), foot- and ulcer-related characteristics (previous lower extremity amputation,

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
						immobilizing offloading, diabetic polyneuropathy, peripheral arterial disease, limb-threatening ischemia, infection, osteomyelitis, and depth, size, duration, and location of ulcer), and co-morbidities (heart failure, neurological disorder, inability to stand or walk without help, visual impairment, and end-stage renal disease)

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Sobol et al (25)	To estimate the effect of severity of DFU on HRQoL	Cross-sectional	NR (assumed to be Poland, based on authors' centres) Between April 2012 and May 2013, 185 patients with DFU who were treated in an ambulatory care clinic were approached, with 179 completing the EQ-5D-3L questionnaire and with information available on severity of ulceration	Health utility using EQ-5D-3L	EQ-5D-3L (version NR), and Polish tariff (34)	NR

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Li et al (26)	To investigate the effect of hyperbaric oxygen therapy on HRQoL in patients with DFU	RCT (note: health utility for patients with chronic, non-healing DFUs available from baseline data for all patients in trial, therefore, the data on non-healing DFUs can be considered cross-sectional)	Ontario, Canada. Patients with Type 1 or 2 diabetes mellitus, Wagner grade 2-4 lower limb lesions that have not healed for at least 4 weeks were enrolled in the trial.	Primary: Health utility using EQ-5D-3L Secondary: SF-36; DFS-SF	EQ-5D-3L (version NR), and Canadian tariff (35)	No adjustment for baseline EQ-5D health utilities

Study	Study objective	Study type (e.g., cross-sectional)	Sampling of study subjects and study location	Study outcomes (Primary, secondary, etc)	Health Utility Data	
					Health State Measurement Tool and Scoring Algorithm (Tariff)	Adjustment for Covariates (including description)
Sothornwit et al (27)	To investigate HRQoL in patients with DFU	Cross-sectional	NR (assumed to be Thailand, based on authors' centres). 254 patients with diabetes; 98 with DFU	Health utility using EQ-5D-5L	EQ-5D-5L (Thai version). Tariffs used were not reported; however, authors stated that utility values were calculated using TTO methods.	NR
Notes: DFS-SF=Diabetic Foot Ulcers Scale-Short Form; NA=not applicable; SF-36=Short From 36; TTO=time trade-off. [A]Patients were asked to report whether they had DFU; however, no information was provided on whether they had healed or non-healed ulcers.						

Redekop et al (20) identified potential participants from the general population using quota-stratified sampling based on age and gender to ensure representativeness with the general Dutch population (Table 2). The health utilities associated with health states involving active DFU and amputations were estimated using a time trade-off approach through group interviews. The authors did not define a minimum amount of time that a patient had to have an active DFU. Health states, defined by ulcer and amputation status, were described to participants, who were then asked to value those states relative to a state of perfect health. The trade off in time was measured relative to a life expectancy of 85 years for all participants.

One study was conducted in Sweden (18), one in the Netherlands (20), one in the UK (21), one in European countries (Table 2) (23,24), one in the US (Coffey), one in Iran (22), and one in Canada (26). Sothornwit et al (27) did not report where the study was conducted; however, it is reasonable to infer that, based on the location of each author's centre, the study was conducted in Thailand. Finally, Sobol et al (25) did not report where the study was conducted, nor could the location be reasonably assumed based on the available information.

In the Li et al study (26,29), a chronic non-healing DFU was defined as patients with an active non-healing DFU for at least 4 weeks duration. Siersma et al included patients with a new DFU, and then reported results from a subset of the

study population who were unhealed after 12 months of follow-up (23,24). The remaining studies included, and reported data on, patients with active non-healing DFUs; however, none clearly reported the duration of the non-healing DFU and none provided a clear definition of chronic non-healing DFU (18-22,25,27).

Of the nine included studies, two were published in abstract form only (25,27) and seven were fully published (Table 1). Siersma et al 2013 (23) and Siersma et al 2017 (24) both reported data for the same sample of participants. Of note, the 2013 publication included some baseline and study design information not reported in the 2017 publication and was used to supplement the 2017 publication.

Baseline Characteristics of Participants with Non-Healing DFU

Siersma et al included 1,232 patients with a new DFU in their study (23,24), with 131 of those patients remaining unhealed after one year (Table 3). The remaining patients were healed over that time and the health utility data reported for them reflected the healed status of their disease. The mean age of the 131 patients with chronic non-healing DFUs was 65.1 years, and 71.8% were male. Approximately 11% of the patients had diabetes for less than five years, 16% had diabetes for 5-10 years, and 73% had diabetes for over 10 years. No data were reported on the severity of DFU. Li et al (26) analyzed data on 103 patients with

chronic DFUs, that were unhealed for at least four weeks, in order to compare two treatment approaches, HBOT (n=49) and a sham procedure (n=54) (Table 3). The mean age was 61 and 62 years, respectively, with a higher proportion of males in both treatment arms (63% and 70%). The mean number of years with diabetes was 19.1 years in the HBOT arm and 12.4 years in the sham arm. Approximately 45% of patients had Wagner Grade 2 DFU, 50% had Wagner Grade 3 DFU, and 6% had Grade 4 DFU, at baseline.

The remaining studies, while they reported data on patients with non-healing DFUs, did not explicitly define the length of time that patients had lived with their DFU. Tennvall et al (18) included 56 patients with "current foot ulcers," Coffey et al (19) included 149 patients with "neuropathic sores," Morgan et al (21) included 661 patients with "DFU," Javanbakht et al (22) included 372 patients with "lower extremity lesions," Sobol et al (25) included 179 patients with DFU "with active foot ulceration," and Sothornwit et al (27) included 98 patients with DFU (Table 3). However, none of those studies provided further details on whether or what treatments those patients may have received for their DFU. The reporting of baseline data for these patients were limited in all of those studies (Table 3). Redekop et al (20) did not limit the inclusion of patients to only those with diabetic foot ulcers; however, the objective of the study was to estimate the health utilities for health states involving DFUs and amputation.

Table 3. Included studies: baseline characteristics of participants with non-healing DFU.

Study	Number of patients with chronic non-healing DFUs	Ethnicity	Mean Age (years)	Gender (% males)	Duration of Diabetes (mean, years)	Disease Severity (%)
Indirectly Elicited Health Utilities						
Tennvall et al (18)	56	NR	NR	NR	NR	NR
Coffey et al (19)	149 [A]	NR	NR	NR	NR	NR
Morgan et al (21)	661	NR	NR	NR	NR	NR
Javanbakht et al (22)	372	NR	NR	NR	NR	NR
Siersma et al (23,24)	131	NR	65.1 ±13.7	71.8	<5 years: 11.1% 5-10 years: 15.9% >10 years: 73.0%	NR
Sobol et al (25)	179	NR	61.9	NR	18.0	Depth/tissue loss:[B] Grade 1: 41.3% Grade 2: 36.3% Grade 3: 22.3% Infection:[B] Grade 1: 46.9% Grade 2: 30.7% Grade 3: 20.1%

Study	**Number of patients with chronic non-healing DFUs**	**Ethnicity**	**Mean Age (years)**	**Gender (% males)**	**Duration of Diabetes (mean, years)**	**Disease Severity (%)**				
Indirectly Elicited Health Utilities										
Li et al (26)						Wagner Grade (%)				
						1	2	3	4	5
	Total N=103	NR	61.5	67.0	15.6	0	44.6	49.5	5.8	0
	HBOT n=49	NR	61	63.3	19.1	0	46.9	44.9	8.2	0
	Sham n=54	NR	62	70.4	12.4	0	42.6	53.7	3.7	0
Sothornwit et al (27)	98	NR	NR	NR	NR	NR				
Directly Elicited Health Utilities										
Study	**Number of Participants**	**Ethnicity**	**Mean Age (years)**	**Gender**	**Duration of Diabetes (mean, year)**	**Disease Severity**				
Redekop et al (20)	107	NR	45[C]	NR	NA	NA				

Notes: NA=not applicable; NR=not reported.

[A]It was assumed that all 149 patients had a non-healing DFU as no information was available on whether patients' DFUs were healed or unhealed.

[B]The grading system used to determine the grade of depth/tissue loss and the grade of infection was not reported.

[C]mean age for all 107 patients included in the study.

Critical Appraisal

The AXIS tool was used as a starting point to assess the quality of the included studies (17), along with the Cochrane Risk of Bias tool (13) and the Newcastle-Ottawa scale (36), where appropriate. The specific critical appraisal results for each study can be found in Appendix 2. Sobol et al (25) and Sothornwit et al (27) were reported in abstract form only, and many details regarding the studies were not reported. Therefore, a proper critical appraisal of those two studies was not possible.

All of the included studies utilized a cross-sectional approach to collect data to estimate the health utility of patients with chronic non-healing DFUs, which is appropriate for determining the health utility of these patients.

Studies That Indirectly Elicited Health Utilities

Eight out of nine of the studies indirectly estimated the health utility of patients with DFUs by using an established and validated questionnaire (e.g., EQ-5D, QWB-SA) and then mapping those responses onto a health utility scale using tariffs (18,19,21-27). Each of those eight studies administered questionnaires to patients with DFUs. Only two of the studies (23,24,26) reported data specifically for patients with chronic non-healing DFUs (i.e., where the duration of the DFU was defined; at least four weeks in Li et al (26) and one year in all patients in

Siersma et al (23,24)). The remaining studies included patients with non-healing DFUs, but did not report the duration of those DFUs (18,19,21,22,25,27). As those studies all included active ulcers, it was assumed, for the purposes of this review, that these patients did not have new DFUs, and therefore had an non-healing DFU. Uncertainty remains in the duration of the DFUs in patients included in each of those studies. The target population was not explicitly stated for six of the studies (18,19,23-27), but for Sobol et al (25) and Li et al (26) the target population could reasonably be inferred from the objectives of each study to be all patients with an active DFU.

Many of the eight studies included only small numbers of participants to estimate the health utility of patients with chronic non-healing DFUs (18,19,23-27). Javanbakht et al (22) included 372 participants with chronic non-healing DFUs, while Morgan et al included 661.

Tennvall et al (18) sent EQ-5D-3L questionnaires, at the end of 1998, to 457 consecutive patients with DFU (Type I or Type II diabetes) who attended a multidisciplinary foot care clinic in Sweden from January 1995 to December 1998. As consecutive patients were included, the risk of sampling bias is low. Baseline patient and disease factor data were collected prospectively using established forms during clinic visits. There is a risk of recall bias as patients were asked to remember details from a DFU that occurred in the past. That risk increases for

those patients who had non-healing DFUs farther in their past. Unfortunately, the impact on the results is unknown, as patients may over- or under-estimate the impact of their DFU on their HRQOL. Furthermore, only patients who were still alive could have completed the questionnaire, thus biasing the results towards patients who are still alive (selection bias). Patients who died between January 1995 and December 1998 may have reported different HRQoL related to their DFU than patients who were alive to participate in the study. A further risk of bias is in the rate of non-responders. Of 457 questionnaires sent, 32% were not returned. While the authors noted that there were no differences in foot ulcer status or other select characteristics between responders to the survey and non-responders, no information was reported in the publication. Therefore, the possibility of bias due to non-response cannot be ruled out. While the extent of this bias is unknown, given that 32% of questionnaires were not returned, the significant potential exists for the results to be biased by those who did not respond.

Coffey et al (19) provided Quality of Well Being index (QWB) questionnaires (31) to patients with diabetes as they attended regularly scheduled visits to the endocrinology, diabetes and ophthalmology clinics at the University of Michigan Health System between June 1998 and March 2001. The risk of recall bias was low as patients were asked to complete their questionnaires at the same time that data on patient and disease characteristics were collected, including information

about their diabetes complications, which included information about whether the patient had a DFU. While a total of 2,048 patients completed questionnaires, only 2,041 were used in the study analyses due to missing values for seven patients. The authors also reported that the response rate to the survey was 88%. A 12% non-response rate carries with it a low risk of bias in the results; however, as noted previously, there remains a risk of bias in the estimates of health utility, as patients with more morbidities—and, possibly poorer HRQoL—may be less likely to complete and return surveys than patients with fewer morbidities (37). Of the 2,041 respondents analyzed, only 149 had a neuropathic sore (assumed, for the purposes of this study, to be a non-healing or active DFU). Therefore, the data from even a few non-responders who had a non-healing DFU could have a substantial effect on the health utility estimates for those patients. Lastly, the authors did not report the specific sampling method (e.g., consecutive sampling, convenience, etc), therefore the risk of sampling bias is unknown, but cannot be ruled out.

Morgan et al (21) sent EQ-5D-3L questionnaires, six weeks after discharge, to 150,113 patients discharged from clinics in Cardiff, UK between January 2002 and July 2005. A total of 50,258 responses were received. Of those, 4,502 patients had diabetes. The authors did not report specific data on non-responders; however, of the 150,113 surveys sent out, 33.5% were returned. The authors state that the demographic characteristics of the responders were

similar to those of the local health care system (from which the sample was drawn). However, the authors also reported a higher proportion of patients with DFU in the health care system population than in the responders (18.1% and 14.7%, respectively). In addition, more males were in the responders (58.0%) than in the health care system population (53.4%). The proportion of smokers and the mean BMI were unknown for the health care system population. Given the limited data reported, the risk of a non-response bias is unknown; however, it cannot be ruled out, as there may or may not exist systematic differences between the baseline characteristics of non-responders and responders that could bias the results. Out of the 4,502 responders with diabetes, 661 had a DFU. Patients may or may not have received a specific treatment for DFU—that information was not reported—which could confound the interpretation of the results, as some patients may have had one or more treatments for their DFU, but with no response to treatment, or some may have had no treatments. Data on baseline patient and disease characteristics were collected routinely (i.e., prospectively) from patients when they attended clinics. There is a risk of recall bias for the health utility data as patients responded to the questionnaire six weeks after discharge.

Javanbakht et al (22) identified a sample of 3,918 patients with Type II diabetes mellitus using a multi-stage cluster sampling method. The risk of sampling bias is low as the authors took steps to ensure that the sample would be representative

of the target population. All study participants were interviewed by study personnel and asked to complete an EQ-5D-3L questionnaire as well as a patient and disease characteristics (i.e., clinical history) questionnaire. A total of 3,472 patients completed both questionnaires. The authors did not report information on the non-responders, therefore the risk of bias is unknown (but cannot be ruled out), as there may or may not exist systematic differences between responders and non-responders that could bias the results. Of the 3,472 respondents, 372 had a lower extremity lesion. Patients may or may not have received a specific treatment for their DFU—that information was not reported—which could confound the interpretation of the results, as some patients may have had one or more treatments for their DFU, but with no response to treatment, or some may have had no treatments. The risk of recall bias is low, as the questionnaires were administered at the same moment in time by trained interviewers.

Siersma et al (23,24) included 1,232 patients with a new DFU who presented at clinics at one of 14 centres in Europe, between September 2003 and October 2004. Of note, the authors excluded patients treated for an ulcer on the ipsilateral foot during the previous 12 months or with a life expectancy less than 1 year. The authors reported that all patients with a new DFU were asked to participate (28), therefore the risk of sampling bias is low. A total of 968 patients were followed for one year (144 were lost to follow-up; 50 had a major amputation, and; 70 died) and data were reported for 131 patients whose DFU

did not heal within one year. The remaining 837 patients had DFUs that healed within the one year follow-up period and were therefore excluded from this review. Of note, no data were reported on the number asked to participate versus the number who responded. Therefore, while the risk of bias due to non-response is unknown, it cannot be ruled out. Patient and disease characteristic data were collected at the time of the patient's initial clinic visit and monthly follow-up visits using standardized case report forms by trained investigators. Also at the time of the first clinic visit and each monthly follow-up visit, patients completed the EQ-5D-3L, thus limiting the risk of recall bias. The authors did not report the types of centres involved in the study. If the clinics from which participants were identified were mostly academic or highly specialized, then there is a risk of selection bias in the sample, as those centres would be more likely to see patients with advanced or difficult to treat disease.

Sobol et al (25) identified 185 patients with DFU, between April 2012 and May 2013, treated in an ambulatory care clinic. The risk of sampling bias is unknown as the authors did not report further details regarding how the patients were sampled (e.g., consecutive or convenience sampling); however, it cannot be ruled out. A total of 179 patients completed an EQ-5D-3L questionnaire and had information available on the severity of ulceration. The authors did not report information on the six non-responders; however, the small number of non-responders is unlikely to substantially influence the results had they responded.

Li et al (26) included 103 patients with a DFU treated at an academic and specialized clinic in Toronto, Canada. Enrolment of patients from one specialized clinic increases the risk of selection bias as such a centre would be more likely to see patients with advanced or difficult to treat disease, which could bias the results, as such patients would be likely to have poorer clinical and patient-reported outcomes than the majority of patients with DFUs. Patient and disease characteristics were collected at baseline, and patients were also asked to complete an EQ-5D-3L questionnaire, thus eliminating the risk of recall bias. Of note, all patients had a minimum duration of DFU of four weeks.

Sothornwit et al (27) included 254 patients with diabetes, of whom, 98 had a DFU. The authors did not report the number of patients who were asked to participate in the study, did not report information on non-responders (if there were in fact any non-responders), and did not report information on how the patients were sampled. Therefore, the risk of bias due to non-response or due to sampling is unknown, but neither can be ruled out. The authors also did not report on the timing of the administration of the EQ-5D-5L questionnaire compared with the timing of the patient's DFU, therefore the risk of recall bias is unknown. Furthermore, the authors did not report how patient and disease characteristics were collected (e.g., retrospectively from patient charts, prospectively at time of enrolment, retrospectively from patient via interview or

questionnaire). The risk of bias associated with the collection of these data is also not possible to determine.

The baseline characteristics of the patients with non-healing DFU included in the studies that indirectly estimated health utilities can be found in Table 4. Of note, baseline information was not reported for patients with non-healing DFUs in five studies (18,19,21,22,27). The number of patients with non-healing DFUs included in those five studies ranged from 56 (18) to 661 (21). For the remaining three studies, the number of patients with non-healing DFUs was 103 in Li et al (26), 179 in Sobol et al (25) and 131 in Siersma et al (23,24). The mean age of patients ranged from 61 years (26) to 65 years (23,24). The proportion of male patients was 72% in Siersma et al (23,24) and 67% in Li et al (26) (Table 4), but this was not reported for Sobol et al (25). The mean duration of diabetes was 18 years in Sobol et al and 16 years in Li et al. Siersma et al (23,24) reported that 11% of patients had diabetes for less than 5 years, 16% had diabetes for 5-10 years, and 73% of patients had diabetes for more than 10 years. Only Sobol et al (25) and Li et al (26) reported information on the severity of the patients non-healing DFUs; however, Sobol et al did not report what grading system was used to determine grade of ulcer depth and tissue loss or grade of infection (25). Li et al (26) reported that 44.6% of patients had Wagner Grade 2 ulcers, 49.5% had Grade 3, and 5.8% had Grade 4. No patients had Grade 1 or 5 ulcers. None of the studies reported the ethnicity of patients with DFUs.

Studies That Directly Elicited Health Utilities

One study directly measured health utility using a time trade off technique to map preferences for various health states directly to a health utility scale.(20) Preferences were obtained from a sample of 107 participants that was representative of the general Dutch population (i.e., the sample was not limited to patients with DFUs). Of note, the authors reported that they used quota-stratified sampling to ensure representativeness to the general population based on gender and age. As other factors besides age and gender may be associated with health utility, there remains a risk of sampling bias if other demographic and disease factors are not similar to the general population. Unfortunately, the authors did not report demographic data for the included participants, therefore the risk of sampling bias is unknown, but cannot be ruled out. All 107 participants were asked to indicate their preference for 13 health states identified by the authors, based on ulcer status (no active ulcer, active uninfected ulcer, and active infected ulcer) and amputation status (no previous amputation, only toes amputated [one or more], one foot amputated, one leg amputated, or both feet or legs amputated). For the purposes of the current study, only the health states with active ulcers (either uninfected or infected) were considered. As the amputation status referred to an amputation that pre-existed the current DFU, all of the health states defined by active uninfected or infected ulcer were included regardless of amputation status. Of note, the authors did not elicit health utility

estimates for the health states with both feet or legs amputated and having either an uninfected or infected active ulcer.

Overall

None of the studies reported a justification for the chosen sample size, with the exception of Li et al (26). In that study, a randomized controlled trial, Li et al reported a sample size that was based on comparing the efficacy of HBOT and high-quality foot care compared with high quality foot care alone on amputation rates. As part of this randomized control trial, HRQoL data was collected. As a result, the baseline EQ-5D data for both groups were included in this systematic review.

