2014-04-23

The Prevalence of Depression and the Accuracy of Depression Screening Tools in Migraine Patients

Amoozegar, Farnaz

UNIVERSITY OF CALGARY

The Prevalence of Depression and the Accuracy of Depression Screening Tools in Migraine Patients

by

Farnaz Amoozegar

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMMUNITY HEALTH SCIENCES

CALGARY, ALBERTA

APRIL, 2014

Abstract

Migraine and major depressive disorders are common comorbid conditions. The purpose of this cross-sectional study was to assess how well the Patient Health Questionnaire (9 items) and the Hospital Anxiety and Depression Scale perform as depression screening tools in migraine patients attending a headache clinic, determine the prevalence of depression in this patient population using a gold standard semi-structured psychiatric interview, and examine disability and quality of life for these patients. The Patient Health Questionnaire (cut-point 14) and the Hospital Anxiety and Depression Scale (cut-point 11) produced an optimal balance of psychometric properties for the studied migraine population. The point prevalence of depression was 25.0% (95% CI 19.0-31.0), and the prevalence of untreated depression was 17.0% (95% CI 10.8 – 23.2). Patients with both migraine and depression had significantly higher degrees of disability and a poorer quality of life as compared to patients without depression.

Acknowledgements

I would like to acknowledge and thank my co-supervisors, Dr. Nathalie Jette and Dr. Scott Patten, for their dedication, guidance, support, and hard work throughout the time period of my Master's degree and during the preparation of my thesis. I would also like to acknowledge and thank Dr. Werner Becker for his mentorship in the field of headache. He has been a role model for me and was one of the motivating factors for me to enter the field of headache neurology. I would also like to acknowledge Dr. Andrew Bulloch for his support and feedback as one of my valued committee members over the last two years.

In addition, I would like to thank the study personnel associated with the NEEDS study for all their hard work and support. I would also like to extent my gratitude to Drs. Jette and Patten for providing the resources and funding for me to perform the research related to my thesis, as a sub-study of the NEEDS project.

Table of Contents

List of Tables

List of Figures and Illustrations

List of Abbreviations

BDNF	Brain-derived neurotrophic factor
CCHS	Canadian Community Health Survey
CHREB	Conjoint Health and Research Ethics Board
CHORD	Canadian Headache Outpatient Registry and Database
CI	Confidence intervals
CSD	Cortical spreading depression
DOB	Date of birth
DSM	Diagnostic and Statistical Manuel of Mental Disorders
DSM-IV	Diagnostic and Statistical Manuel of Mental Disorders, 4th edition
HADS	Hospital Anxiety and Depression Scale
HADS – A	Hospital Anxiety and Depression Scale – Anxiety score
HADS – D	Hospital Anxiety and Depression Scale - Depression score
HPA	Hypothalamic-pituitary adrenal
HRQoL	Health-related quality of life scale
ICHD-II	International Classification of Headache Disorders – Edition II
MAOI	Monoamine oxidase inhibitors
MCS	Mental component score
MIDAS	Migraine Disability Assessment Questionnaire
miRNA	micro-RNA
NEEDS	Neurological disease and Depression Study
NPV	Negative predictive value
P	Prevalence
PCS	Physical component score
PHQ-2	Patient Health Questionnaire, 2 items
PHQ-9	Patient Health Questionnaire, 9 items
PPV	Positive predictive value
RNA	Ribonucleic Acid

ROC	Receiver-Operator Characteristic
SE	Standard error
SCID-I	Structured Clinical Interview for DSM-IV for axis I disorders
SCID	Structured Clinical Interview for DSM-IV
SD	Standard deviation
Sens	Sensitivity
SF-12	Short Form 12-Item Health Survey
SF-36	Short Form 36-Item health survey
SSRIs	Selective serotonin reuptake inhibitors
Spec	Specificity

Chapter 1: Migraine

1.1 Introduction

Migraine is a major health problem with an enormous impact on the individual and society. Migraine can cause a great deal of disability, and disrupt family life, relationships, and careers [1, 2]. It can also be associated with a number of conditions, including psychiatric, neurologic, vascular, and cardiac conditions [2]. The societal cost of migraine is high due to a combination of missed work, decreased functionality, medication use and visits to physicians and to the emergency [1]. The Global Burden of Disease study showed that migraine alone accounts for 325 years lived with disability per 100,000 people globally [3]. Given this substantial burden of disease, it is important to have an understanding of what constitutes a migraine, its causes, comorbid conditions and its management.

1.2 Diagnosis and clinical presentation

Migraine is defined as a headache that typically lasts 4 to 72 hours with at least two of the following criteria: unilateral location, pulsating quality, moderate or severe pain intensity, and aggravation by routine physical activity [4, 5]. One or more of the following must also be present: nausea and/or vomiting, or light sensitivity and sound sensitivity (Appendix A). Migraine headaches can be classified as episodic or chronic. Episodic refers to a headache frequency less than 15 days per month. Chronic refers to a headache frequency of 15 or more days per month for a minimum of three months [4, 5].

1.3 Stages of migraine

Migraine can be associated with a number of other symptoms before, during and after the headache. These time periods or phases are the prodromal phase, the headache phase and the postdromal phase [6]. Before a migraine attack, some patients experience prodromal or premonitory symptoms, such as fatigue, irritability, change in appetite, or excessive yawning. This can occur for hours to days before a migraine attack and is thought to be related to hypothalamic involvement [6]. During the attack, patients have a wide range of headache intensity and disability and can experience focal neurological deficits as described below. After the attack, or during the postdromal phase, patients can still have lingering headache for hours to days, as well as fatigue and mood changes [4, 6].

1.4 Aura

About a third of patients with migraine experience one or more auras, which are transient and focal neurological symptoms that usually occur in association with a migraine [6]. Aura symptoms can sometimes be as disturbing as the migraine headache itself. Most commonly, patients experience visual auras, such as coloured lights, zigzag lines, a kaleidoscope, or areas of visual loss (scotomas) [6]. Sensory symptoms such as numbness or tingling in the face, arm or leg are also quite common. Aphasia or difficulty with language function is the third most common aura. Vertigo and imbalance, as well as motor auras, such as hemiparesis are the least common [6]. Auras by definition are transient and usually last 5-60 minutes [5].

1.5 Epidemiology

Migraine can affect individuals throughout the lifespan. The prevalence and incidence of migraine vary by age and gender [2]. Before puberty, the prevalence of migraine is higher in boys than girls (by a ratio of about 1.14:1) [7]. After puberty, female prevalence is higher than male prevalence by a ratio of about three to one; this has been demonstrated in several population-based studies [2, 8-11]. The prevalence of migraine increases in both genders during the forties, after which time it starts to decline [2]. In population-based studies, the point prevalence of migraine is about 15-18% in women and 6-7% in men [10, 12]. Migraine occurs in all races and countries, but seems to be lowest in Africa and Asia [2].

1.6 Pathophysiology

Migraine is considered a functional disorder of the brain, whereby a hyperexcitable brain state makes the patient more susceptible to headaches and other associated features [6]. A phenomenon called cortical spreading depression (CSD) is felt to be an important part of the pathophysiology of migraine. Cortical spreading depression is characterized by an initial wave of synchronized stimulation of cortical electrical activity, followed by a depression [6]. This wave spreads across the cortex at a rate of 3 mm per minute [6]. There is now clinical and experimental evidence that CSD is the biological substrate for aura [13]. CSD is associated with profound and focal disruptions in ionic homeostasis as well as changes in blood flow [13]. Many experts feel that CSD also occurs in patients who do not experience migraine aura. This may be because CSD in these patients occurs in non-eloquent areas of the brain [6].

Cortical spreading depression occurs in the cerebral cortex, cerebellum and hippocampus. As the wave propagates forward, nitric oxide, arachidonic acid, protons and potassium are released extracellularly [6]. These substances activate meningeal nociceptors or pain fibres which are peripheral branches of the trigeminal nerve located on the meninges. These trigeminal neurons, which also innervate meningeal blood vessels, then in turn release calcitonin gene-related peptide, substance P, and neurokinin A [6]. This results in dilatation and inflammation of these blood vessels with plasma protein extravasation. This process is known as sterile neurogenic inflammation [6].