The applicability of the results of the identified studies to a Canadian population is variable. Li et al (26), while they sampled from a Canadian centre, were limited to a single centre which may have more patients with advanced and/or difficult to treat disease with a poorer prognosis and therefore may not be representative of the general Canadian population. The authors used Canadian tariffs to obtain health utility values. Six of the remaining studies (18-21,23-25) included participants from Western nations; however, similarly to Li et al (26), all six also identified patients from diabetes clinics or more specialized DFU clinics. The remaining two studies were conducted in Thailand (27) and Iran (22), which would limit applicability to a Canadian population. Further complicating the

results is the fact that the seven studies that indirectly elicited health utilities used varying tariffs to map the reported preferences to a health utility scale. Tariffs used included those for a UK population (18,22), a US population (19), or a Polish population (25). Three of the studies did not report the tariffs used to calculate the health utilities (21,23,24,27).

There was a high degree of heterogeneity between the included studies due to several factors. First, there was a lack of reporting of patient and disease characteristics for the included participants in the studies, which makes establishing the homogeneity or heterogeneity of the study populations difficult; however, based on the countries where the studies were conducted, there would exist differences in ethnicity, the health systems, and the continuum of care for patients (e.g., timing of diagnosis; available treatments; identification and treatment of comorbidities; etc) between the studies. There also existed differences between the studies in how utilities were estimated (i.e., direct or indirect estimates; tariffs used to calculate indirect estimates), differences in patient populations for indirectly estimated health utilities (e.g., differences in the geographic [country] location of each study), differences in the objectives of each study, and differences in the applicability of the studied population of patients with DFU between each study as well as with the current study. As a result, combining the study results in a meta-analysis was not appropriate.

Study Results

Unadjusted Estimates of Health Utility for Patients with Non-Healing DFUs

Unadjusted mean (and variance) estimates of health utility for patients with non-healing DFUs for each study can be found in Table 4. The estimates ranged from an indirectly obtained mean of 0.44 in 56 patients included in Tennvall et al (18) to a directly obtained mean of 0.89 for active uninfected ulcers with no previous amputation estimated from 107 participants included in Redekop et al (20). Redekop et al provided some of the highest estimates of health utility for various health states for uninfected or infected active ulcers and amputation status (Table 4). Among studies that indirectly elicited health utilities, mean estimates ranged from 0.44 in Tennvall et al (18) to 0.703 in Sothornwit et al (27). Sothornwit et al (27) reported health utility data for 98 patients with DFUs and 43 patients with amputations combined; however, the authors also noted that there was no difference in the mean utility value between the two groups of patients. Siersma et al (23,24) reported data for 131 European patients with DFUs unhealed within one year and reported that the mean health utility was 0.645 (standard deviation, 0.308); however, the tariff used was not reported. Li et al (26) included 103 patients with chronic non-healing DFUs and reported a mean health utility of 0.585, using a Canadian tariff. Both Tennvall et al (18) and Javanbakht et al (22) reported that they used a UK tariff to determine the health utility of patients with non-healing DFUs, with mean health utility scores of 0.44 and 0.62, respectively.

Table 4. Included studies: health utility estimates for patients with chronic non-healing diabetic foot ulcers.

Study	**Measure-ment Tool**	**Scoring Algorithm**	**Health Utility Estimate, unadjusted mean (95% CI)**	**Adjustment for Covariates (description of covariates used, if any)**	**Health Utility Estimate (adjusted)**
Tennvall et al (18)	EQ-5D-3L	UK tariffs	0.44 (-0.59 to 1) (min-max) n=56	Amputations and "healed" status examined. No other factors were reported.	NA
Coffey et al (19)	QWB	US (QWB)	Type 1: 0.504 (NR) Type 2: 0.474 (NR)	Regression model included neuropathic sores as one of many factors assessed (e.g., sex, BMI, nephropathy, neuropathy, stroke, etc)	For a Type 1 diabetes patient without any complications except neuropathic sores, mean health utility was 0.596 for males and 0.563 for females For a Type 2 diabetes patient without any complications except neuropathic sores, mean health utility was 0.590 for males and 0.552 for females

Study	**Measure-ment Tool**	**Scoring Algorithm**	**Health Utility Estimate, unadjusted mean (95% CI)**	**Adjustment for Covariates (description of covariates used, if any)**	**Health Utility Estimate (adjusted)**
Redekop et al (20)	Directly measured-TTO	Directly measured-TTO	<u>Active, uninfected ulcers</u> No previous amputation: 0.89 (0.86-0.91) 1 or more toes amputated: 0.8 (0.76-0.84) 1 foot amputated: 0.74 (0.70-0.78) 1 leg amputated: 0.66 (0.62-0.71) <u>Active, infected ulcers</u> No previous amputation: 0.82 (0.79-0.85) 1 or more toes amputated: 0.75 (0.71-0.79) 1 foot amputated: 0.68 (0.64-0.72) 1 leg amputated: 0.62 (0.57-0.67)	As patients with a previous amputation would return to a diabetic health state, the utilities for each state were adjusted by combining the utility for diabetes (0.84) with the utility value for each amputation state (because no patient would return to a perfect utility)	<u>Active, uninfected ulcers</u> No previous amputation: 0.75 (0.71-0.79) 1 or more toes amputated: 0.68 (0.64-0.73) 1 foot amputated: 0.63 (0.59-0.68) 1 leg amputated: 0.57 (0.53-0.62) Active, infected ulcers No previous amputation: 0.70 (0.66-0.75) 1 or more toes amputated: 0.65 (0.60-0.69) 1 foot amputated: 0.59 (0.54-0.63) 1 leg amputated: 0.55 (0.50-0.59)

Study	Measure-ment Tool	Scoring Algorithm	Health Utility Estimate, unadjusted mean (95% CI)	Adjustment for Covariates (description of covariates used, if any)	Health Utility Estimate (adjusted)
Morgan et al (21)	EQ-5D-3L	NR	DFU without other complications: 0.512 (SD=0.325), n=239 DFU + other complication(s): 0.424 (SD=0.340), n=422 DFU with or without other complications: 0.455 (SD=0.337), n=661 DFU+CHD+CVD+RET: 0.293 (SD=0.418), n=12 DFU+CHD+CVD: 0.373 (SD=0.348), n=26 DFU+CHD+RET: 0.389 (SD=0.343), n=62 DFU+CHD:	Regression model using sex, age, BMI, CHD, stroke, DFU, ESRD, peripheral vascular disease, and RET.	Regression model: Intercept 1.068, p<0.001 Male 0.052, p<0.001 Age (years) -0.002, p<0.001 BMI (kg/m^2) -0.012, p<0.001 CHD -0.066, p<0.001 Stroke -0.114, p<0.001 DFU -0.069, p<0.001 ESRD -0,082, p<0.063 PVD -0.063, p<0.030 Retinopathy -0.029, p<0.031

Study	Measure-ment Tool	Scoring Algorithm	Health Utility Estimate, unadjusted mean (95% CI)	Adjustment for Covariates (description of covariates used, if any)	Health Utility Estimate (adjusted)
			0.414 (SD=0.343), n=188 DFU+RET: 0.497 (SD=0.314), n=78 DFU+CVD: 0.58 (SD=0.31), n=25		
Javanbakht et al (22)	EQ-5D-3L	UK VAS tariff	0.62 (0.59-0.65)	No adjustments made to account for the effect of baseline factors on health utility of patients with DFU.	NA
Siersma et al (23,24)	EQ-5D-3L	NR	0.645 (SD=0.308)	Investigated several baseline factors using a linear regression to determine association with EQ-5D score; however, no results of the linear regression were reported (i.e., coefficients, p-values, etc)	NA

Study	Measure-ment Tool	Scoring Algorithm	Health Utility Estimate, unadjusted mean (95% CI)	Adjustment for Covariates (description of covariates used, if any)	Health Utility Estimate (adjusted)
Sobol et al (25)	EQ-5D-3L	Polish tariff	0.618 (SD=0.320[A])	Ulcer size and severity investigated for effect on health utility estimate	Very weak negative correlation found between ulcer size and EQ-5D utility value. No Correlation found between ulcer severity and utility value.
Li et al (26)	EQ-5D-3L	Canada tariff	Placebo (n=54): 0.59 (SD=0.25) HBOT (n=49: 0.58 (SD=0.19)	Correlations or adjustments were not investigated for baseline EQ-5D utility values	NA
Sothornwit et al (27)	EQ-5D-5L	Directly measured-TTO	0.703 (SD=0.28[A]) Note: included 98 patients with DFU and 43 with amputations; however, the authors reported no difference in the mean utility value between those two subgroups.	NR	NR

Notes: BMI=body mass index; CHD=cardiovascular heart disease; CVD=cerebrovascular disease; DFU=diabetic foot ulcer; ESRD=end-stage renal disease; RET=retinopathy; SD=standard deviation.
[A]The authors did not specifically state that this value was a standard deviation; instead, it was assumed from the available information.

Adjusted Estimates of Health Utility for Patients with Non-Healing DFUs

While six of the nine included studies reported that they investigated factors related to health utility scores (e.g., using regression analysis) for patients with non-healing DFUs, only three of the studies included information on these investigations in the results (19,21,25). Coffey et al (19) investigated the health utility of patients with diabetes, and conducted regression analyses—one for Type 1 diabetes and a second for Type 2—that included the following factors: sex, BMI, diabetes intervention, retinopathy, nephropathy, neuropathy (including DFU), stroke, cardiovascular disease, and high blood pressure. The authors estimated, based on the results of their regression analyses, that for a male patient with Type 1 diabetes without any complications except DFU, the mean health utility was 0.596; and, for a female, 0.563. For a male patient with Type 2 diabetes, also without any complications except a DFU, the mean health utility was 0.590; and, for a female, 0.552. The authors reported that the regression models for Type 1 and Type 2 diabetes fit the data moderately well, with R^2=45.0% and 36.3%, respectively. They also noted that both models did not show a significant lack of fit (p=0.664 and p=0.643, respectively).

Morgan et al (21) aimed to derive and compare the health utility of patients with diabetes and single and multiple morbidities. The authors conducted multiple regression analysis using sex, age, BMI, coronary heart disease (CHD), stroke, DFU, end-stage renal disease (ESRD), peripheral vascular disease and

retinopathy. The authors reported that a general linear model with EQ5D as the dependent variable, predicted by gender ($p<0.001$), age ($p<0.001$), BMI ($p<0.001$), CHD ($p<0.001$), stroke ($p<0.001$), DFU ($p<0.001$), ESRD ($p<0.063$), peripheral vascular disease ($p<0.030$), and retinopathy ($p<0.031$), explained 9.2% of the variance in the EQ5D health utility estimates. The authors reported that among 422 patients with a DFU and other complications, the health utility was 0.424 (standard deviation, 0.340). As seen in Table 5, patients with more co-morbidities, in addition to a DFU, had lower mean health utility scores than patients with less co-morbidities. Of note, the regression model indicated that DFU resulted in a decrement of 0.069 to the predicted health utility score, BMI a decrement of 0.012 for every kg/m^2, age a decrement of 0.002 for every year of age, CHD a decrement of 0.066, stroke a decrement of 0.114, ESRD a decrement of 0.082, peripheral vascular disease a decrement of 0.063, and retinopathy a decrement of 0.029.

Sobol et al (25) investigated the impact of ulcer size and severity on the health utility of patients with DFU. The authors reported that they found a very weak negative correlation between ulcer size and health utility and no correlation between ulcer severity and health utility; however, no correlation statistics were reported.

The remaining studies did not adjust the health utility values or investigate factors that may be associated with health utility scores for patients with non-healing DFUs. Of note, Siersma et al (23,24) reported that several baseline factors were investigated, using linear regression, to determine if any were associated with EQ-5D score in patients DFU; however, no results of the linear regression were reported. Finally, Redekop et al (20) adjusted the health utility for health states with a previous amputation by combining the utility for diabetes (0.84) with the utility value for each amputation state, because the authors felt that no patient would return to a perfect utility after an amputation. The adjusted utility values for each infected and uninfected DFU state, by type of amputation, can be found in Table 4. However, while the authors reported that they planned to check for association between sex and age of participants and the utility values that they assigned to health states, no data were reported on the association of those factors and health utility.

2.4 DISCUSSION

Several studies were identified that provided estimates of the health utility of patients with non-healing DFUs. The heterogeneity between the studies made a meta-analysis of the results not possible. Unfortunately, the heterogeneity also makes the interpretation of the disparate results difficult. Firstly, the included studies had very different objectives, not all of which aligned perfectly with the objectives of this systematic review. The objectives of the current study were to identify estimates of health utility for patients with chronic, non-healing DFUs and to identify factors associated with health utility scores in those patients. Six of the included studies investigated the health utility of patients with non-healing DFUs (18,20,23-27). The remaining studies had broader objectives and, more generally, they investigated the health utility of patients with diabetes (19,21,22). However, those studies also reported data for a subgroup of patients with chronic non-healing DFU, two of which also investigated factors (e.g., comorbidities, demographics, etc) associated with health utility scores (19,21). All nine studies had objectives that at least partially aligned with the objectives of this study. Eight of the studies estimated health utility indirectly, with seven studies using a version of the EQ-5D (18,21-27) and one using the QWB-SA (19). Of note, the EQ-5D-3L, the EQ-5D-5L and the QWB-SA are validated and appropriate instruments to measure health utility. These instruments estimate health utilities in two general steps. The first step utilizes a questionnaire and/or interview to define the health state that each patient is in. The second step is to use tariffs

(i.e., preference weights obtained from a sample of the general population, generally using a time trade off or standard gamble technique) to map the health state to a health utility score. The QWB-SA used tariffs determined from a sample of individuals from San Diego, United States (31), whose preferences are not identical to a Canadian population, but may be similar. For the EQ-5D instruments, the -3L and -5L require different sets of tariffs. Of note, tariffs are available for various countries for both the -3L and -5L instrument. Li et al (26) was the only study that used a Canadian tariff to calculate health utility scores. Of note, Tennvall et al (18) used a UK tariff, and while the preferences of a UK population are not identical to a Canadian population, they may be similar. Sobol et al (25) used a Polish tariff, which also may have different preference values from a Canadian population. While Siersma et al (23,24) and Sothornwit et al (27) both indirectly elicited health utilities, neither reported the tariff used to calculate the health utility scores. As Siersma et al (23,24) conducted their study in several European countries, and as tariffs for multiple European countries are available for the EQ-5D-3L, it is not possible to reasonably hypothesize which tariff the authors may have used. Sothornwit et al (27) used the EQ-5D-5L to estimate health states, but there is no published tariff for Thailand (38), therefore, no reasonable assumption about which tariff was used can be made.

The choice of preference weight is important as different populations often value a given health state differently, which is why there exist several tariffs for the EQ-

5D instruments, with each tariff specific to a certain country. In general, and all else being equal, if one wants to estimate the health utility of a given health state for a Canadian population, estimates based on preference weights derived from a Canadian population would be most appropriate and only the study reported by Li et al (26) used a Canadian preference set.

One study, Redekop et al (20), used a time trade off technique to directly elicit health utilities from a Dutch population. An advantage of this approach is that the health utilities for each specific health state defined by ulcer status and by previous amputation status were obtained directly from a sample of the general population. In short, the health utility scores reflect the preference for each health state that the participants were asked to value (uninfected and infected active ulcers, by previous amputation status). As the participants were sampled from the general population, they were not stating a preference for a health state that they were living, rather they were stating a preference for an imagined health state. One of the disadvantages of this approach is related to the ability of the interviewers to accurately describe to the participants the experience of living in each health state. In contrast, the indirect elicitation approach asks patients with the condition of interest, who are living in the health states of interest, to complete a questionnaire which equates the patient's health state to a standardized health state. The tariffs then map that health state to a health utility based on preference weights for those health states based on validated preference studies.

While both indirect and direct techniques are valid, Redekop et al (20) reported health utilities for patients with active ulcers that are infected (0.70, adjusted for utility of diabetes without ulcers or amputation) or uninfected (0.75, adjusted), that were higher than those obtained in the remaining studies. Even when accounting for previous amputations, the range of mean health utilities for infected ulcers (0.55-0.65) and uninfected ulcers (0.57-0.68) was still high. Unfortunately, the remaining studies did not report data on health utilities for patients with DFU and previous amputations, making it difficult to compare results. However, even if those results were available, such cross-study comparisons would need to be interpreted with a high degree of caution as there are differences between the studies other than the valuation method that would confound any observed differences (or similarities). Of note, the preferences of a Dutch population may not be the same as a Canadian population. As Redekop et al (20) seemed to be the single outlying study with higher health utilities for these patients, it is less likely to be a useful estimate of the health utility of Canadian patients with chronic non-healing DFU.

Siersma et al (23,24) included a sample of patients with DFUs that had all experienced their DFU for one year (n=131). However, the authors did not report the tariffs used to calculate the health utility scores, which makes interpreting the results difficult. Without knowing the tariff applied to the sample of European patients, one cannot put the health utility results into context. More succinctly, it is

not possible to fully judge the applicability of the results of this study to a Canadian population.

All of the studies had limitations in their designs. Furthermore, they all suffered from a lack of reporting of study and patient details which precluded the possibility of conducting a full critical appraisal and made it difficult to determine the generalizability of the results of each study to a Canadian population with chronic non-healing DFUs. Of note, six of the included studies did not report the length of time that patients had experienced their non-healing DFU. This increases the uncertainty in the results of these studies, which further confounds the interpretation of the results. Although Li et al (26) included a small sample of patients with chronic non-healing DFUs (n=103), the authors conducted the study in a Canadian population, used a validated instrument to measure health utility (EQ-5D-3L) and used a Canadian tariff to obtain health utility scores. The health utility of 0.585 obtained from that study may provide the best currently available estimate of health utility for Canadian patients with chronic non-healing DFUs.

Only three of the included studies reported any information on factors that may be associated with health utility in patients with non-healing DFUs (25) or with diabetes (19,21). Sobol et al (25) reported a very weak negative correlation between ulcer size and health utility and no correlation between ulcer severity and health utility among 179 patients with DFUs. Coffey et al (19) and Morgan et

al (21) investigated factors associated with the health utility of patients with diabetes. They both included DFU as one of the factors in several multiple regression analyses. Coffey et al (19) concluded that a combination of sex, BMI, diabetes intervention, retinopathy, nephropathy, neuropathy (including DFU as a factor), stroke, cardiovascular disease, and high blood pressure predicted the health utility of 2,041 patients with diabetes (149 had a DFU). The model for Type 1 diabetes explained 45% of the variability in the health utility data, while the Type 2 model explained 36.3% of the variability. Morgan et al (21) noted that the health utility of 4,502 patients with diabetes (661 had a DFU) was predicted by gender, age, BMI, coronary heart disease, stroke, DFU, end-stage renal disease, peripheral vascular disease, and retinopathy; however, this model accounted for only 9.2% of the variance in the estimates of health utility. The results of these studies suggest that several factors may be associated with the health utility of patients with DFUs. While Sobol et al (25) found only a weak negative correlation between ulcer size and health utility and no correlation between ulcer severity and health utility, these results may be due to the small size of the study. That is, the sample size may have been too small and the study underpowered to find a correlation; however, due to the reporting of this study in abstract form only, a full critical appraisal was not possible, therefore the results should be interpreted with a great amount of caution. Coffey et al (19) and Morgan et al (21) both investigated DFU as one of several factors associated with the health utility of patients with diabetes. While both studies suggested that

gender, age, BMI, cardiovascular disease, stroke, retinopathy, DFU were associated with health utility, it is not known if the same factors (excluding DFU) are associated with the health utility of patients with a DFU. Given the uncertainty in the results of the identified studies, further study is warranted to determine what factors are associated with the health utility of patients with chronic non-healing DFUs.

There are some limitations to this systematic review. While every effort was made to ensure that both reviewers that screened articles for inclusion were applying the eligibility criteria equally, the kappa score (0.52) suggests a moderate amount of agreement also suggests that there was at least some disagreement. This increases the risk that some possibly relevant citations could have been excluded. The risk of this was mitigated by having the two reviewers retrieve the full texts for all of the disagreements and review them along with all of the citations where the reviewers agreed that full text review was required. However, the disagreement does suggest that there is a risk that the study eligibility criteria may have been open to some misinterpretation by the reviewers. Therefore, there remains a low risk that the reviewers both classified some citations as not relevant when, in fact, they may have been.

A second limitation is the risk of publication bias. The current study attempted to limit the risk of this bias by including several databases in the planned literature

search (Medline, Embase, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, Database of Abstracts of Reviews of Effects, the Health Technology Assessment Database, and the NHS Economic Evaluation Database). Embase itself indexes the conference proceedings from a large number of medical professional societies and these abstracts were not excluded. Furthermore, publication bias is a risk for studies investigating and comparing interventions to identify the treatment effect. For example, RCTs of interventions, which provide a comparative effect estimate for an intervention to some comparator, may produce "negative" results, that is, they fail to reject the null hypothesis of the study (that there is no statistically significant difference in the effect of the interventions being compared). These "negative" trials have been shown to be less likely to be published than trials that have demonstrated statistically significant differences in treatment effect. This means that published literature for a given intervention may be biased in favour of that intervention; however, this study was not investigating the comparative effects of treatments. Instead, it was focused on identifying estimates of health utility for patients with chronic non-healing DFUs. This alone reduces the risk of publication bias as the results would not be considered "positive" or "negative". Therefore, given this, as well as the inclusion of multiple databases in the search strategy, the risk of publication bias in this review is considered to be low.