At this point, the peripheral trigeminal nerve fibres or the first-order trigeminal neurons have become activated. This is known as peripheral sensitization [6]. Pain signals are then carried centrally to the trigeminal nucleus caudalis or trigeminocervical complex, which then sends fibres to the thalamus, autonomic nuclei in the brainstem and to the hypothalamus [6]. Thalamic neurons then project to the somatosensory cortex and parts of the limbic system, where pain is perceived and experienced emotionally by the patient [6].

If a migraine attack is treated early, when only peripheral sensitization has occurred, the migraine may be aborted and full activation of trigeminothalamic and thalamocortical neurons does not occur [6]. If the attack progresses however, the trigeminothalamic and thalamocortical neurons are fully activated, leading to central sensitization [6]. This is manifested clinically as cutaneous allodynia, or experiencing pain spontaneously or with a non-painful stimulus [6]. Patients often describe that they have scalp tenderness or that their hair hurts. It is indeed well known that patients who treat their migraine attack once central sensitization has occurred are

less likely to have success with the treatment than those who treat before central sensitization has occurred [6].

1.7 Genetics and the environment:

Migraine is felt to occur as a result of genetic susceptibility and environmental triggers. This combination lowers the threshold for CSD to occur, leading to the processes described above [6]. In the vast majority of cases, migraine is polygenic; multiple genes contribute to its manifestation clinically. There are a few rare monogenic forms that result in familial hemiplegic migraine [6]. These monogenic forms lead to alterations in ionic channels, resulting in increased glutamate, the main excitatory neurotransmitter in the brain [6]. Other data also reveal that migraine patients have dysfunctional pain pathways. There is impairment of descending pain modulatory circuits, leading to a loss of pain inhibition [14]. Overall, these alterations are felt to result in hyperexcitability in nociceptive areas of the brain [14]. In addition to a genetically susceptible host, environmental factors can contribute to the development of a migraine [6]. Common triggers include dietary factors, sleep disturbance, stressful or emotional situations, hormonal fluctuations in women (such as occurs during or just prior to menstruation), as well as weather changes [6].

1.8 Management of migraine

Migraines can be very disabling and may require treatment. First, environmental and lifestyle factors need to be optimized [15]. For example, the patient's diet, sleep, stress level, exercise routine and posture need to be assessed. In some, that may be sufficient to reduce migraine headaches. For those where this is insufficient, medications can be tailored to the

patient's headache characteristics [15]. For individual headache attacks, acute or symptomatic therapy is used. This consists of simple analgesics, such as acetaminophen or ibuprofen, stronger anti-inflammatories, such as naproxyn, diclofenac, or indomethacin, and migraine specific medications, such as triptans and dihydroergotamine [15]. The triptans, which are now the mainstay of treatment for moderate to severe migraine attacks are serotonin receptor agonists [15]. Serotonin receptors are located on trigeminal neurons, the meningeal blood vessels and in the trigeminal nucleus caudalis [6]. Stimulation of these serotonin receptors on the trigmeninal nerve endings prevents the release of calcitonin gene-related peptide [6]. Stimulation of the receptors on the meningeal blood vessels results in a mild level of vasoconstriction, and stimulation of the receptors in the trigeminal nucleus caudalis decreases central neuronal signalling [6].

When headaches are occurring frequently, are prolonged or severe, or patients are using symptomatic therapy too frequently, a preventative medication is appropriate in addition to treatment of individual headache attacks [16]. Several preventative medications have been shown to be effective for migraine prophylaxis in large randomized controlled trials [16]. These medications come from various classes, such as antidepressants, anti-epileptics, anti-hypertensives, supplements and herbs, and botulinum toxin [16]. The specific medication is chosen based on the patient's overall health profile, comorbidities, other medications, and patient preference [16]. In general, preventative medications are felt to raise the threshold for initiating CSD and reduce the frequency of CSD events once the process is triggered [6]. However, the exact mechanism of action of preventative medications are not yet known.

1.9 Disability and quality of life

Migraine headaches can be associated with significant disability and reduced quality of life. It is therefore important to have an understanding of these concepts and some of the tools used to assess these factors. Disability refers to the consequences of a health condition on a patient's ability to function at work, school, or in other roles [17]. In contrast, health-related quality of life is a patient's perceived health status and overall physical and mental well-being [18].

Headache-related disability is an important factor to assess in migraine patients because it can guide the type of treatment used and determine the success of treatment over time [1]. A number of tools have been designed to assess disability in migraine. One of the most commonly used tools is the Migraine Disability Assessment Questionnaire (MIDAS). This tool is simple to administer, easily interpreted and has been validated in population-based samples of headache patients, including migraine patients [17]. The reliability and internal consistency of the MIDAS are high. Spearman's and Pearson's correlations, used to assess reliability, are 0.84 and 0.75, respectively [17]. Cronbach's alpha, used to assess internal consistency, is 0.83 [17]. Discriminant validity has also been shown to be high with the MIDAS, resulting in significantly higher scores for patients with migraine compared to those without migraine [17].

The MIDAS is a simple 5-item questionnaire that can be self-administered and collects data on time lost from work, household duties and social activities over the past 3 months [17]. Patients are asked to answer questions on the number of days missed from these activities over the past 3 months, and number of days where productivity was reduced by half or more in the

last 3 months, because of headaches. The score is then a simple summation of the number of days for question 1-5 (Appendix B) [1]. The questionnaire also asks about frequency and intensity of headaches over the past 3 months but these are not counted toward the total score [1]. A 3 month recall interval has been chosen for the MIDAS to balance the recall of information by patients with a long enough time period to accurately represent the patient's headache disability [17].

Once the scores are summed for the MIDAS, patients can be placed into four categories: grade I to IV. Grade I indicates little or no disability (score of 0-5), grade II indicates mild disability (score of 6-10), grade III indicates moderate disability (score of 11-20), and grade IV indicates severe disability (score of 21+). Because of its simplicity in regards to scoring and interpreting the scores, quick and easy administration, and high reliability and internal consistency, the MIDAS was used in this study to assess disability related to migraine.

A patient's quality of life is also an important determinant of the patient's overall outcome [19]. As a result, measures of quality of life are useful tools for assessing a patient's perceived health. Similar to disability, a number of tools exist to measure quality of life. In this study, the Short Form 12-Item health survey (SF-12) was used because it is a general health survey not specific to any disease, age, or gender [18]. The SF-12 is a shorter version of the Short Form 36-Item health survey (SF-36), which has been extensively validated [18]. The SF-12 contains 12 questions that measure health and well-being as perceived by the patient (appendix not included for copyright reasons). The SF-12 measures eight health domains which encompass physical and emotional well-being, over the last 4 weeks [18]. As such, it provides two summary scores: a physical component score (PCS) and a mental component score (MCS).

A computer algorithm is required to score the SF-12. Higher scores indicate better health-related quality of life [18].

The SF-12 survey has been validated in general populations and has high internal consistency: Cronbach's alpha ≥ 0.82 and 0.75 for the PCS and MCS scales respectively [18]. It also has adequate test-retest reliability: $r = 0.89$ for the PCS and $r = 0.76$ for the MCS [18]. It has been used to assess quality of life in migraine patients but has not been specifically validated in a migraine population. Compared to the more detailed SF-36, the SF-12 performs well in the general population, showing a high criterion-related validity [18]. The SF-12 PCS and MCS have been tested to discriminate between groups of patients that differ in physical and mental health according to a number of clinical factors, such as the seriousness or severity of a health condition [20]. The clinical factors assessed are the same as the SF-36 and correlations have then been determined. The SF-12 physical and mental component scores correlate 0.95 and 0.96 with the SF-36 physical and mental component scores respectively [18]. For this reason, and the fact that the SF-12 is much faster and more practical to administer, we chose to use the SF-12, rather than the SF-36, for the current study.

1.10 Migraine and comorbid conditions

Migraine is comorbid with a number of health conditions, including stroke, cardiovascular disease, as well as a number of psychiatric diseases [2]. The term comorbidity refers to the greater than coincidental association of two health conditions in an individual [21]. Of the psychiatric conditions, depression is the most common comorbidity seen in migraine patients [2, 8, 10]. The next chapter provides background information about depression and the third chapter discusses the associations that exist between migraine and depression.

Chapter 2: Depression

2.1 Introduction

Major depressive disorder is one of the most common mental health disorders worldwide. It is characterized by feelings of sadness and/or loss of interest in enjoyable activities. It is accompanied by a number of other emotional and physical symptoms. Mental and substance use disorders are the fifth leading cause of disability in the world [22], with depression alone ranking in the eleventh place [23]. Therefore, major depressive disorder contributes substantially to the global burden of disease [22]. The World Health Organization has declared depression as a health priority for the coming years [24]. It is therefore important to recognize depression and provide effective management for this condition whenever possible.