A third limitation of the study was the inability to obtain one possibly relevant full text article for review (39). There exists a possibility that the study may have contributed relevant information to this review; however, the study appears to have been conducted in Libya, and the applicability of the results from that population and health system to the Canadian context may be limited.

An additional limitation in conducting this review was the lack of guidance on the critical appraisal and proper conduct of studies that are designed to measure, analyze and provide estimates of health utility. The authors of this review selected the AXIS tool to assist with the critical appraisal of the included studies, as eight of the studies had a cross-sectional design. While it is acknowledged that the AXIS tool was not designed to specifically appraise studies investigating health utility, it was designed to appraise cross-sectional studies. As this review's objective was to identify an estimate of the health utility of patients with chronic non-healing DFUs, the authors considered that health utility estimates for these patients would be collected from subjects in a cross-sectional manner, that is, subjects would be asked to complete a questionnaire or participate in an interview at a single point in time. This would provide the investigator with an estimate of health utility for that particular moment in time for that patient. Therefore, the authors felt that the principles of conducting a cross-sectional study would be a good starting point for the critical appraisal of studies collecting

health utility data for the patient population of interest to this thesis. For the RCT, the Cochrane Risk of Bias tool was also used.

While Li et al (26) may provide the best available estimate of health utility for patients with non-healing DFU, several factors may affect the estimates of health utility. Li et al (26) included a small population of patients, therefore the health utility estimate may not accurately capture the variance in the parameter that would be seen in the Canadian population. Given the uncertainty in the estimate, other studies identified in this review may also provide a reasonable estimate of health utility for this group of patients, therefore, future health economic evaluations in the Canadian setting, and utilizing a health state defined by chronic non-healing DFUs should explore uncertainty in this parameter through scenario analyses utilizing health utility scores from other studies. The selection of which study's estimate to use should take into consideration whether the study sample is representative of the target population of the economic evaluation as well as the population from which the sample was drawn; the number of non-responders; the appropriateness of the health utility measure; the choice of tariff (for indirectly elicited health utilities); the sample size, both in terms of the precision of the estimate as well as the likelihood that the sample captured patients with relevant co-morbidities and fully represents the target population. At the current time, the best available evidence for the health utility of Canadian patients with non-healing DFUs comes from the RCT conducted by Fedorko et al (29) and presented by Li

et al (26). Given the small sample size, and the limited evidence regarding factors that may be associated with health utility of these patients, further research is warranted to validate the estimate, to provide more precision in the estimate, and to identify factors that are associated with health utility (as well as to estimate the impact of those factors).

CHAPTER 3: EXPLORATORY ANALYSIS OF ONTARIO-SPECIFIC (HBOT TRIAL) DATA TO ESTIMATE THE HEALTH UTILITY OF PATIENTS WITH DIABETES AND CHRONIC, NON-HEALING UCLERS OF THE LOWER LIMB

3.1 INTRODUCTION

Economic evaluations of healthcare technologies provide decision makers with the ability to compare the incremental costs and benefits of a number of health technologies. One commonly used type of economic evaluation is the cost-effectiveness analysis, which provides a ratio of the incremental cost over the incremental effect of a treatment relative to that of another treatment, known as the incremental cost-effectiveness ratio (ICER). Put another way, it is the difference in cost of one treatment versus another treatment divided by the difference in effectiveness of one treatment versus another treatment. The challenge with cost-effectiveness analyses is that the effectiveness may be measured differently for one disease or treatment than another, thus making comparisons using the ICER difficult if the effectiveness measure varies across treatments. Furthermore, some outcomes do not fully capture the effect on the individual; for instance, patients with DFUs may experience healing of their ulcer or they may continue in a non-healing state with an uninfected wound or that wound could become infected, which may eventually require either a minor amputation of a toe or major amputation of the lower limb. Each of those outcomes for the patient would be measured differently (i.e., one could count the outcome as whether the diabetic foot ulcer (DFU) healed, became infected or

was amputated). Each of those, in turn, would affect the denominator (i.e., the effectiveness) of the ICER as the effectiveness would differ depending on what outcome was used to measure effectiveness; if the outcome were number of healed ulcers the ICER would be different than if the outcome were number of amputations, which would be different than the ICER based on the outcome being the number of infections prevented.

For patients with non-healing DFUs, infections or amputations prevented may not be the only important outcome to consider when evaluating the cost-effectiveness of interventions. The quality of life that patients experience while they have a non-healing DFU is also an important consideration, as is the impact that an intervention may or may not have on a patient's quality of life. Nabuurs-Fransen et al reported that patients with a healed DFU had a higher HRQoL than patients with a non-healing, chronic DFU (40).

A health utility is a generic measure of HRQoL. A health utility score is a preference weight for a given health state relative to other health states, where 1 represents perfect health and 0 represents death. As it is possible for some health states to be considered worse than death, it is possible for health utility scores for certain health states to be less than 0; however, health utility scores are bound at the upper end at 1.

There are several methods to derive health utility scores or preferences. The simplest method is to ask subjects to rank health states from most preferred to least on a scale such that the spacing between the states indicates the subject's perceived preference between the states (9). Variations of this approach provide a line with intervals marked from 0-100 (a rating scale), or a visual analogue scale where a line of specific length (e.g., 10 cm) is placed on a page and is anchored with clearly defined endpoints (e.g., one end is death and the other end is perfect health). Another method is the standard gamble method. For a chronic health state, subjects are asked their preference between two alternatives. One is the certainty of being in a given health state (e.g., non-healing chronic DFU). The second alternative has two possible outcomes; a healthy state with perfect health and a defined additional number of years of life or immediate death, where the probability of death is 1 minus the probability of being in the healthy state. The probability of perfect health for the second alternative is varied until the point where the subject is indifferent between the two alternatives, which is the preference score for the given health state (9).

Another method is the time trade-off (TTO) approach developed by Torrance et al (41). For a chronic state, the subject is given two choices; the first is the life expectancy (time=*t*) of an individual with a given health state (e.g., non-healing chronic DFU), followed by death; and the second is being in a perfectly healthy state for time less than *t*, followed by death. The time in the perfectly health state

is varied until the subject is indifferent between the two choices, at which point the preference score for the given health state is the ratio of the time in the perfectly health state over the life expectancy in the given health state, *t* (9). Both the standard gamble and the TTO approaches are time and resource consuming as trained interviewers are required to explain the health states and the process to subjects. This is often prohibitive when conducting clinical trial research for several reasons; limited resources to hire and train interviewers; the length of time that interviews require within study visits; and reluctance of subjects in trials to participate in a complicated and lengthy interview process. In order to simplify the collection of health utility data, instruments such as the EuroQol-5 dimension (EQ-5D) were developed (i.e., indirect methods). The EQ-5D can be self-administered or administered with the guidance of an interviewer, thus easing the burden on subjects with the aim to increase the number of patients willing to provide health utility data. The EQ-5D asks individuals to self-report whether they have experienced problems in the following five dimensions; mobility, self-care, usual activities, pain or discomfort, and anxiety or depression (42) and is available in two versions, the 3L and the 5L. In the EQ-5D-3L (43), each question or item has three possible responses, including, “no problems,” “some problems,” and “extreme problems.” There are 243 possible combinations of responses to the 5 dimensions in the EQ-5D-3L, each describing a health state. In the EQ-5D-5L, there are five possible response, resulting in a total of 3,125 possible combination of responses, thus increasing the sensitivity of the

instrument. Once the subject's responses to the five questions are obtained, it can then be mapped to a health utility score using preference weights, which are based on preference weight studies conducted in various countries (38). These preference weight studies may use a standard gamble or TTO approach to determine health utility scores for various health states for a given population.

While health utility scores can be measured for a given health state, it may not provide a complete estimate of the health utility for all individuals in that health state. For instance, while one could define a health state as having a given condition, and health utility data could be collected from a sample of subjects who have that condition, they may have other co-morbidities or characteristics that are also associated with HRQoL, thus leading to a high degree of variation and potential confounding in the health utility scores within the defined health state. In the previously conducted systematic review, two studies presented results suggesting that age, BMI, gender, and DFUs may be associated with the health utility of patients with diabetes (19,21).

A cost-utility analysis is a specific type of cost-effectiveness analysis that makes comparisons possible between different diseases and interventions, with different outcomes by measuring effectiveness as a quality-adjusted life year (QALY) (9). In this type of analysis, the disease course that individuals may take is mapped into several health states. These health states represent changes in the disease

course that have a measurable and meaningful impact on the individual. A cost-utility analysis takes into consideration the time spent in a given health state along with the health utility associated with that health state in order to determine the amount of QALYs that could be expected for a given treatment. Therefore, a QALY of 0 would represent death, 1 would represent a year of perfect health, and a QALY of 0.5 could represent a full year at half quality of life or half a year at perfect quality of life. In this way, the QALY is a measure of the effectiveness of a given treatment in terms of both the time in each health state as well as the quality of that time. A cost-utility analysis, therefore, allows comparisons between treatments that may impact disparate outcomes, by providing a common denominator upon which to make comparisons (9).

Without reliable estimates of the incremental cost-utility ratio (ICUR) of a treatment compared with another treatment, decision makers may not be able to make funding decisions. To provide the best estimate of an ICUR, it is critical to have reliable estimates for the change in health utility for each intervention and the associated health states of the population included in the economic evaluation. Without reliable health utilities, and hence, incremental cost-utility estimates, decision makers are unable to make appropriate decisions regarding which treatments to fund. Several studies have reported health utility estimates for patients with chronic non-healing DFUs, with mean values ranging from 0.293 to 0.89 (18-27). The difference in the estimates across the studies is large and

suggests a high degree of uncertainty in the available estimates for health utility in this patient population. This uncertainty could be due to several factors, including differences in the patient demographics and disease characteristics across the studies, the valuation technique, and the method of assessment. Of these studies, the majority of studies were conducted in a population and health care system other than Canada. Only one study, Li et al (26) included patients with chronic non-healing DFUs. While the study included only 103 patients, it was conducted in a Canadian setting and used a Canadian preference weight on data obtained using the EQ-5D-3L in order to obtain health utility scores.

This study aims to estimate the health utility of Canadian patients with chronic non-healing DFUs by utilizing baseline data from an RCT originally reported by Fedorko et al (29) and to; 1) explore the baseline characteristics such as age, gender, BMI, smoking status, type of diabetes, the duration of diabetes, duration of the DFU, number of DFU wounds, wound severity as measured by the Wagner grade, and hemoglobin A1c (HbA1c) with health utility values, using multiple regression analysis; and, 2) create a descriptive regression model to calculate health utility estimates adjusted for the baseline characteristics noted above.

3.2 METHODS

The data used in this analysis were collected as part of a previously published randomized controlled trial (26,29,44). Fedorko et al reported the methods used in the trial, but in short, 107 patients with a chronic, non-healing DFU in patients with Type 1 or Type 2 diabetes, were randomized to receive treatment with hyperbaric oxygen therapy or to a sham procedure (29). Patients had to have a DFU with a Wagner grade of 2-4 of at least 4 weeks' duration. All patients receive comprehensive wound care, consisting of weekly clinical assessments and care provided by a multidisciplinary team and included infection control, debridement, prescriptions for offloading devices and advanced wound care dressings. Treatment with HBOT or sham was provided over 30 treatment sessions delivered over a 6-week period. Data were available for 103 patients in the trial. The primary outcome of that study was freedom from having or meeting the criteria for below-knee or metatarsal amputation at 12 weeks from start of treatment.

Secondary outcomes included wound measurements, wound assessments (i.e., Bates-Jensen wound assessment tool), wound classification (i.e., Wagner grade), and patient-reported outcomes (i.e., SF-36, EQ-5D-3L, Diabetic Foot Ulcers Scale-Short Form).

Outcome data were collected at randomization (baseline), following the intervention period (i.e., at week 6), and at the end of the follow-up period (i.e., at week 12). Only the baseline data for each patient are relevant for the current study; therefore, the baseline patient characteristics and EQ-5D-3L scores were used in this analysis, as this thesis was not evaluating the effect of the intervention on health utility scores.

The EQ-5D-3L provides a generic measure of patient's HRQoL. It is an indirect preference-based instrument. The EQ-5D-3L was self-administered by patients, but in the presence of a trained researcher. In this way, the patient could seek clarification on the instrument questions as they went through the questionnaire. In order to map the subjects responses to a health utility score, a Canadian preference weight was applied (35). The utility index for the Canadian preference weight value set ranges from -0.340 to 1 (35).

Statistical Analyses

Descriptive analyses of categorical variables were presented as frequencies and percentages. Continuous data (e.g., age, BMI) were presented as a mean and standard deviation (SD). Where data were missing, a multiple imputation technique was used (five imputations with regression analysis, with linear regression for continuous variables and logistic regression for categorical variables)(26). Potential associations between each categorical variable and EQ-

5D-3L-derived health utility scores were explored using boxplots and by testing for differences in the mean health utility score (t-test for dichotomous variables and Analysis Of Variance [ANOVA] for categorical variables with 3 or more levels); differences were considered statistically significant when $p<0.05$. Potential associations between each continuous variable and health utility scores were explored using scatterplots.

A descriptive model, using linear regression analysis, was used to calculate health utility scores (as derived from EQ-5D-3L) adjusted for baseline characteristics (age, BMI, type of diabetes, duration of diabetes, duration of DFU, gender, HbA1c, number of DFU wounds, smoking status, and Wagner grade) using the ordinary least squares model. Additionally, a forward selection multiple linear regression analysis was planned to assess which baseline characteristics (variables) may be associated with health utility score, using an ordinary least squares model. The first step was to perform a univariate linear regression to determine which variables individually were associated for health utility score. Those variables with an overall F-test $p<0.05$ were included and tested in the next iteration of the model. Significant variables were combined in an iterative fashion, starting with the two variables that had the highest R^2 (an indicator of the amount that the model explained the variability of the health utility data around its mean). If the overall F-test for the new model was still significant, and the R^2 of the new model was higher than the R^2 of the previous model, the new model was

kept and the variable with the model with the next highest R^2 was added. This process was continued until the adjusted R^2 did not increase compared with the previous model. Once a variable was added to the model, it was not removed at later stages.

Alternative methods for linear regression analysis include the Tobit model and the censored least absolute deviance (CLAD) model, both of which can be used for censored data. As the EQ-5D-derived health utility data are bound above at 1, a Tobit or CLAD model may be reasonable approaches; however, Pullenayegum et al (45) investigated the use of ordinary least squares, Tobit and CLAD models to analyze health utility data that are bound above at 1, and concluded that Tobit and CLAD models produced biased estimates which are not appropriate when the data are health utility data that are to be used to calculate QALYs for use in economic evaluations. Therefore, an ordinary least squares model was utilized.

All statistical analyses were conducted using R (46) and the MASS package (47).

3.3 RESULTS

In total, 103 patients with chronic non-healing DFUs were included in this analysis of patients with chronic non-healing DFUs. The baseline characteristics, without imputation for missing data, can be found in Table 5. Of note, two patients were missing data on age, nine patients were missing HbA1c values, and 17 patients were missing a baseline EQ-5D-3L health utility score. The majority of the study population, 93.2% of patients, had Type 2 diabetes. The patients had their diabetes for a mean of 15.6 years (SD=11.2), which was slightly left-skewed, with a median of 12.8 years and an interquartile range of 7.4 years to 15.6 years (Figure 2). The mean age of participants was 61.4 years (SD=11.8) and the mean BMI was 30.4 kg/m^2 (SD=5.96). Sixty-nine (67%) study participants were male and 54 (52.4%) were smokers. The majority of patients had a non-healing DFU with a Wagner grade of 2 or 3, with 5.8% of patients having a grade 4 wound. Of note, the trial only included patients with Wagner Grade 2-4 ulcers, where Wagner Grade 2 ulcers are defined as ulcers with extension to ligament, tendon, joint capsule, or deep fascia without abscess or osteomyelitis, Grade 3 ulcers are defined as a deep ulcer with abscess, osteomyelitis, or joint sepsis, and Grade 4 ulcers have gangrene localized to a portion of the forefoot or heel (48). The mean HbA1c was 8.097 mmol/mol (SD=1.954). The mean duration of DFU was 288 days (SD=414), or 9.5 months. Of note the duration of DFU data were highly left-skewed, with a median of 180 days and an interquartile range of 90 days to 365 days (Figure 3).

Figures 4-8 show boxplots of the EQ-5D-3L-derived health utility scores by type of diabetes, gender, smoking status, Wagner grade, and number of DFUs. The boxplots show the distribution of EQ-5D-3L scores by each variable and demonstrate that for each value of a given variable, the EQ-5D-3L scores in the trial do not differ by a great amount. Table 6 shows the mean EQ-5D-3L-derived health utility scores for each of the variables; however, no statistically significant differences in health utility scores were noted for any of the categorical variables (Table 6). While no statistically significant differences were demonstrated, for some values of certain variables, very low numbers of patients were included (i.e., only seven patients had Type 1 diabetes, six patients had Wagner Grade 4 ulcer, and two patients each had 3 and 4 wounds), which reduces confidence in the mean estimate for particular values of those variables.

Figures 9-13 show scatterplots of EQ-5D-3L-derived health utility scores by age, BMI, HbA1c, number of years with diabetes, and the number of years with a DFU. By visual inspection, no strong linear relationships between these variables and health utility score were apparent.

The results of the first step in the multiple linear regression analysis conducted to find variables associated with health utility score can be found in Table 8. No statistically significant associations of EQ-5D-3L-derived health utility were found

with any of the variables tested: age, BMI, type of diabetes, duration of diabetes, duration of wound (DFU), gender, HbA1c, number of wounds (DFUs), smoking status, or Wagner grade. Age was the most closely correlated variable with EQ-5D-3L health utility score, with r=0.171 and p=0.0841 (F-test for overall significance), however, this was not statistically significant at the *a priori* alpha=0.05.

While none of the tested independent variables were statistically significantly associated with EQ-5D-3L-derived health utility, a multiple linear regression of all of the identified baseline characteristics was conducted (i.e., a descriptive model) to provide health utility estimates adjusted for baseline characteristics. Table 7 provides the coefficients for each characteristic. For a male patient, 61 years old, with Type 1 diabetes for 15.6 years, who is a non-smoker and has one DFU wound with Wagner grade of 2 for 288 days, and who has an HbA1c of 8.10 mmol/mol and a BMI of 30.4 kg/m^2, the EQ-5D-3L-derived health utility score was 0.647.

Table 5. HBOT trial: baseline patient characteristics.

Characteristic	**All enrolled patients, N=103**
Age, years[A] mean (SD) median (25^{th},75^{th})	 61.4 (11.8) 61.0 (54.2,69.8)
Males, n (%)	69 (67)
BMI, kg/m^2 mean (SD) median (25^{th},75^{th})	 30.4 (5.96) 29.7 (25.5,35.5)
Wagner grade: n (%) 1 2 3 4 5	 0 (0) 46 (44.7) 51 (49.5) 6 (5.8) 0 (0)
Type of Diabetes: n (%) Type 1 Type 2	 7 (6.8) 96 (93.2)
Smoker, n (%)	54 (52.4)
HbA1c, mmol/mol mean (SD) median (25^{th},75^{th})	 8.10 (1.95) 7.70 (6.55,9.25)
Duration of Diabetes, years mean (SD) median (25^{th},75^{th})	 15.6 (11.2) 12.8 (7.4,15.6)
Duration of Wound (DFU), days mean (SD) median (25^{th},75^{th})	 288 (414.1) 180 (90,365)
Number of Wounds: n (%) 1 2 3 4	 70 (68.0) 29 (28.2) 2 (1.9) 2 (1.9)
EQ-5D-3L Health Utility: mean (SD) For 86 patients (17 had missing data) For 103 patients (imputed values for missing data)	 0.591 (0.223) 0.609 (0.226)
Notes: DFU=diabetic foot ulcer; n=number of patients; SD=standard deviation. [A]Two patients were missing data for age and were not included in the calculation of mean age.	

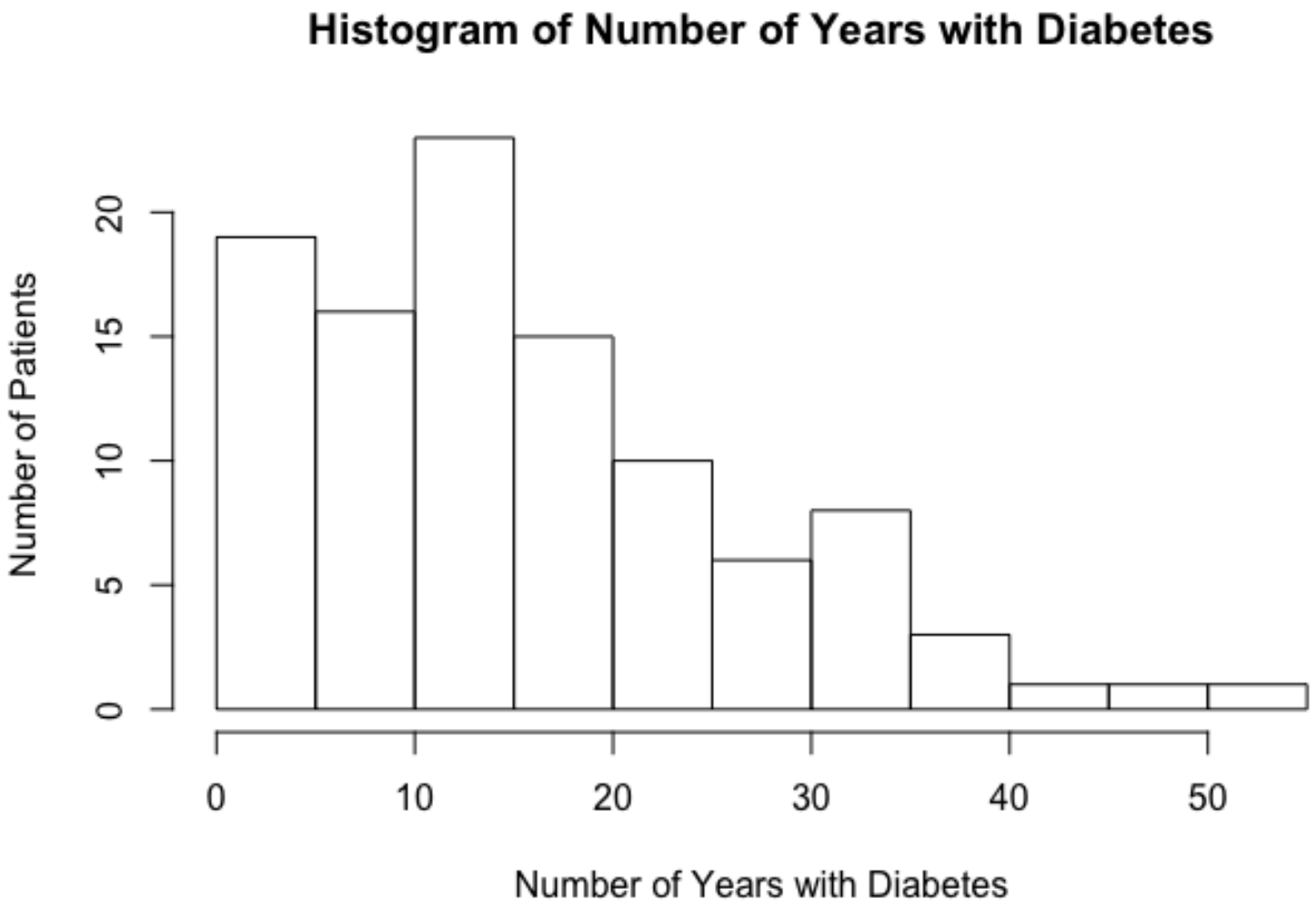

Figure 2. Histogram of number of years with diabetes.

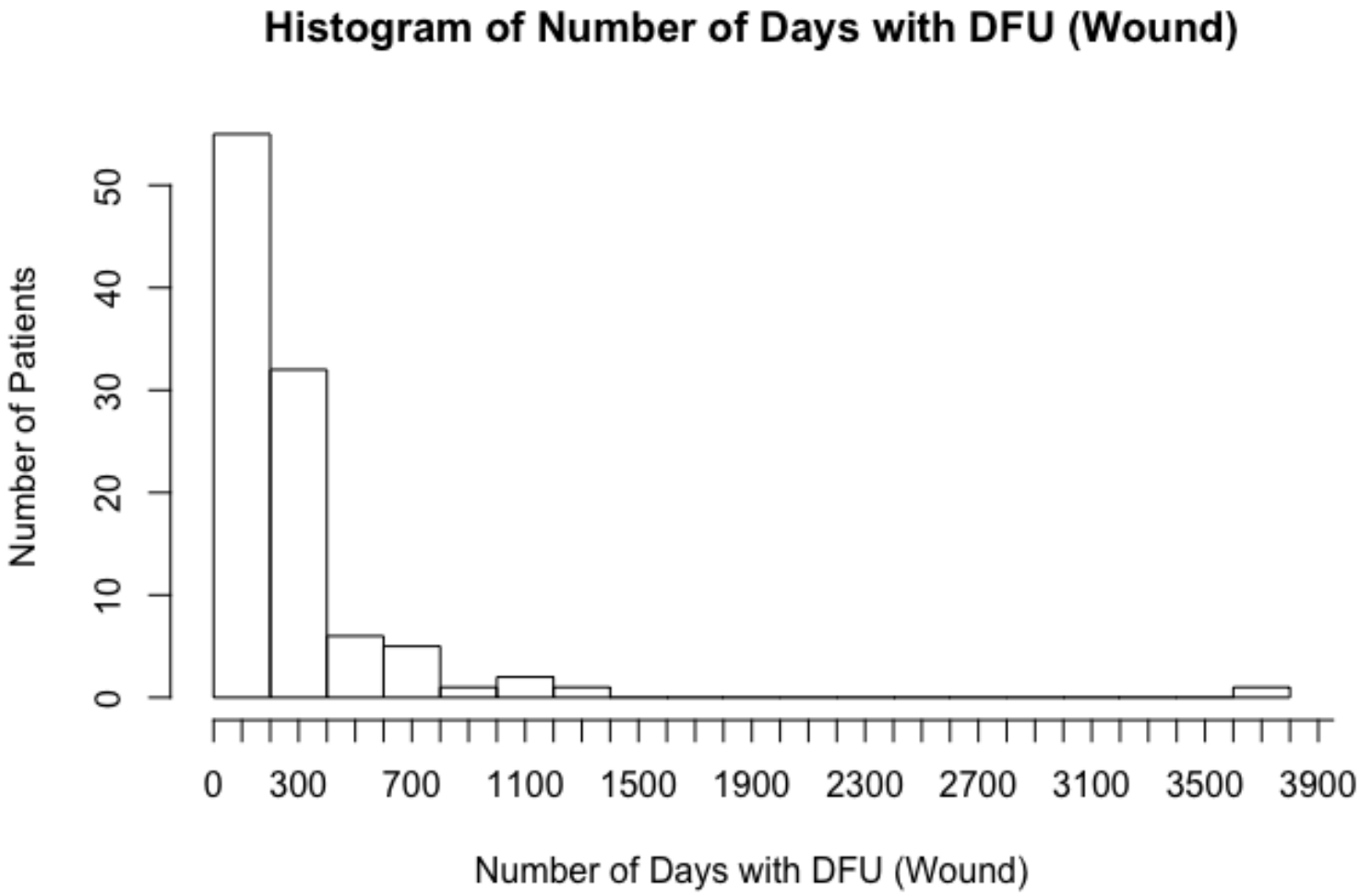

Figure 3. Histogram of number of days with chronic non-healing diabetic foot ulcer.

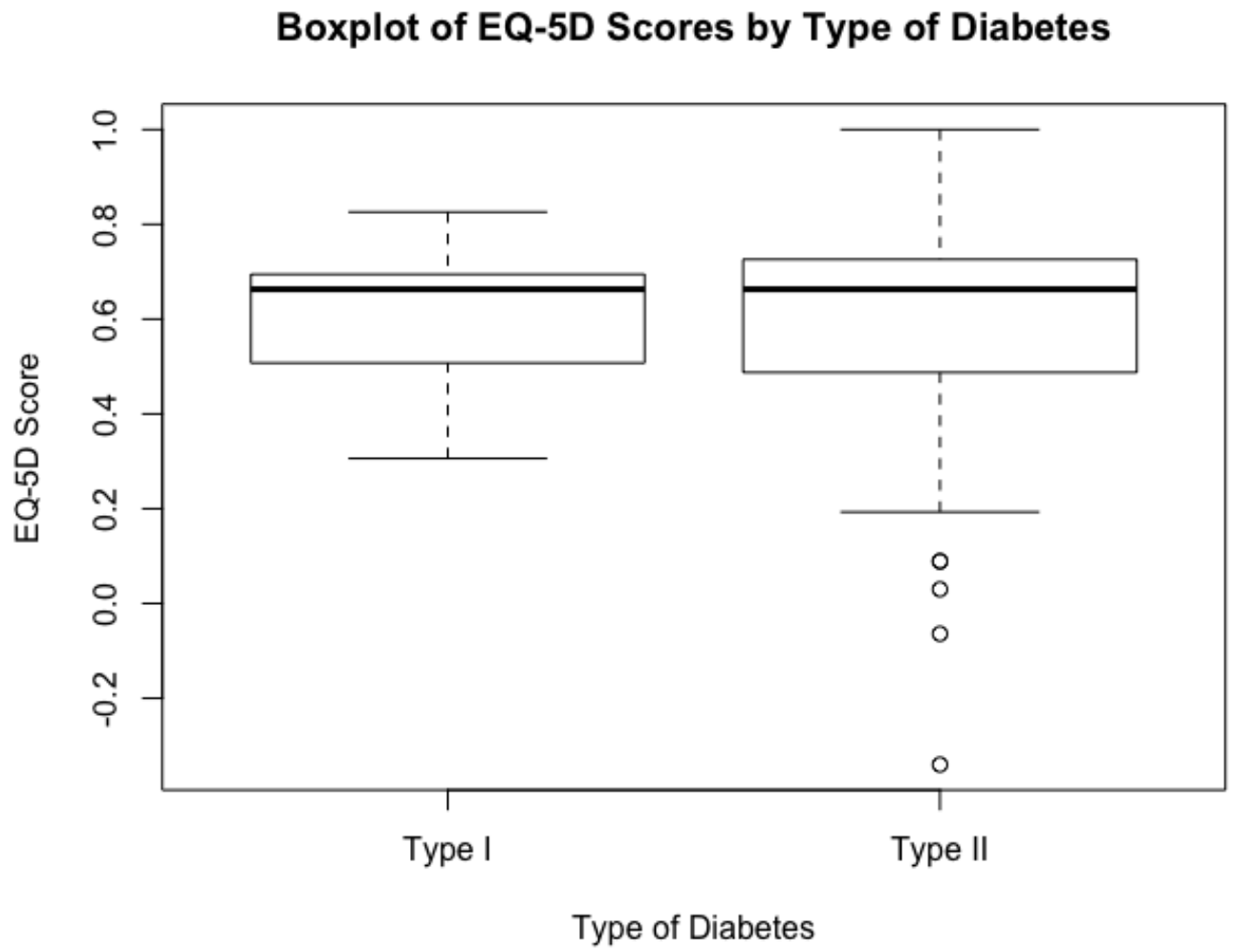

Figure 4. Boxplot of EQ-5D-3L-derived health utility scores by type of diabetes.

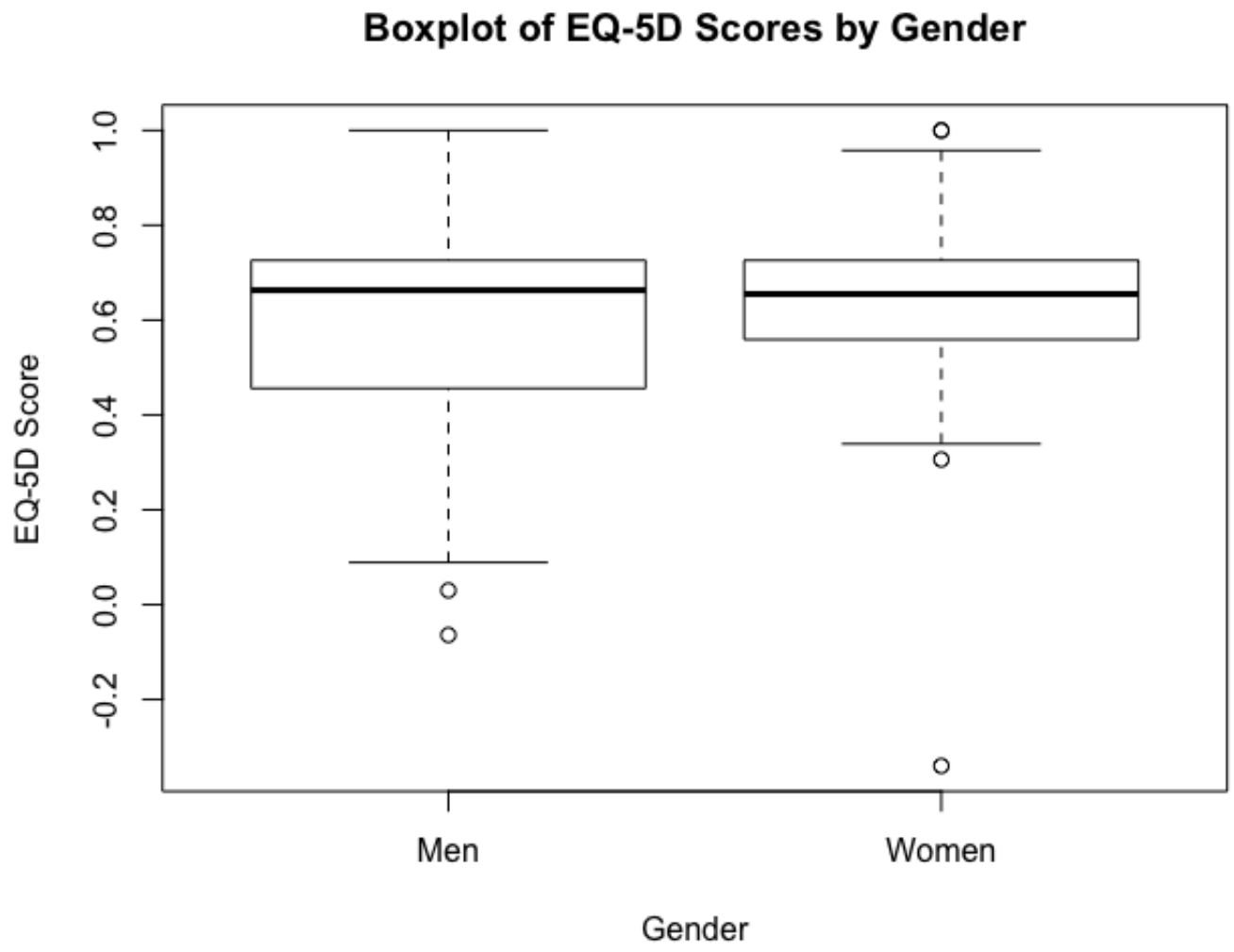

Figure 5. Boxplot of EQ-5D-3L-derived health utility scores by gender.

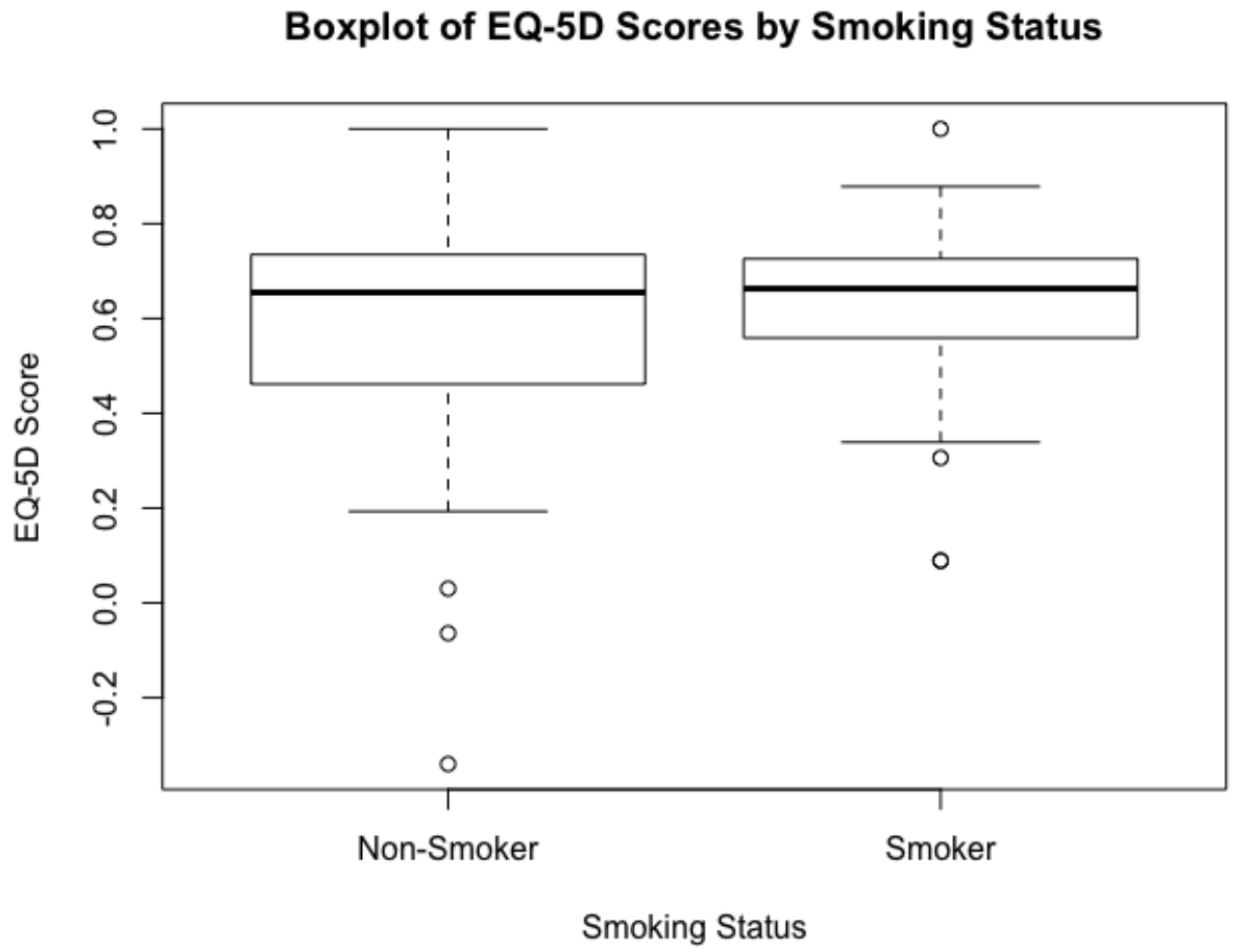

Figure 6. Boxplot of EQ-5D-3L-derived health utility scores by smoking status.

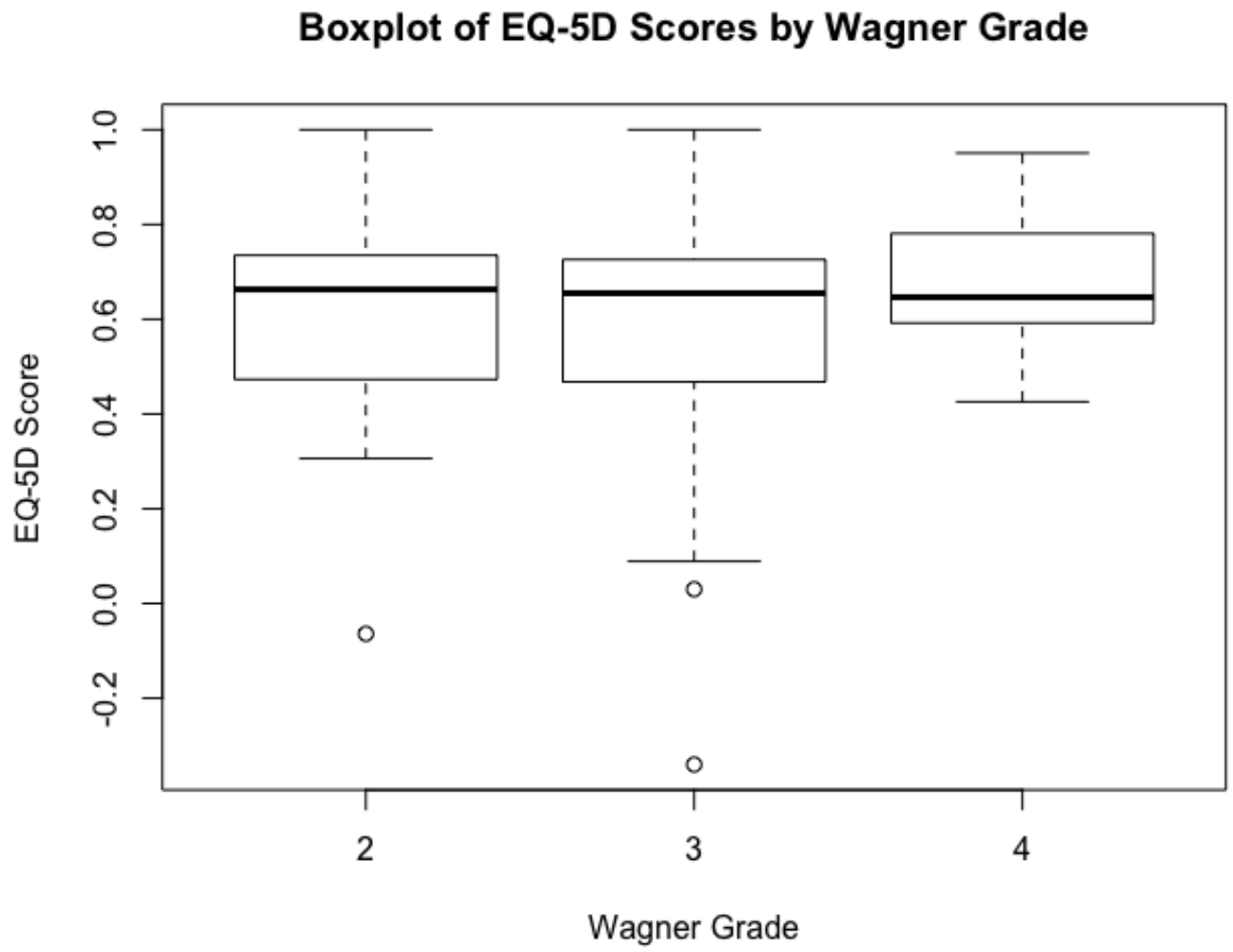

Figure 7. Boxplot of EQ-5D-3L-derived health utility scores by Wagner grade.

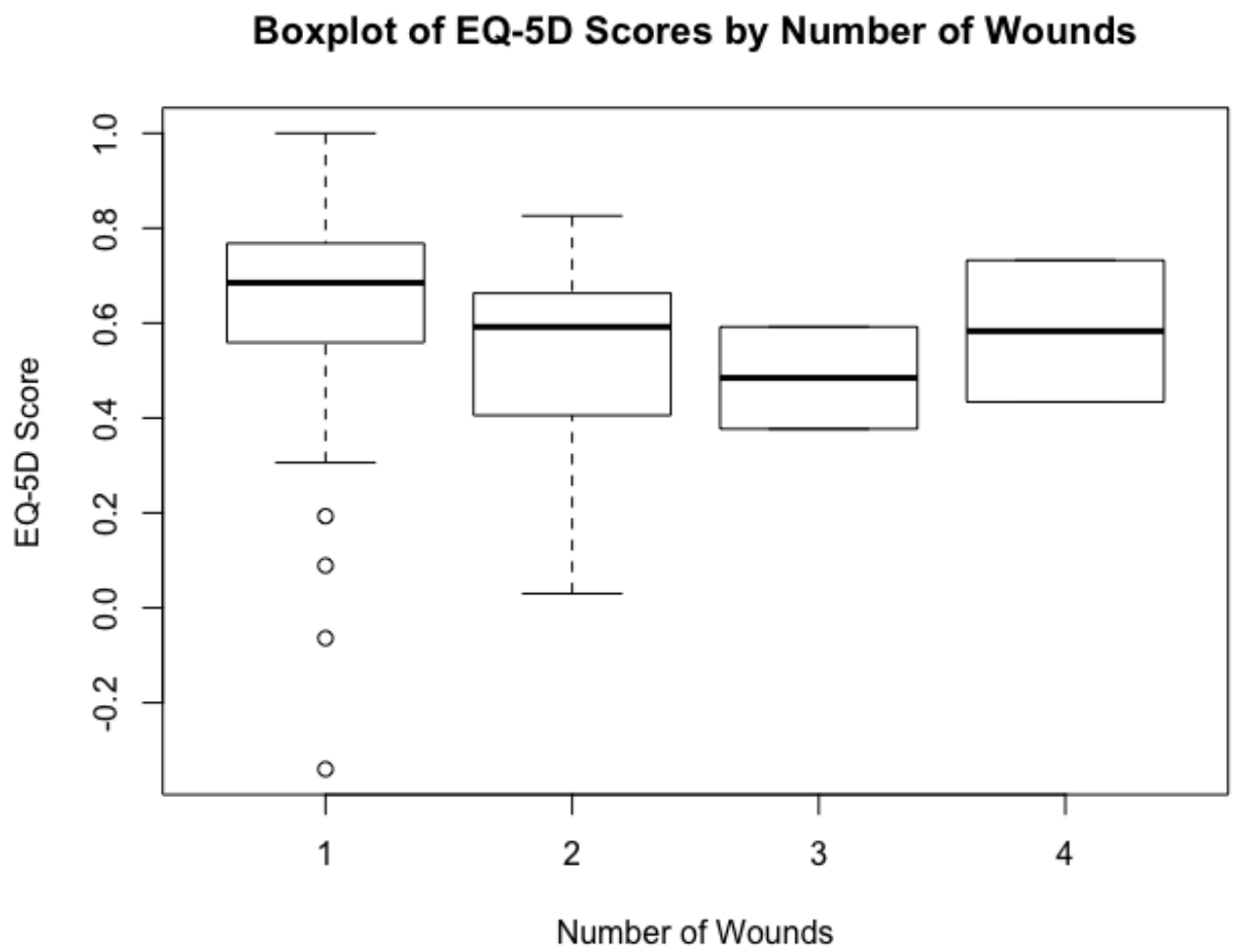

Figure 8. Boxplot of EQ-5D-3L-derived health utility scores by number of wounds.

Table 6. HBOT trial: EQ-5D-3L health utility scores by baseline characteristic.

Baseline Characteristic	Number of Patients	Mean EQ-5D-3L Score	SD EQ-5D-3L Score	Statistical Comparison
Gender				
Male	69	0.6060	0.2233	
Female	34	0.6154	0.2348	p=0.8476 (t-test)
Diabetes Type				
Type I	7	0.5999	0.1750	
Type II	96	0.6098	0.2300	p=0.891 (t-test)
Smoking Status				
Smoker	54	0.6252	0.1794	
Non-smoker	49	0.5914	0.2691	p=0.4601 (t-test)
Wagner Grade				
2	46	0.6430	0.2063	
3	51	0.5710	0.2444	
4	6	0.6738	0.1780	p=0.228 (ANOVA)
Number of Wounds				
1	70	0.6377	0.2369	
2	29	0.5505	0.1968	
3	2	0.4845	0.1520	
4	2	0.5831	0.2111	p=0.298 (ANOVA)

Table 7. HBOT Trial: Multiple Regression and Univariate Regression Results.

Independent Variable	**Univariate Regression Equation Terms**	**Multivariate Regression Equation Terms**
	Coefficient Estimates (95% CI)	**Coefficient Estimates (95% CI)**
Intercept	NA	0.662 (0.174 to 1.15) Overall model p<0.652†
Age (years)	Intercept=0.409 (0.178 to 0.641) Age=+0.00326 (-0.000446 to 0.00696) Overall model p<0.0841†	+0.00259 (-0.00184 to 0.00701)
BMI (kg/m^2)	Intercept=0.750 (0.520 to 0.980) BMI=-0.00464 (-0.0121 to 0.00279) Overall model p<0.218†	-0.00320 (-0.0118 to 0.00543)
Diabetes type (Type I or Type II*)	Intercept=0.600 (0.430 to 0.770) Type II*=+0.00994 (-0.166 to 0.186) Overall model p<0.911†	-0.0254 (-0.248 to 0.197)
Duration of Diabetes (years)	Intercept=0.615 (0.538 to 0.691) Years Diabetes=-0.000346 (-0.00432 to 0.00363) Overall model p<0.863†	-0.00142 (-0.00623 to 0.00340)

Independent Variable	**Univariate Regression Equation Terms**	**Multivariate Regression Equation Terms**
	Coefficient Estimates (95% CI)	**Coefficient Estimates (95% CI)**
Duration of Wound (days)	Intercept=0.617 (0.563 to 0.671) Days Wound=-0.0000270 (-0.000135 to 0.0000806) Overall model p<0.620†	-0.0000359 (-0.000152 to 0.0000800)
Gender (Male or Female*)	Intercept=0.606 (0.552 to 0.660) Female*=+0.00934 (-0.0851 to 0.104) Overall model p<0.845†	+0.0178 (-0.0855 to 0.121)
HbA1c (mmol/mol)	Intercept=0.669 (0.479 to 0.859) HbA1c=-+0.00742 (-0.0302 to 0.0154) Overall model p<0.520†	-0.00539 (-0.0298 to 0.019)
Number of wounds (1, 2*, 3*, or 4*)	Intercept=0.638 (0.584 to 0.691) 2 wounds*=-0.0872 (-0.186 to 0.0115) 3 wounds*=-0.153 (-0.474 to 0.167) 4 wounds*=-0.0596 (-0.375 to 0.266) Overall model p<0.298†	2 wounds=-0.0751 (-0.181 to 0.0308) 3 wounds=-0.125 (-0.461 to 0.210) 4 wounds=+0.0451 (-0.295 to 0.385)
Smoking status (Smoker* or Non-Smoker)	Intercept=0.591 (0.527 to 0.656) Smoker*=+0.0338 (-0.0549 to 0.122) Overall model p<0.451†	+0.0458 (-0.0489 to 0.141)

Independent Variable	**Univariate Regression Equation Terms**	**Multivariate Regression Equation Terms**
	Coefficient Estimates (95% CI)	**Coefficient Estimates (95% CI)**
Wagner grade (Grade 2, 3* or 4*)	Intercept=0.643 (0.577 to 0.709) Grade 3*=-0.0720 (-0.163 to 0.0188) Grade 4*=+0.0308 (-0.163 to 0.224) Overall model p<0.228†	Grade 3=-0.0531 (0.151 to 0.0444) Grade 4=+0.0864 (0.119 to 0.292)
Notes: NA=not applicable. †F-test for overall significance of model. ***Indicates a dummy variable**		

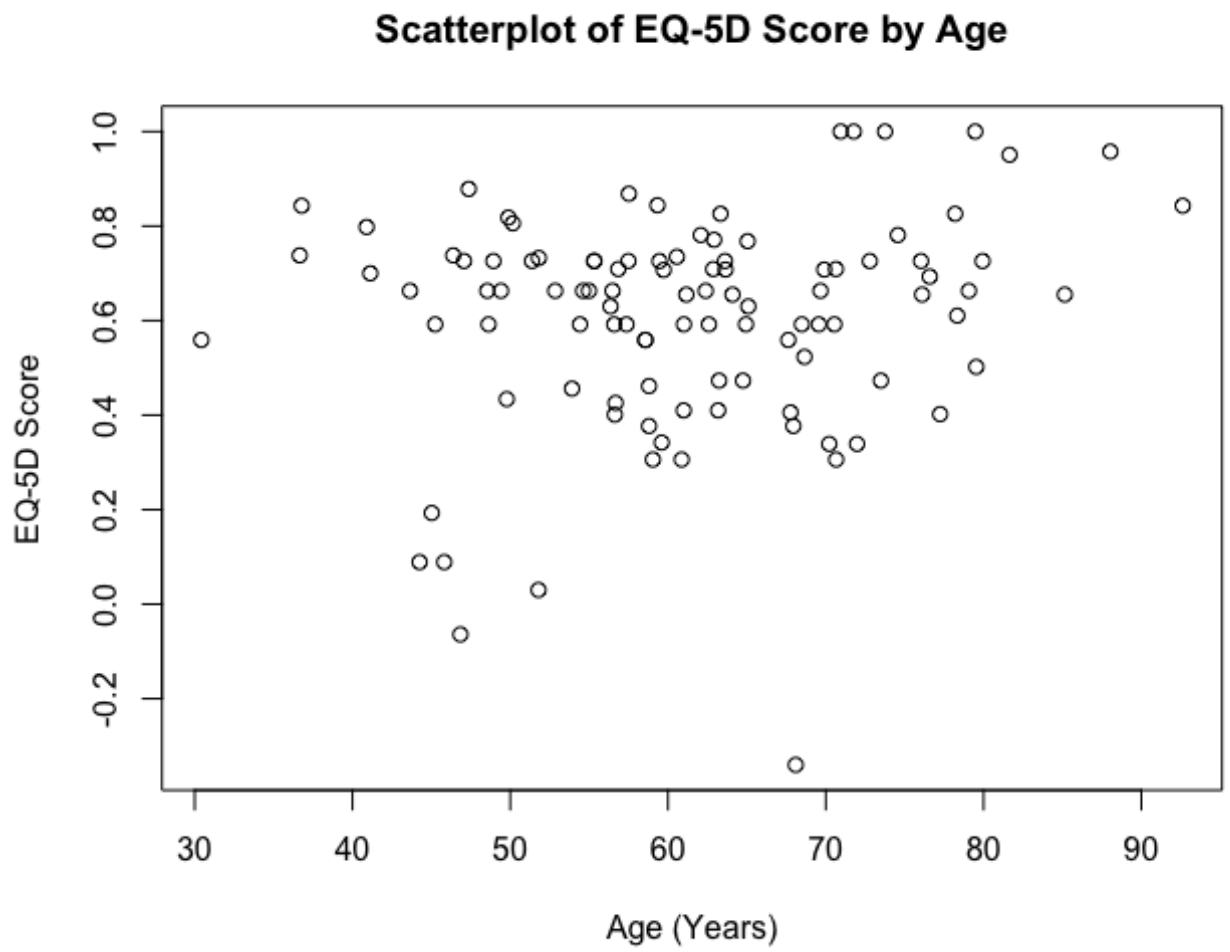

Figure 9. Scatterplot of EQ-5D-3L-derived health utilities by age.

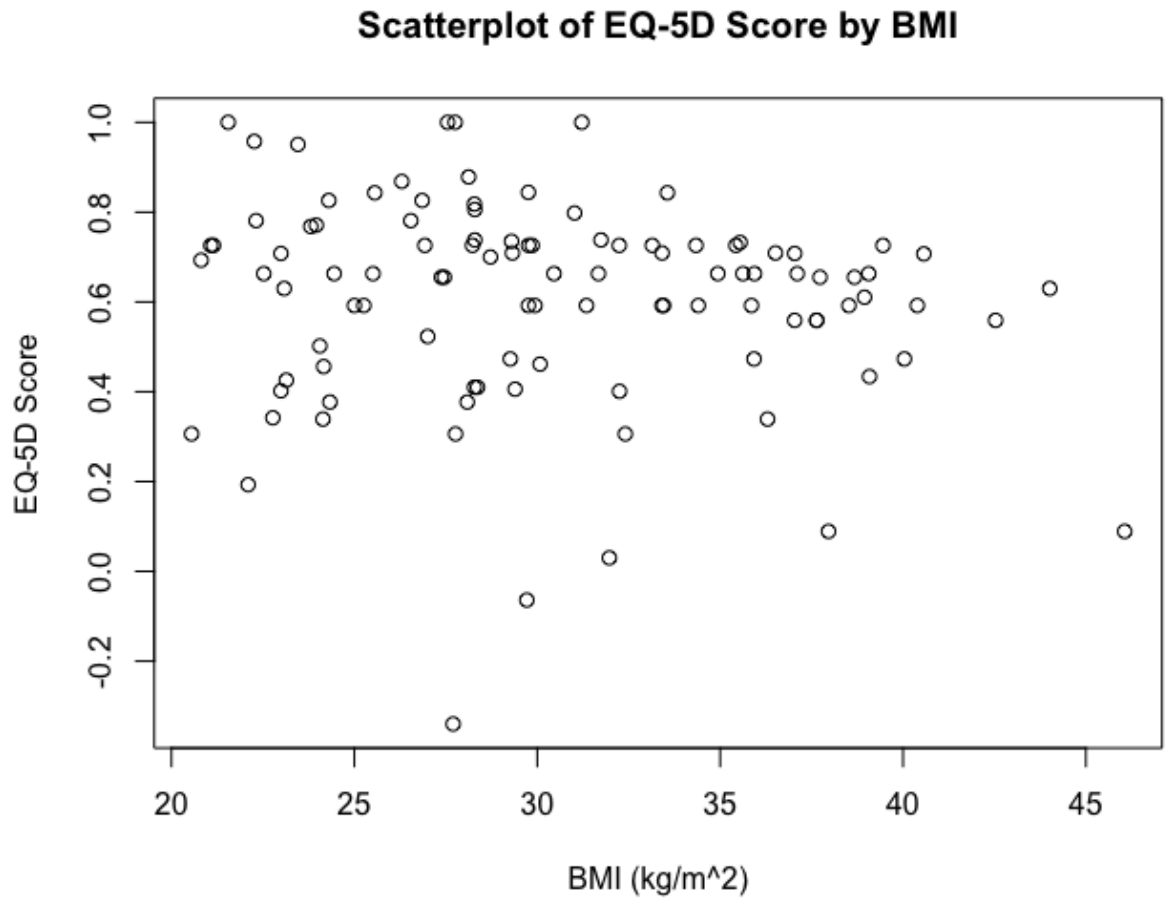

Figure 10. Scatterplot of EQ-5D-3L-derived health utilities by body mass index (BMI).

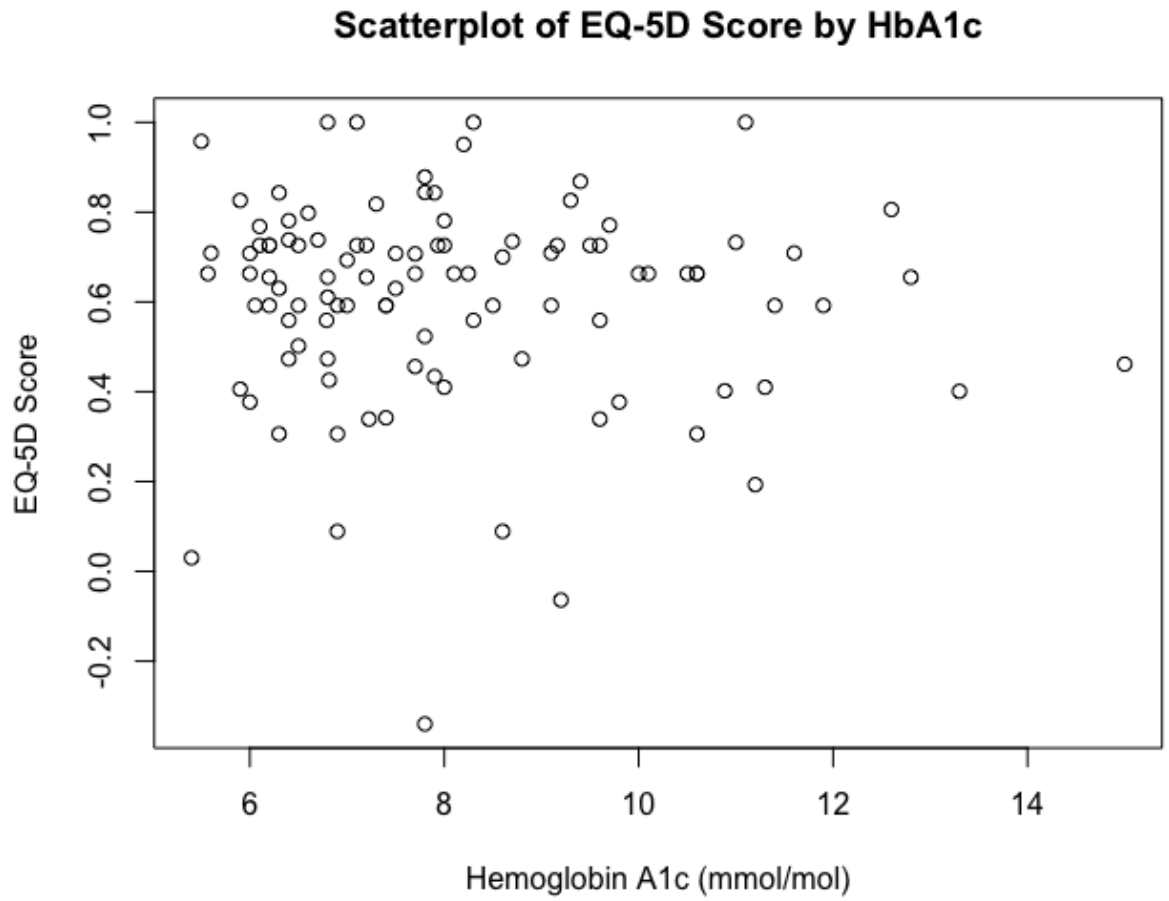

Figure 11. Scatterplot of EQ-5D-3L-derived health utilities by HbA1c.

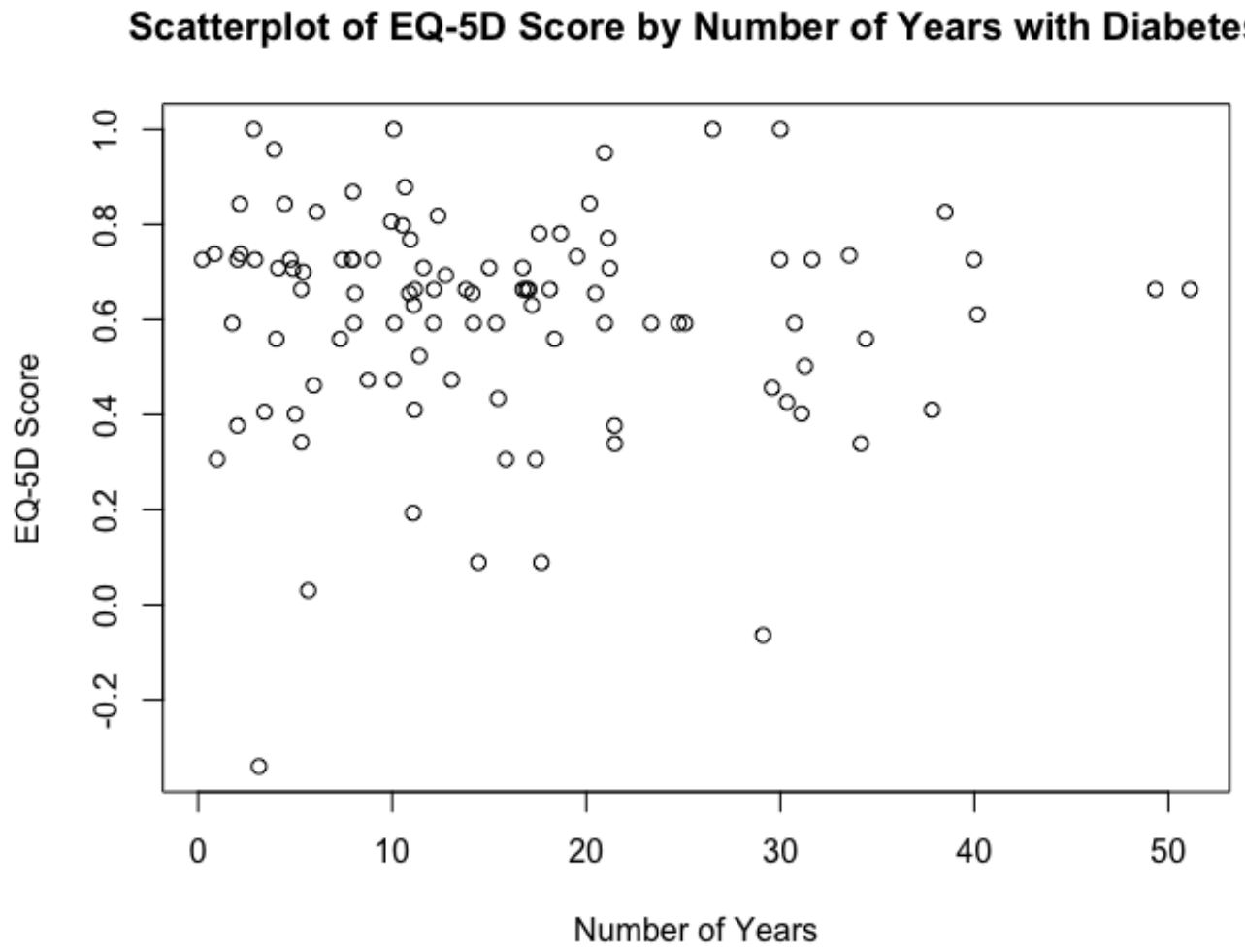

Figure 12. Scatterplot of EQ-5D-3L-derived health utilities by number of years with diabetes.

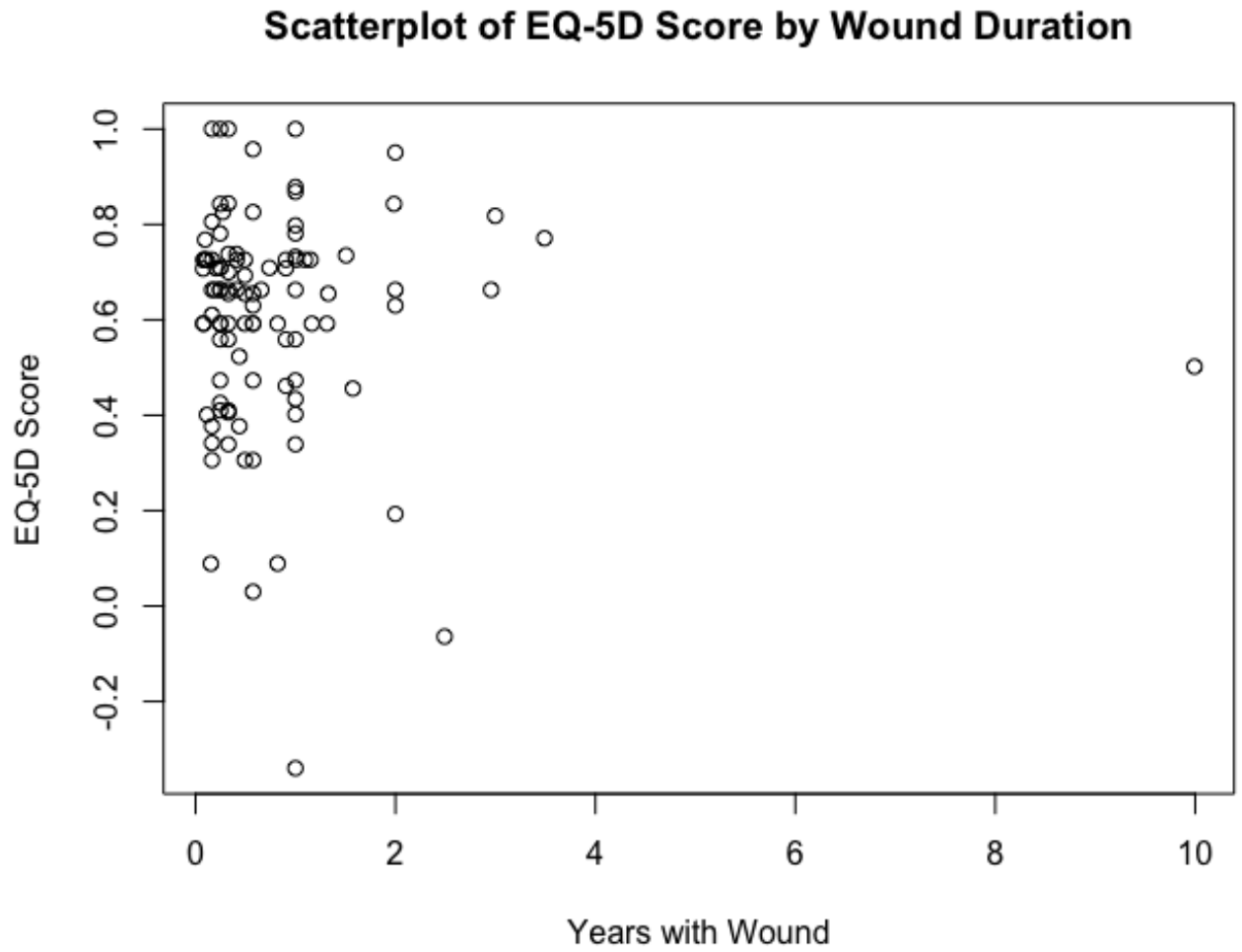

Figure 13. Scatterplot of EQ-5D-3L-derived health utilities by wound duration.

3.4 DISCUSSION

The results of the exploratory multiple regression analysis suggest that the baseline characteristics (gender, age, BMI, type of diabetes, duration of diabetes, duration of wound, HbA1c, number of wounds, smoking status and Wagner grade) may not be associated with EQ-5D-3L-derived health utility scores in patients with chronic non-healing DFUs, as the analysis failed to reject the null hypothesis that, for each of the 10 variables, the variable is associated with health utility. This aligns with the examination of the health utility scores by level of categorical variable, where no statistically significant differences were noted and with the scatterplots of health utility data against age, BMI, HbA1c, duration of diabetes and duration of wound.

Of note, two of the variables used had counter-intuitive results in both single-variable linear regressions and in the multiple linear regression: the number of wounds and Wagner grade. The adjustment to the model for a Wagner grade of 3 lowered the health utility score compared with a Wagner grade of 2, which is reasonable, given that a Wagner grade of 3 represents a more severe ulcer than grade 2, and one would expect a lower health utility. However, the adjustment to the model for a Wagner grade of 4 increased the health utility score (i.e., the adjustment was positive) compared with a Wagner grade of 2. Interestingly, this also occurred in the single-variable linear regression for Wagner grade. Given that a Grade 4 ulcer is more severe than a Grade 2 ulcer, this result appears

counterintuitive. The most likely explanation is the small sample size, as only six patients had grade 4 disease. Small sample sizes can result in an unreliable estimate of the effect that can be higher or lower than the true effect. Therefore, the mean health utility of the six patients with Grade 4 ulcers from this study may not be reflective of the larger population of patients with Grade 4 ulcers. So, while the six patients in this study had a higher mean health utility than those with Grade 2 or 3 ulcers, it is likely that the estimate is not reflective of the population of patients with Grade 4 ulcers. Given a larger sample size of patients with Grade 4 disease, the adjustment to the health utility score in the single and multiple linear regression models may have been negative rather than positive. With respect to the adjustment to the model for the number of wounds, patients with two wounds (n=29) had a lower health utility score, using either the single linear regression or the multiple linear regression, than patients with a single wound (n=70). Similarly, patients with three wounds (n=2) had a lower health utility score than patients with two wounds. However, the single-variable regression model provided a smaller detrimental adjustment to a patient's health utility score if they had four wounds than if the patient had two wounds. Furthermore, the descriptive regression model provided a small positive adjustment (i.e., the coefficient for this term was positive, in the direction of better health utility) in a patient's health utility score if they had four wounds. Of note, the adjustments in both models for a patient with four wounds were small in an absolute sense, and while one (the single-variable linear regression model) was in the direction of

effect that one would expect (i.e., as the number of wounds increases, one's health utility would decrease), the other (for the multiple linear regression) was not. Importantly, only two patients in the trial had four wounds, which means that this suffers from the same problems as the subgroup of patients with Wagner Grade 4 ulcers. It is, however, possible that patients with more wounds would not necessarily have poorer health utility, especially if other factors are associated with health utility. For instance, a patient with three or more ulcers that are all Wagner Grade 2 may have better HRQoL, and therefore a higher health utility, than a patient with only a single ulcer that is Wagner Grade 4. Of note, none of the studies identified in the previous systematic review reported that the number of wounds was negatively correlated with health utility score. A larger sample size of patients with multiple DFUs would be required to make any determination of the impact of the number of wounds on a patient's health utility. It is important to note that the single-variable and multiple linear regression models are exploratory (i.e., they suggest associations, but one cannot infer causality) due to the post-hoc nature of the analyses and the small sample size available from the trial. Furthermore, the health utility estimate adjusted for a Wagner Grade of 4 and the estimate adjusted for four or more ulcers should both be used with caution given the small sample sizes in these subgroups and the exploratory nature of the analyses.

The current study suffers from a small sample size of 103 patients, which means that it may lack sufficient power to find statistically significant differences in health utility scores for potential predictor variables. Coffey et al (19) and Morgan et al (21) both identified that BMI and gender were associated with health utility score, along with DFU status (with one level of that variable being a non-healing, chronic ulcer). Each of those studies included a large number of patients with diabetes (Coffey et al included 2,048 subjects, including 149 with chronic, non-healing DFU, and Morgan et al included 4,502 subjects, including 661 with chronic, non-healing DFU), therefore the power in those studies to detect significant differences in variables and associations between variables would have been much greater than in the current study. Coffey et al (19) estimated that the health utility for patients with a chronic non-healing DFU is 0.504 for Type 1 diabetes and 0.474 for Type 2 diabetes. For a male with Type 1 diabetes without any complications (except having a DFU), the mean health utility was 0.596. Morgan et al (21) estimated that the health utility for patients with a DFU without other complications is 0.512. The authors investigated the effects of age, gender, BMI, coronary heart disease, stroke, DFU, end-stage renal disease, peripheral vascular disease, and retinopathy. The authors reported that for patients with DFU, coronary heart disease, cerebrovascular disease, end-stage renal disease, and retinopathy, the health utility was 0.293 and for patients with DFU and no other complications, the health utility was 0.512. Both Coffey et al (19) and Morgan et al (21) suggested that there are variables that are associated with

health utility in patients with chronic DFUs. An important difference between those studies and the current study is how the health utility data were determined. Similar to the current study, Morgan et al (21) used the EQ-5D-3L; however, they used UK preference weights to obtain health utility scores. Coffey et al (19) used the Quality of Well-Being Index, which uses preference weights obtained from a US population. While the current study estimated an unadjusted health utility of 0.609, Coffey et al and Morgan et al both reported lower unadjusted estimates. Given the differences between how the estimates were obtained from the subjects as well as differences in the preference weights applied to obtain health utility scores between the studies, a comparison of the unadjusted health utility scores is difficult. Interestingly, Morgan et al and Coffey et al both identified different variables that predicted health utility score, with the exception of DFU, gender, and BMI, which were used in both. The most similar estimates of the health utility of patients with DFU between the two studies is 0.512 for patients with DFU not adjusted for other complications from Morgan et al (21), and 0.596 for patients with DFU and not adjusted for other factors from Coffey et al (19). These estimates differ quite meaningfully from each other. Furthermore, the unadjusted estimate from the current study is 0.609, which is higher than in Morgan et al and slightly higher than in Coffey et al.

There are some limitations to the current study. The first is that the patients who were included in this study were enrolled from a single centre in Toronto,

Canada. The centre was a specialized wound care clinic, providing HBOT, that saw patient referrals from physicians in the Toronto area and from other wound care clinics in the surrounding area. It is likely that patients referred to this centre have poorer prognosis, more severe wounds, and generally would be more complicated cases than is seen in regular clinical practice. Therefore, the applicability of the health utility estimates obtained from this trial to patients encountered in regular clinical practice may be limited. A second limitation is that the RCT was not designed *a priori* to estimate health utility for patients with chronic non-healing DFUs. This creates two issues related to the small sample size: 1) that the estimates of health utility suffer from a lack of precision, highlighted by a SD=0.226; and that 2) the ability of the study to identify factors that may predict health utility scores is limited, due to the small sample size. The latter point is important; while the study did not demonstrate that any variables in this sample predicted health utility score, that does not mean that in the highly selected population from which the sample was drawn, these variables do *not* predict health utility score. Without an adequately powered study, it is difficult to draw conclusions from these data. Furthermore, as the study sample was drawn from a highly selected population (i.e., a highly specialized clinic), there is a high risk that the results suffer from selection bias. Lastly, the study suffers from the problem of multiple comparison testing and the fact that the analyses were not planned at the outset of the trial. With each comparison, the risk of a Type 1 error (i.e., the risk of a false-positive result) increases. As no statistically

significant predictors of health utility were identified, this is not a concern, but if the analyses had identified a potential predictor(s) of health utility, the result would need to be interpreted with caution.

The current study did not identify any statistically significant associations for health utility in patients with chronic non-healing DFU. However, due to the small sample size, the results do not rule out that some baseline factors may be associated with health utility in these patients. Given results in prior studies, further research is required to determine what factors predict for health utility in patients with chronic non-healing DFU. Such a study should be adequately designed and powered to identify such variables in order to provide a model that accurately predicts health utility score.

CHAPTER 4: DISCUSSION

Variation in the estimates of health utility for patients exists with chronic non-healing DFUs in the current literature (18-27). The variation in currently available estimates makes it difficult to determine the best estimate for use in a cost-utility analysis investigating treatments for Canadian patients with non-healing DFUs.

It is important to note that there exists little guidance on the appropriate methods to identify the health utility of a population in a given health state, and the author is unaware of any tools that have been specifically developed to critically appraise such studies. This may explain some of the variation in the reported estimates, and made it challenging to ascertain which studies provided the best estimate of health utility for patients with DFUs. Given the limitations identified in each of the studies and the lack of reporting of important study and patient characteristics, the quality of the studies and the estimates that they provided was difficult to ascertain.

Only one fully published study, reported by Li et al (26) and further examined in the Chapter 3 of the current study, provided health utility estimates for Canadian patients (mean=0.609, SD=0.226). The variance in this estimate is quite large, which would result in a substantial amount of parameter uncertainty in a cost-utility analysis. Of note, the health utility score of the HBOT trial, estimated from the baseline demographic variables, may provide a better estimate for health

utility if the baseline characteristics of the population under consideration in the cost-utility analysis is well-defined. Caution should be exercised in the use of these estimates due to the exploratory nature of the analysis, and even more so for the estimates adjusted for a Wagner Grade of 4 and for four or more wounds due to the very small sample sizes for both subgroups.

However, even with the uncertainty in the estimates due to these issues, these estimates are likely the best available for a Canadian population. Given that the effectiveness component of a cost-utility analysis is driven by the time in each health state AND the health utility, the variance in the health utility could result in a large degree of variation in ICUR estimates in a probabilistic analysis; therefore, using an estimate from a study population that more closely matches the population in the decision problem may help to reduce the uncertainty. A substantial amount of uncertainty can be a barrier to decision makers as it leads to difficulty in interpreting the cost-effectiveness acceptability curve and plots of the ICUR estimates on the cost-effectiveness plane.

While other estimates of health utility for patients with chronic non-healing DFUs have been reported, none are from Canadian patients. Some studies have reported similar utility values as those obtained in Chapter 3 of the current study; however, those same studies have also reported similar, or greater, variance for the estimates. Furthermore, not all studies reported health utility estimates solely

for patients with chronic non-healing DFUs. While patients in these studies had chronic non-healing DFUs, it was unclear if some may have had a healed ulcer or even an amputation due to the length of time from identification for inclusion in the study to the time that the subject completed the health utility instrument.

CHAPTER 5: CONCLUSIONS

In order to ensure that decision makers in Canada can make better-informed decisions around funding of treatments for non-healing DFUs, additional research is needed to provide a more accurate and precise estimate of the health utility of this group of patients. Such a study should be designed to investigate what factors (i.e., comorbidities, patient or demographic characteristics) predict for health utility score, as it would help to identify groups of patients in whom the health utility varies and, therefore, the subsequent impact on the ICUR and the value of treatment.

Currently, there is a lack of guidance on using health utility estimates in economic evaluations. More specifically, no tools exist to guide the critical appraisal of studies measuring the health utility of patients with a specific health state for use in economic evaluations and there is limited guidance on the appropriate methods of conducting such studies. In order to obtain accurate and precise estimates of health utilities for use in economic evaluations, guidance on the best study methodologies is needed, as well as critical appraisal tools to aid analysts in choosing the best health utility estimates.

REFERENCES

1. Public Health Agency of Canada. [Internet] Ottawa (Canada): Public Health Agency of Cancer, Government of Canada. Diabetes in Canada: facts and figures from a public health perspective. [updated 2011 Dec 15; cited 2018 Feb 3]. Available from: https://www.canada.ca/en/public-health/services/chronic-diseases/reports-publications/diabetes/diabetes-canada-facts-figures-a-public-health-perspective.html.
2. Health Analytics Branch, Ontario Ministry of Health. Key performance measures for the Ontario Diabetes Strategy final report. June 2013. Ontario Ministry of Health and Long-Term Care, Health System Information Management and Investment Division, Health Analytics Branch; Toronto, Canada: 2013 [cited 2018 Feb 5]. Available from: http://www.champlaindrcc.ca/downloads/toolbox_providers/Key Performance Measures_Ontario Diabetes Strategy_Final Report 2013.pdf.
3. Registered Nurses' Association of Ontario. Assessment and management of foot ulcers for people with diabetes [Internet]. 2nd ed. Toronto, Canada: Registered Nurses' Association of Ontario; 2013 Mar [cited 2018 Feb 4]. 156p. Avaiable from: http://rnao.ca/sites/rnao-ca/files/Assessment_and_Management_of_Foot_Ulcers_for_People_with_Diabetes_Second_Edition1.pdf.
4. Boulton AJM. The diabetic foot: a global view. Diabetes/Metabolism Research and Reviews. 2000;16(Suppl 1):S2-S5.

5. Reiber GE, Vileikyte L, Boyko EJ, Del Aguila M, Smith DG, Lavery LA, et al. Causal pathways for incident lower extremity ulcers in patients with diabetes from two settings. Diabetes Care. 1999;22(1):157-62.
6. Gordois A, Scuffham P, Shearer A, Oglesby A, Tobian JA. The health care costs of diabetic peripheral neuropathy in the U.S. Diabetes Care. 2003;26(6):1790-5.
7. Mancini L, Ruotolo V. The diabetic foot: epidemiology. Rays. 1997;22(4):511-23.
8. Higgins, JPT, Altman DG. Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated 2011 Mar; cited 2018 Feb 1]. The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org/.
9. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes 3rd ed. New York (US): Oxford University Press; c2005. Chapter 6, Cost-utility analysis; p. 137-209.
10. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of Internal Medicine. 2009;151(4):264-9.
11. Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960;20(1):37-46.

12. NICE: National Institute for Health and Care Excellence. [Internet] London (UK): NICE. Developing NICE guidelines: the manual. Process and methods guides (PMG20). [updated 22 Jul 2015; cited 02 Mar 2018]. Available from: https://www.nice.org.uk/process/pmg20/resources/developing-nice-guidelines-the-manual-appendix-h-pdf-2549711485.

13. Higgins, JPT, Altman DG. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated 2011 Mar; cited 2018 Feb 1]. The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org/.

14. Scottish Intercollegiate Guidelines Network (SIGN). [Internet] Edinburgh (Scotland): Healthcare Improvement Scotland. SIGN: Critical appraisal notes and checklists. [cited 2018 Jan 05]. Available from: http://www.sign.ac.uk/checklists-and-notes.html.

15. Sanderson S, Tatt ID, Higgins JPT. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. International Journal of Epidemiology. 2007;36:666-76.

16. Zeng X, Zhang Y, Kwong JSW, Zhang C, Li S, Sun F, et al. The methodological quality assessment tools for preclinical and clinical studies,

systematic reviews and meta-analysis, and clinical practice guideline: a systematic review. Journal of Evidence-Based Medicine. 2015;8(1):2-10.

17. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open. 2016;6:e011458. doi:10.1136/bmjopen-2016-.
18. Tennvall GR, Apelqvist J. Health-related quality of life in patients with diabetes mellitus and foot ulcers. Journal of Diabetes & its Complications. 2000 Sep-Oct;14(5):235-41.
19. Coffey JT, Brandle M, Zhou H, Marriott D, Burke R, Tabaei BP, et al. Valuing health-related quality of life in diabetes. Diabetes Care. 2002 Dec;25(12):2238-43.
20. Redekop WK, Stolk EA, Kok E, Lovas K, Kalo Z, Busschbach JJ. Diabetic foot ulcers and amputations: estimates of health utility for use in cost-effectiveness analyses of new treatments. Diabetes & Metabolism. 2004 Dec;30(6):549-56.
21. Morgan CL, McEwan P, Morrissey M, Peters JR, Poole C, Currie CJ. Characterization and comparison of health-related utility in people with diabetes with various single and multiple vascular complications. Diabetic Medicine. 2006 October;23(10):1100-5.
22. Javanbakht M, Abolhasani F, Mashayekhi A, Baradaran HR, Jahangiri noudeh Y. Health Related Quality of Life in Patients with Type 2 Diabetes

Mellitus in Iran: A National Survey. PLoS ONE. 2012 30 Aug;7 (8) (no pagination)(e44526).

23. Siersma V, Thorsen H, Holstein PE, Kars M, Apelqvist J, Jude EB, et al. Importance of factors determining the low health-related quality of life in people presenting with a diabetic foot ulcer: the Eurodiale study. Diabetic Medicine. 2013 Nov;30(11):1382-7.
24. Siersma V, Thorsen H, Holstein PE, Kars M, Apelqvist J, Jude EB, et al. Diabetic complications do not hamper improvement of health-related quality of life over the course of treatment of diabetic foot ulcers - the Eurodiale study. Journal of Diabetes & its Complications. 2017 Jul;31(7):1145-51.
25. Sobol E, Macioch T, Krakowiecki A, Mrozikiewicz-Rakowska B, Mienicka A, Mlynarczuk M, et al. Health-related quality of life and utility in patients with diabetic foot syndrome - An EQ-5D survey in Poland. Value in Health. 2013 November;16 (7):A445-A6.
26. Li G, Hopkins RB, Levine MAH, Jin X, Bowen JM, Thabane L, et al. Relationship between hyperbaric oxygen therapy and quality of life in participants with chronic diabetic foot ulcers: data from a randomized controlled trial. Acta Diabetologica. 2017 Jun 12:12.
27. Sothornwit J, Suwannakin A, Srisawasdi G, Sriwijitkamol A. Decreased health-related quality of life in diabetes patients with foot problems. Diabetes. 2017 June;66:A28.

28. Prompers L, Huijberts M, Apelqvist J, Jude E, Piaggesi A, Bakker K, et al. Optimal organization of health care in diabetic foot disease: introduction to the Eurodiale study. International Journal of Lower Extremity Wounds. 2007 Mar;6(1):11-7.

29. Fedorko L, Bowen JM, Jones W, Oreopoulos G, Goeree R, Hopkins RB, et al. Hyperbaric oxygen therapy does not reduce indications for amputation in patients with diabetes with nonhealing uclers of the lower limb: a prospective double-blind, randomized controlled clinical trial. Diabetes Care. 2016;39:392-9.

30. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey. York (UK): The University of York, Centre for Health Economics;1995. Discussion Paper: 138 Available from: https://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE Discussion Paper 138.pdf.

31. Seiber WJ, Groessl EJ, David KM, Ganiats TG, Kaplan RM. Quality of Well Being Self-Administered (QWB-SA) Scale: User's Manual. San Diego (US): University of California, San Diego, Health Services Research Center;2008 Available from: https://hoap.ucsd.edu/qwb-info/QWB-Manual.pdf.

32. Dolan P. Modeling valuations for EuroQol health states. Medical Care. 1997;35(11):1095-108.

33. EQ-5D value sets: inventory, comparative review and user guide. Szende A, Oppe M, Devlin N, editors: Springer Netherlands; 2007.

34. Golicki D, Jakubczyk M, Niewada M, Wrona W, Busschbach JJV. Valuation of EQ-5D health states in Poland: first TTO-based social value set in Central and Eastern Europe. Value in Health. 2010;13(2):289-97.

35. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian valuation of EQ-5D health states: preliminary value set and considerations for future valuation studies. PLOS one. 2012;7(2):e31115. doi:10.1371/journal.pone.0031115.

36. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Ottawa Hospital Research Institute. [Internet] Ottawa (Canada): The Ottawa Hospital Research Institute. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [updated 2018; cited 2018 Sep 7]. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

37. Coste J, Quinquis L, Audureau E, Pouchot J. Non response, incomplete and inconsistent responses to self-administered health-related quality of life measures in the general population: patterns, determinants and impact on the validity of estimates--a population-based study in France using the MOS SF-36. Health and Quality of Life Outcomes. 2013;11:44.

38. EuroQol. [Internet] Netherlands: EuroQol. EQ-5D-3L: Valuation. [updated 2017 Apr 18; cited 2018 Feb 5]. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/valuation/.

39. Abufalgha F, Badi H, Gaddafi A, Shilabi H. The aetiological factors, management and outcome of diabetic foot in Misurata central hospital. Jamahiriya Medical Journal. 2009 Autumn;9(3):191-5.

40. Nabuurs-Franssen MH, Huijberts MSP, Nieuwenhuijzen Kruseman AC, Willems J, Schaper NC. Health-related quality of life of diabetic foot ulcer patients and their caregivers. Diabetologia. 2005;48:1906-10.

41. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Services Research. 1972;7(2):118-33.

42. EQ-5D. [Internet] Rotterdam (NED): EUROQOL. c2017. EQ-5D-5L About. [cited 2018 Feb 3]. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/.

43. van Reenen M, Oppe M. EQ-5D-3L: User Guide Version 5.1. [Monograph on the Internet]. Netherlands: EuroQol2015 2015 Apr Contract No.: Feb 5. Available from: https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-3L_UserGuide_2015.pdf. [Accessed: 2018 Feb 5].

44. O'Reilly D, Linden R, Fedorko L, Tarride JE, Jones WG, Bowen JM, et al. A prospective, double-blind, randomized, controlled clinical trial comparing standard wound care with adjunctive hyperbaric oxygen therapy (HBOT) to standard wound care only for the treatment of chronic, non-healing ulcers of the lower limb in patients with diabetes mellitus: a study protocol. Trials [Electronic Resource]. 2011 Mar 07;12:69.

45. Pullenayegum EM, Tarride J-E, Xie F, Goeree R, Gerstein HC, O'Reilly D. Analysis of health utility data when some subjects attain the upper bound of 1: Are Tobit and CLAD models appropriate? Value in Health. 2010;13(4):487-94.

46. R Core Team. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2017. Available at: https://www.r-project.org/.

47. Veneables WN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York: Springer; 2002. Available at: http://www.stats.ox.ac.uk/pub/MASS4.

48. Wagner FWJ. The dysvascular foot: a system for diagnosis and treatment. Foot Ankle. 1981;2(2):64-122.

APPENDICES

Appendix 1. Literature Search Strategy

Medline (OVID)

1. Exp leg ulcer/
2. ((foot or feet or leg or legs or plantar or lower limb* or lower extremit* or toe or toes or ankle or ankles or heel or heels or varicose) adj3 (ulcer* or wound*)).ti,ab.
3. exp *lower extremity/ or (feet or foot or leg or legs or plantar or lower limb* or lower extremit* or toe or toes or ankle or ankles or heel or heels or varicose).ti,kf.
4. skin ulcer/ or Buruli ulcer/ or (ulcer* or wound*).ti,ab,kf.
5. 1 or 2 or (3 and 4)
6. exp diabetes mellitus/ or diabet*.ti,kf,hw.
7. exp diabetic foot/
8. (diabet* adj2 (foot or feet)).ti,ab. or diabet* foot.kf. or diabet* feet.kf.
9. exp skin ulcer/ or (ulcer* or wound*).ti,ab,kf.
10. (7 or 8) and 9
11. (5 and 6) or 10
12. limit 11 to english language
13. "Value of Life"/
14. quality of life/

15. quality of life.ti,kf.
16. ((instrument or instruments) adj3 quality of life).ab.
17. quality adjusted life.ti,ab,kf.
18. (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kf.
19. disability adjusted life.ti,ab,kf.
20. daly*.ti,ab,kf.
21. (sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab,kf.
22. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6).ti,ab,kf.
23. (sf8 or sf 8 or sf eight or shortform 8 or shortform 8 or short form 8 or short form8 or shortform8 or shortform eight or short form eight).ti,ab,kf.
24. (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,kf.
25. (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab,kf.
26. (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,kf.
27. (hql or hqol or h qol or hrqol or hr qol).ti,ab,kf.
28. (health* adj2 year* adj2 equivalent*).ti,ab,kf.
29. (pqol or qls).ti,ab,kf.

30. (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb).ti,ab,kf.
31. nottingham health profile*.ti,ab,kf.
32. sickness impact profile*.ti,ab,kf.
33. exp health status indicators/
34. (health adj3 (utilit* or status)).ti,ab,kf.
35. (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)).ti,ab,kf.
36. (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)).ti,ab,kf.
37. disutilit*.ti,ab,kf.
38. rosser.ti,ab,kf.
39. willingness to pay.ti,ab,kf.
40. standard gamble*.ti,ab,kf.
41. (time trade off or time tradeoff).ti,ab,kf.
42. tto.ti,ab,kf.
43. (hui or hui1 or hui2 or hui3).ti,ab,kf.
44. (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual).ti,ab,kf.
45. duke health profile.ti,ab,kf.
46. functional status questionnaire.ti,ab,kf.
47. dartmouth coop functional health assessment*.ti,ab,kf.
48. quality-adjusted life years/

49. (hye or hyes).ti,ab,kf.
50. quality of life/ or health status/ or ((quality adj2 life) or (health adj1 utilit*) or (patient adj1 reported adj1 outcome*)).ti,ab,kf.
51. or/13-50
52. 12 and 51

EMBASE (OVID)

1. exp leg ulcer/
2. ((foot or feet or leg or legs or plantar or lower limb* or lower extremit* or toe or toes or ankle or ankles or heel or heels or varicose) adj3 (ulcer* or wound*)).ti,ab.
3. exp *lower extremity/ or (feet or foot or leg or legs or plantar or lower limb* or lower extremit* or toe or toes or ankle or ankles or heel or heels or varicose).ti,kw.
4. skin ulcer/ or Buruli ulcer/ or (ulcer* or wound*).ti,ab,kw.
5. 1 or 2 or (3 and 4)
6. exp diabetes mellitus/ or diabet*.ti,kw,hw.
7. exp diabetic foot/
8. (diabet* adj2 (foot or feet)).ti,ab. or diabet* foot.kw. or diabet* feet.kw.
9. exp skin ulcer/ or (ulcer* or wound*).ti,ab,kw.
10. (7 or 8) and 9
11. (5 and 6) or 10

12. limit 11 to english language
13. socioeconomics/
14. exp Quality of life/
15. quality of life.ti,kw.
16. ((instrument or instruments) adj3 quality of life).ab.
17. quality-adjusted life year/
18. quality adjusted life.ti,ab,kw.
19. (qaly* or qald* or qale* or qtime* or life year or life years).ti,ab,kw.
20. disability adjusted life.ti,ab,kw.
21. daly*.ti,ab,kw.
22. (sf36 or sf 36 or short form 36 or shortform 36 or short form36 or shortform36 or sf thirtysix or sfthirtysix or sfthirty six or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab,kw.
23. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six or shortform6 or short form6).ti,ab,kw.
24. (sf8 or sf 8 or sf eight or shortform 8 or shortform 8 or short form 8 or short form8 or shortform8 or shortform eight or short form eight).ti,ab,kw.
25. (sf12 or sf 12 or short form 12 or shortform 12 or short form12 or shortform12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,kw.
26. (sf16 or sf 16 or short form 16 or shortform 16 or short form16 or shortform16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab,kw.

27. (sf20 or sf 20 or short form 20 or shortform 20 or short form20 or shortform20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,kw.
28. (hql or hqol or h qol or hrqol or hr qol).ti,ab,kw.
29. (hye or hyes).ti,ab,kw.
30. (health* adj2 year* adj2 equivalent*).ti,ab,kw.
31. (pqol or qls).ti,ab,kw.
32. (quality of wellbeing or quality of well being or index of wellbeing or index of well being or qwb).ti,ab,kw.
33. nottingham health profile*.ti,ab,kw.
34. nottingham health profile/
35. sickness impact profile*.ti,ab,kw.
36. sickness impact profile/
37. health status indicator/
38. (health adj3 (utilit* or status)).ti,ab,kw.
39. (utilit* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or weight)).ti,ab,kw.
40. (preference* adj3 (valu* or measur* or health or life or estimat* or elicit* or disease or score* or instrument or instruments)).ti,ab,kw.
41. disutilit*.ti,ab,kw.
42. rosser.ti,ab,kw.
43. willingness to pay.ti,ab,kw.
44. standard gamble*.ti,ab,kw.

45. (time trade off or time tradeoff).ti,ab,kw.
46. tto.ti,ab,kw.
47. (hui or hui1 or hui2 or hui3).ti,ab,kw.
48. (eq or euroqol or euro qol or eq5d or eq 5d or euroqual or euro qual).ti,ab,kw.
49. duke health profile.ti,ab,kw.
50. functional status questionnaire.ti,ab,kw.
51. dartmouth coop functional health assessment*.ti,ab,kw.
52. quality of life/ or health status/ or ((quality adj2 life) or (health adj1 utilit*) or (patient adj1 reported adj1 outcome*)).ti,ab,kw.
53. or/13-52
54. 12 and 53

Appendix 2.

Table 2-1. Critical Appraisal of Tennvall et al, 2000 (18), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. Study aim was to determine health utilities of patients with (1) chronic non-healing DFU's, (2) healed ulcers, and (3) amputations. Also sought to determine the differences between non-healing DFU patients and those with healed ulcers and those with amputations.
Methods 2. Was the study design appropriate for the stated aims?	Yes.
3. Was the sample size justified?	No. No rationale for sample size was reported.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	No. Authors did not clearly state if the target population was all patients with DFU's. Can be assumed that authors intended the target population to be all patients with DFU's.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Do not know. The target population was not defined. Sample was from one academic centre in Sweden; therefore, not representative of all patients, worldwide.
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Do not know. The target population was not defined. All consecutive patients from one academic centre in Sweden were asked to participate. Selecting from all consecutive patients reduces risk of bias; however, selecting from only one academic centre may increase bias as an academic centre is more likely to have patients with more complicated and difficult to

	treat cases, etc. All patients were from Sweden, therefore, not representative of all patients, worldwide. The results from this population may not be generalizable to a Canadian population due to differences in demographics and health care systems between Sweden and Canada.
7. Were measures undertaken to address and categorise non-responders?	Yes. The authors compared the responders to the survey with the non-responders and noted that there were no differences in foot ulcer status or other select characteristics. See note for Q#13 and #14, below.
8. Were the risk factor and outcome variables measured appropriate to the aims of the study?	Outcome variable was HRQoL. Measurement was appropriate; used EQ-5D to measure health utility.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	Yes; EQ-5D (Swedish version) was used to measure the outcome variable.
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	No. The authors did not provide confidence intervals for the EQ-5D score for patients with current ulcers. The range was provided (minimum and maximum scores). Of note, Tennvall et al described methods to use multiple linear regression to identify independent factors related to EQ-5D score. Their model included patients with current ulcers, healed ulcers and amputations, which was beyond the scope of the current study.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Yes.
Results 12. Were the basic data adequately described?	No. There are limited data (baseline characteristics, precision estimates for EQ-5D)

	reported for patients with non-healed ulcers.
13. Does the response rate raise concerns about non-response bias?	Yes. Of 457 surveys sent, 310 were returned; therefore, 30% of surveys were not returned. Given the number of non-responders, the estimates of health utility could be biased, as patients with poorer HRQoL are less likely to complete and return surveys than patients with better HRQoL. Furthermore, for the subgroup of interest for the current study, 56 respondents had a current ulcer, and it is possible that of the 147 non-returned surveys, many of those may have had current ulcers and did not return a survey due to poor HRQoL as a result of their ulcers. The possibility of bias due to non-response cannot be ruled out; however, the extent of the impact is unknown.
14. If appropriate, was information about non-responders described?	No, information about non-responders to the survey were not described.
15. Were the results internally consistent?	Authors used a validated and reliable tool to assess health utility (EQ-5D-3L). No assessment of consistency of results within the study was reported.
16. Were the results for the analyses described in the methods, presented?	Yes.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	The authors conclusions (that HRQoL in patients with current DFUs is lower than those with healed or amputated DFUs) is possible; however, the health utility of patients with current (non-healing) DFUs was based on data from only 56 patients. The results for these patients should be interpreted with caution due to uncertainty from

	the small sample size.
18. Were the limitations of the study discussed?	The authors discussed the limitations of the study (using a single generic instrument; one QoL rating for each patient; few patients in the major amputation group); however, the authors felt that their survey response rate was high, given the age distribution in the study (i.e., that 19% of respondents were 80 years of age or older). Of note, 30% of surveys were not returned, which may have biased the results (see question 13, above).
Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	Funding sources were declared (National Corporation of Swedish Pharmacies and the Swedish Diabetes Association), but none could be considered a clear conflict of interest.
20. Was ethical approval or consent of participants attained?	Yes, ethics approval was attained. No mention of obtaining consent; however, only data from patients who completed and returned a survey were included (returning a survey can be considered implied consent).

Table 2-2. Critical Appraisal of Coffey et al, 2002 (19), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. The aim of the study was to describe the health utilities associated with diabetes and its treatments, complications and comorbidities.
Methods 2. Was the study design appropriate for the stated aims?	Yes.
3. Was the sample size justified?	No. No rationale for the sample size was reported.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	No. Authors did not clearly state if the target population was all patients with DFU's. Can be assumed that authors intended the target population to be all patients with DFU's.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Do not know. The target population was not clearly defined. Sample was from the University of Michigan health system; therefore, not representative of all patients, worldwide. May be reasonable to generalize this population to Canadian patients; however, there are differences in the health care systems between Michigan (being in the US, and a private-payer system) and Canada (being a public-payer system).
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Do not know. The target population was not defined. All consecutive patients from one academic centre in Sweden were asked to participate. Selecting from all consecutive patients reduces risk of bias; however, selecting from only one academic centre may increase bias

	as an academic centre is more likely to have patients with more complicated and difficult to treat cases, etc. All patients were from Sweden, therefore, not representative of all patients, worldwide. The results from this population may not be generalizable to a Canadian population due to differences in demographics and health care systems between Sweden and Canada
7. Were measures undertaken to address and categorise non-responders?	Unknown. The authors did not provide any details regarding the non-responders and did not report if any assessment of non-responders versus responders was conducted.
8. Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes. Outcome variable was health utility. Measurement was appropriate: used the QWB-SA. DFU status and demographic characteristics were also measured appropriately.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	Yes. The outcome variable was measured using QWB-SA.
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	Authors included all variables in the multiple regression. Interaction terms that were significant (i.e., $p \leq 0.05$) were included in the model.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Yes.
Results 12. Were the basic data adequately described?	No. There are limited data reported for patients with non-healed (chronic) DFUs. Patients were categorized as having neuropathic sores or amputations.
13. Does the response rate raise concerns about non-response bias?	Yes. The authors noted that 88% of all subjects approached participated in the study. The

	authors did not report exactly how many patients were approached. However, as 2,048 subjects participated, it is estimated that, based on an 88% response rate, approximately 2,327 individuals were asked to participate. Of those, 279 did not complete the health utility assessment. Of concern is that only 149 of 2,048 respondents had a neuropathic sore (DFU), therefore the data from even a few non-responders could impact the estimates of health utility due to the small sample size.
14. If appropriate, was information about non-responders described?	No. Information about non-responders was not reported.
15. Were the results internally consistent?	The authors used a validated tool to assess health utility. No assessment of consistency of results within the study was reported.
16. Were the results for the analyses described in the methods, presented?	Yes.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	Not all. While the authors did point out 2 limitations of their study (see Q#18, below), the authors concluded that the results of their study can be used to inform economic evaluations of health technologies for patients with type 1 and 2 diabetes. However, the sample size of patients with DFUs was small (n=57 for type 1 diabetes and n=92 for type 2 diabetes), which increases the uncertainty in the results for these subgroups. Of note, the penalty for DFUs in type 1 diabetic patients was -0.076 with a standard error of 0.016, and the penalty for DFUs in type 2 patients was -0.099, with a standard error of 0.013.

18. Were the limitations of the study discussed?	The authors discussed 2 limitations: 1) the selection of the sample from a tertiary clinic that was more likely to have advanced complications—but the authors felt that this not bias the results as the tariffs come from a general population; and, 2) the self-reporting of diabetes health status—but the authors felt it important to collect diabetes staging data from the patient (by questionnaire) rather than by a review of medical records or by objective testing.
Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	Not apparent conflicts that would seem to bias the results. The research was supported by a grant from the Michigan Diabetes Research and Training Center, by the Centers for Disease Control and Prevention, by an unrestricted grant from Aventis Pharmaceuticals, and by a fellowship training award from Eli Lilly.
20. Was ethical approval or consent of participants attained?	Yes. Approval obtained from the University of Michigan and the Centers for Disease Control.

Table 2-3. Critical Appraisal of Redekop, 2004 (20), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	To estimate health utility values for health states in patients with DFUs.
Methods 2. Was the study design appropriate for the stated aims?	Yes.
3. Was the sample size justified?	No justification for the sample size was reported.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	No, as the authors did not clearly state if they were attempting to estimate health utilities for all patients with DFUs or just those in a particular jurisdiction. The study population was sampled from the general public in Rotterdam, Netherlands.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Do not know, as the target population was unclear. The sample was obtained from the general public in Rotterdam, Netherlands.
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	No. The authors reported that they used quota sampling to ensure representativeness in terms of gender and age (for the Dutch population), the non-random nature of the sampling method means that the selection of participants may have been biased.
7. Were measures undertaken to address and categorise non-responders?	All responders were included in the study.
8. Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes. The authors measured health utility using the time trade-off method.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	Yes. The authors used a time trade-off method to measure health utility. They did adapt it though, by interviewing the study participants in groups and not individually. However, the authors

	reported that they used a method that was proven valid in previously reported studies. The authors determined health states by ulcer status (no active ulcer, active uninfected, or active infected) and by amputation status (no previous amputation, 1 or more toes amputated, one foot amputated, one leg amputated, or both feet or legs amputated).
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	Yes. The authors reported that they would calculate mean and standard deviations for utility scores for each health state.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Yes.
Results 12. Were the basic data adequately described?	 Yes.
13. Does the response rate raise concerns about non-response bias?	No.
14. If appropriate, was information about non-responders described?	Not appropriate.
15. Were the results internally consistent?	Yes. The results appear to be consistent within the study.
16. Were the results for the analyses described in the methods, presented?	Yes.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	 Yes. The authors concluded that the impact of a DFU was less in the presence of a previous amputation compared to when there was no previous amputation.
18. Were the limitations of the study discussed?	Yes.
Other	

19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	Funding was provided by an unrestricted grant from Novartis Pharma AG. Did not likely affect the authors' interpretation of the study results.
20. Was ethical approval or consent of participants attained?	Do not know. The authors did not report if ethical approval of consent of participants was attained.

Table 2-4. Critical Appraisal of Morgan et al, 2006 (21), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. The authors objective was to understand and compare health utilities in patients treated in-hospital who have diabetes and single and multiple comorbidities.
Methods 2. Was the study design appropriate for the stated aims?	Yes.
3. Was the sample size justified?	No. No justification for the sample size was reported.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	Yes. The reference population was patients treated in Cardiff and the Vale of Glamorgan in the UK.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Yes. The sample was taken from a hospital trust in Cardiff and the Vale of Glamorgan in the UK.
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	No. The authors used a consecutive sampling strategy, where all patient discharged from hospital between January 2002 and July 2005 were sent a questionnaire.
7. Were measures undertaken to address and categorise non-responders?	Do not know. The authors reported demographic and disease-specific data for the responders and previously-reported data for the hospital trust in order to indicate similarities and differences; however, no further details were reported.
8. Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes. Health utilities were measured using the EQ-5D-3L.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been	Yes. Health utilities were measured using the EQ-5D-3L. The definition of “diabetic foot” used

trialled, piloted or published previously?	in the study was not clear.
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	The authors were interested in determining what complications were associated with reducing health utility in diabetic patients. The authors reported that only those complications with a p≤0.05 would be included in the regression model.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Yes.
Results 12. Were the basic data adequately described?	For the purposes of this study, the baseline data on patients with DFUs were not adequately described.
13. Does the response rate raise concerns about non-response bias?	A total of 150,113 questionnaires were sent out and 50,258 were returned. Of those, 41,234 were from unique patients of which, only 4,052 (9.8%) were from patients with diabetes. The response rate is very low, therefore the potential for non-response bias exists.
14. If appropriate, was information about non-responders described?	No.
15. Were the results internally consistent?	The results appeared to be internally consistent.
16. Were the results for the analyses described in the methods, presented?	Yes.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	The authors conclusions (that diabetes is associated with decreased quality of life, which in turn, is associated with the extent of comorbidity) are justified by the results.
18. Were the limitations of the study discussed?	Yes. The authors noted that there may exist differences between responders and non-

	responders.
Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	No. Funding provided by AstraZeneca (but likely did not affect interpretation).
20. Was ethical approval or consent of participants attained?	Yes (both).

Table 2-5. Critical Appraisal of Javanbakht et al, 2012 (22), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. The objective of the study was to measure HRQoL in Iranian people and to determine what demographic and disease-related characteristics are associated with better HRQoL.
Methods 2. Was the study design appropriate for the stated aims?	Yes.
3. Was the sample size justified?	No justification for that sample size was provided.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	Yes. The reference population was Iranians.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Yes. The sample frame included 30 Iranian provinces (there are currently 31).
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Yes. Sample size was 3,918 patients with Type 2 diabetes mellitus with fasting plasma glucose ≥7.0 mmol/L or with a glucose tolerance test 2 hours after the oral dose a plasma glucose of ≥11.1 mmol/L, and 16 years of age or older. A multi-stage cluster sample was used. 50 clusters in each of 30 provinces. The authors did not distinguish between rural and urban in the clusters; however, they reported that the samples were selected proportional to urban and rural populations.
7. Were measures undertaken to address and categorise non-responders?	Don't know. The authors did not report how many individuals were contacted to participate in the study; only the number of respondents to the questionnaire (and included in the final analysis)

	was reported (n=3472).
8. Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes. HRQoL was appropriately measured using EQ-5D-3L. Demographic and disease-related characteristics were also appropriate.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	HRQoL was measured using the EQ-5D-3L. The demographic and disease-related characteristics were not explicitly defined; however, most have common definitions. Those that do not may be defined differently in various countries (e.g., under employment status, one is defined as "employed" or "housewife + students" or "unemployed"; the definition of housewife is not provided and could be defined differently in Canada compared with Iran)
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	Health utilities were expressed as a mean +/- a standard deviation. The authors described how they determined what factors to include in their regression models, and reporting of precision (95% CIs).
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Yes. The methods for the study and analysis were sufficiently described such that they could be repeated.
Results 12. Were the basic data adequately described?	No. The study data were well described; however, there are no data on non-responders.
13. Does the response rate raise concerns about non-response bias?	No. Out of 3,918 patients asked to complete a questionnaire, a total of 3,472 (88.6%) responded. The response rate was high.
14. If appropriate, was information about non-responders described?	No. Information about non-responders was not described.
15. Were the results internally consistent?	Yes. The results appeared to be internally

	consistent.
16. Were the results for the analyses described in the methods, presented?	Yes. The authors reported results for the analyses that were described in the methods.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	Yes. The authors' conclusions were justified by the results.
18. Were the limitations of the study discussed?	Yes. The authors discussed the limitations of their study. The authors noted that the database from which patients demographic and disease-related characteristics were obtained was not designed specifically for patients with diabetes, therefore, data for some important complications of diabetes were not available. Also, the study was collected HRQoL data at one point in time and the HRQoL of patients with diabetes is likely to fluctuate over time. The authors also noted that the associations that were suggested in their study are not necessarily causal due to the cross-sectional design of the study.
Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	No. The authors reported no sources of funding that would affect their interpretation of the results. The authors reported no conflicts of interest.
20. Was ethical approval or consent of participants attained?	Yes. The authors sought and received ethics approval and consent of participants.

Table 2-6. Critical Appraisal of Sobol et al, 2013 (25), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. The objective was to estimate the impact of the severity of DFU on HRQoL.
Methods 2. Was the study design appropriate for the stated aims?	Don't know. Insufficient information to determine.
3. Was the sample size justified?	Don't know. Insufficient information to determine.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	Don't know. The authors did not explicitly report what the target population was; however, inferring from the objective, the target population appeared to be all patients with active DFU.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Don't know. The sample frame was not reported; however, it can be inferred from the author's affiliations that the sample was derived from patients in Poland.
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Don't know. The selection process was not described.
7. Were measures undertaken to address and categorise non-responders?	Don't know. A total of 185 patients were questioned directly (between April 2012 and May 2013), and 179 (96.8%) of those completed an EQ-5D-3L questionnaire.
8. Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes. HRQoL was the outcome variable and was measured using the EQ-5D-3L (and used a Polish value set). Risk factors included severity of patients' DFUs, ulcer size, years from diagnosis of diabetes, grade of tissue loss, and grade of infection.
9. Were the risk factor and outcome variables measured	The outcome (HRQoL) was measured

correctly using instruments/measurements that had been trialled, piloted or published previously?	appropriately using the EQ-5D-3L. The severity of DFU was appropriately measured using the PEDIS scale. How grade of tissue loss and grade of infection were classified was not reported.
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	Yes. For precision around the HRQoL estimates, the authors reported mean values and the standard deviation. No details were reported on how correlation between HRQoL and risk factors was assessed.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	No.
Results 12. Were the basic data adequately described?	No. The authors reported basic data for the risk factors and the EQ-5D estimate of HRQoL; however, there were no data on the correlation statistics, and other demographic information about the patients was not reported.
13. Does the response rate raise concerns about non-response bias?	No. The response rate was high.
14. If appropriate, was information about non-responders described?	No. The authors did not report information about the non-responders.
15. Were the results internally consistent?	Don't know. Insufficient information to determine.
16. Were the results for the analyses described in the methods, presented?	Don't know. Insufficient information to determine.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	The authors did not report results for the correlations; therefore, their conclusion regarding finding little or no correlation between ulcer severity (PEDIS) and HRQoL (EQ-5D-3L) was not justified (by data).
18. Were the limitations of the study discussed?	No. No limitations were discussed.

Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	Don't know. The authors did not report funding or a statement on conflicts of interest.
20. Was ethical approval or consent of participants attained?	Don't know.

Table 2-7. Critical Appraisal of Siersma et al, 2013 (23) and Siersma et al, 2017 (24), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. The objective of the study was to identify factors responsible for the low HRQoL associated with DFUs and the relative importance of those factors.
Methods 2. Was the study design appropriate for the stated aims?	Yes.
3. Was the sample size justified?	No. No justification was provided.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	No. The target population was not explicitly stated, but it is inferred from the objective to be all patients with new DFUs.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Yes. The sample frame was from a European population of patients that attended clinics. There may be some concerns with generalizing to all patients; however, a European population is similar to a Canadian population (the focus of this study). There may be differences in the continuum of care in a European population compared with patients from other countries.
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Yes. However, the authors did not report the types of centres where the clinics were based. If the clinics were highly specialized or part of academic centres, the sample may be over representative of severe cases.
7. Were measures undertaken to address and categorise non-responders?	Don't know. The authors did not report information on non-responders, and they did not state how many patients were asked to participate.

8. Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes. HRQoL was measured appropriately using EQ-5D-3L. Risk factors (patient characteristics; disease characteristics; foot and ulcer related characteristics; co-morbidities) were appropriate.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	Yes. The outcome measure was by EQ-5D-3L, which is an appropriate and standardized measure. The risk factor variables are appropriate and the measures are standard.
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	Yes. The authors reported mean and standard deviations for estimates of EQ-5D. Authors also reported the mean increase in the coefficient of determination for each individual factor into a model comprising all factors.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Yes.
Results 12. Were the basic data adequately described?	Yes.
13. Does the response rate raise concerns about non-response bias?	Don't know. No data reported on non-responders.
14. If appropriate, was information about non-responders described?	No.
15. Were the results internally consistent?	Yes. The results appear to be internally consistent. Of note, no data on non-responders.
16. Were the results for the analyses described in the methods, presented?	Yes.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	Yes. The conclusions are aligned with the study's results.
18. Were the limitations of the study discussed?	No. The authors did not discuss the limitations of

	the the study and the impact on the results.
Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	No. The funding sources were reported (none would be considered competing interests) and the authors declared they had no conflicts of interest.
20. Was ethical approval or consent of participants attained?	Yes.

Table 2-8A. Critical Appraisal of Li et al, 2017 (26), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. The objective was to investigate the effect of hyperbaric oxygen therapy (HBOT) on HRQoL in patients with DFUs.
Methods 2. Was the study design appropriate for the stated aims?	Yes. Randomized trial to compare the HRQoL of HBOT with standard therapy.
3. Was the sample size justified?	No. A sample size justification was not provided in the publication. While it was provided in another publication for the trial (29), the sample size was not determined based on HRQoL as an outcome. Therefore, no sample size justification was provided with regard to HRQoL.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	No. The target population was not clearly defined; however, it can be inferred from the conclusions that the authors' target population was all patients with DFUs.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Yes. The sample was taken from patients in southern Ontario, Canada. There may be some concerns with generalizing to all patients; however, the sample is from a subset of the Canadian population (the focus of this study).
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Yes. Consecutive patients from an academic/specialized diabetic foot clinic in Toronto, Canada.
7. Were measures undertaken to address and categorise non-responders?	Yes. All patients in the trial were asked to complete the EQ-5D questionnaire.
8. Were the risk factor and outcome variables measured	Yes. The outcome measure was appropriate for

appropriate to the aims of the study?	the study (EQ-5D). The study did not assess the association of risk factors with HRQoL; however, the outcome estimates were adjusted for baseline index value, age, sex, BMI, HbA1c, type of diabetes, diabetes and foot ulcer duration, and Wagner grade.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	Yes. The outcome measure was appropriately measured (EQ-5D-3L).
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	Yes. The EQ-5D estimates for HRQoL were reported using means and standard deviations. The authors planned to compare the results between treatment groups and over time; however, that is beyond the scope of the present study.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Yes.
Results 12. Were the basic data adequately described?	Yes.
13. Does the response rate raise concerns about non-response bias?	No.
14. If appropriate, was information about non-responders described?	N/A.
15. Were the results internally consistent?	Yes. The results appeared to be internally consistent.
16. Were the results for the analyses described in the methods, presented?	Yes. The authors reported results for the analyses that they described.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	Yes. The conclusions were justified by the results.

18. Were the limitations of the study discussed?	Yes. The authors discussed the limitations of their study.
Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	No funding sources or conflicts of interest were reported that would appear to affect the authors' interpretation and conclusions.
20. Was ethical approval or consent of participants attained?	Yes.

Table 2-8B. Critical Appraisal of Li et al, 2017 (26), Fedorko et al (29), and O'Reilly et al (44), Using the Cochrane Risk of Bias tool for Randomized Controlled Trials.

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk of bias	Quote: "…patients are randomized used a computerized block randomization schedule with a multiple block size of four." Comment: Probably done.
Allocation concealment (selection bias)	Low risk of bias	Quote: "…technician obtains the treatment allocation through an internet-based automated randomization system." (O'Reilly et al) "Only the technician controlling the hyperbaric oxygen chamber was aware of group allocation for each patient, which was maintained in sequential unique opaque envelops [*sic*] opened as participants were randomly assigned." (Fedorko et al) Comment: Despite the discrepancy in the reported method used to conceal allocation, either method would prevent the researcher from allocating a specific patient to a particular treatment arm.
Blinding of participants and personnel (performance bias)	Low risk of bias	Quote: "Researchers and patients are blinded to treatment allocation; the only unblinded individual is the technician responsible for controlling the hyperbaric oxygen chamber." (O'Reilly et al) "…each study participant is placed into the hyperbaric chamber, but only those patients allocated to active HBOT receive 90 minutes of oxygen at 2.4 ATA with the patients breathing 100% oxygen inside the chamber. Those patients randomized to placebo will be compressed on air to 0.3 ATA (10 feet) and kept at that level." (O'Reilly et al)

		Comment: All patients were placed into the HBOT chamber, which lessened the risk that that randomization was broken. Although the technician would need to know to which treatment arm each patient was allocated (in order to have the chamber at the correct setting based on the arm to which the patient was randomized), it is unknown if steps were taken to prevent the technician from deliberately or inadvertently informing the patient or other study personnel of the arm to which the patient was assigned. Given the steps taken to keep patients and personnel blinded (i.e., sham HBOT treatment), it is possible that the technicians were kept (or dissuaded) from informing the remaining study team of the patient's allocation. While more information is needed to determine the risk of bias, it is likely that there is a low risk of performance bias due to blinding.
Blinding of outcome assessment (detection bias); EQ-5D-3L (secondary outcome)	Low risk of bias	Quote: "Participants self-administered all the [*sic*] three HRQoL instruments in the presence of a trained researcher who could assist with the confusion about the questions being asked in the instruments. Data on HRQoL were all collected at baseline before randomization to treatment group, end of intervention (i.e., week 6) and end of follow-up (i.e., week 12)." (Li et al) "Once all study data for each patient are received, the randomization assignment may be revealed." (O'Reilly et al) Comment: As patients were blinded to treatment allocation, their self-assessments of HRQoL using the EQ-5D-3L would be blinded to treatment allocation as well. Of note, the baseline measures were collected prior to randomization, therefore there is no possible way for knowledge of treatment assignment to bias the baseline measures of EQ-5D-3L.

Incomplete outcome data addressed (attrition bias)	Unknown risk of bias	Comment: the authors did not address the follow-up of patients with respect to EQ-5D-3L data. Li et al reported that missing data were imputed; however, the number of patients with missing data were not reported. EQ-5D-3L data may be missing unequally between treatment groups, which could bias the comparative results of the trial, especially if the reasons for missing data are associated with both the treatment received and EQ-5D-3L (i.e., confounding variables). Without further information on missing data, the risk of bias is unknown.
Selection reporting (reporting bias)	Low risk of bias	Comment: O'Reilly et al clearly outlined the primary and secondary outcomes of the study, and these were reported in the publications by Fedorko et al and Li et al.

Table 2-9. Critical Appraisal of Sothornwit et al, 2017 (27), Using the Appraisal tool for Cross-Sectional Studies (AXIS tool)

	Yes/No/Do not know/Comment
Introduction 1. Were the aims/objectives of the study clear?	Yes. The objective of the study was to investigate the HRQoL of patients with DFUs and to compare HRQoL between patients with diabetes who have a DFU or amputation, who have other diabetic complications, or who have no diabetic complications.
Methods 2. Was the study design appropriate for the stated aims?	Yes.
3. Was the sample size justified?	No. No justification was reported; however, this was an abstract publication.
4. Was the target/reference population clearly defined? (Is it clear who the research was about?)	No. The target population was not clearly defined, nor could it be inferred easily.
5. Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Don't know. As the target population was not clearly defined, it is not possible to determine if the sample frame represents the target population. Also where the sample was taken from was not clearly described; however, it is likely from a population in Bangkok, Thailand, which does not generalize well to the population of interest for the present study (Canada).
6. Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Don't know. The selection process was not described.
7. Were measures undertaken to address and categorise non-responders?	Don't know. The number of people asked to participate was not reported, and no information on non-responders was provided.
8. Were the risk factor and outcome variables measured	Yes. The outcome variable, HRQoL (using EQ-

appropriate to the aims of the study?	5D-5L) was appropriate to the aims of the study. The risk factor groups were those with a DFU or amputation; those with other diabetic complications; and those with no diabetic complications.
9. Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	Yes. HRQoL was measured using the EQ-5D-5L (Thai version), and using time-trade-off methods.
10. Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p-values, CIs)	Yes. The authors noted that mean values were reported; however, they did not clearly report whether the variance estimates were standard deviations, standard error, or some other measure of variance.
11. Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	No. Insufficient methods were reported.
Results 12. Were the basic data adequately described?	No. Baseline data for the participants were not reported.
13. Does the response rate raise concerns about non-response bias?	Don't know. No data on non-responders.
14. If appropriate, was information about non-responders described?	No. No information about non-responders.
15. Were the results internally consistent?	Yes. The results appeared internally consistent.
16. Were the results for the analyses described in the methods, presented?	Yes. The methods section was missing a lot of details regarding the study; however, the authors reported results for those analyses that were described in the abstract.
Discussion 17. Were the authors' discussions and conclusions justified by the results?	No. The results were hypothesis-generating, but the conclusions were much stronger than warranted by the results.

18. Were the limitations of the study discussed?	No. The authors did not discuss the limitations.
Other 19. Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	No. The authors did not report funding sources or conflicts of interest.
20. Was ethical approval or consent of participants attained?	Don't know. The authors did not report whether ethics approval or consent of participants was sought or attained.