2.2 Diagnosis and clinical presentation

Major depressive disorder is defined as one or more major depressive episodes with at least two weeks of depressed mood and/or loss of interest in previously enjoyed activities, plus three to four additional symptoms from the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th edition) criteria including, but not limited to, changes in appetite, changes in sleep or concentration, decreased energy, feelings of worthlessness or excessive guilt, indecisiveness, and recurrent thoughts of death or suicide [25]. In addition, the symptoms must cause clinically significant impairment in social, occupational or other areas of functioning, must not be due to a general medical condition or the direct effects of a substance, and not be better accounted for by bereavement [25]. The words depression, major depression and major

depressive episode are often used interchangeably. In this manuscript, the term depression will be used, with reference to the definition stated above.

2.3 Epidemiology

The Canadian one-year prevalence population-based estimates for depression are between 4.5 and 4.8%. This is based on the Canadian Community Health Survey: Mental Health and Well-being, the National Population Health Survey, and the Edmonton Survey [26]. The Canadian Community Health Survey showed lifetime prevalence estimates to be substantially higher than point prevalence and varying with age. The lifetime prevalence was 8.8% in those 15-25 years, 12.2% in those 26 to 45, 12.4% in those 46 to 64 and 6.4% in those 65 years and over [27]. Depression is more common in women than in men, in the unemployed, in unmarried individuals, in those with lower income, and in those with a chronic medical condition [27].

2.4 Pathophysiology

The exact cause of depression is not fully understood. Depression can occur spontaneously or after a traumatic life event [28]. Histological and neuroimaging studies have shown that alterations in neuroplasticity occur in the brains of patients with depression. Several areas of the brain, including the frontal cortex and hippocampus are involved and research has shown that depression involves major disruptions within the brain's reward circuitry [28]. The brain's reward system is important for survival because it guides the individual's attention and energy towards behaviours that lead to rewards, such as seeking food and shelter. This reward circuitry spans a large and complex interconnection of brain regions with innervations from dopaminergic, serotonergic and noradrenergic neurons [28]. Through disruptions of this system

therefore, dysregulation of the mono-amines: serotonin, norepinephrine and dopamine, play an important part in the development of depression [28].

Over the last several years, neuroscientists have also focused their attention on neurotrophins and their role in depression. Neurotrophins are critical growth factors for the survival and maintenance of neural circuits during development and in adulthood [29]. Research over the last decade indicates that certain neurotrophins (e.g. brain-derived neurotrophic factor, BDNF) and their associated gene expression and binding receptors are decreased in patients with depression [29]. Antidepressant treatments can increase the levels of these neurotrophins [29].

In addition to polymorphisms and alterations in the genes of neurotrophic factors and their associated binding sites, units of non-coding RNA or micro-RNA (miRNA) are felt to play a role in depression as well [29]. Micro-RNA are regulatory molecules which are important for synaptogenesis, neuron survival and neural plasticity [29].

Many studies have shown that depression is also associated with dysfunction of the hypothalamic-pituitary-adrenal axis (HPA axis), leading to a maladaptive response to stressful stimuli [28, 29]. There is an increase in the release of stress hormones, such as cortisol, in patients with depression as compared to those not suffering from depression. Patients with depression also have an overactive response to environmental stressors [28, 29]. The increased stress and glucocorticoid hormones can adversely affect hippocampal neurogenesis [29]. Other hormones such as estrogen are also felt to play a role in depression, explaining the higher incidence of depression in women [28, 29].

The immune system and inflammation are further factors implicated in the pathogenesis of depression. Elevated levels of pro-inflammatory cytokines are found in patients with depression [29]. These cytokines in turn activate the HPA axis, stimulate secretion of growth

hormone and inhibit release of thyroid stimulating hormone, all of which may contribute to the depressive state [29].

In summary, the various factors discussed above, including dysregulation of mono-amines and disruption in the brain's reward circuitry, alterations in neurotrophins, abnormalities in miRNA, dysregulation of the HPA axis, and the increased levels of pro-inflammatory cytokines, are all functionally interconnected [29]. An alteration in one factor can result in changes in other factors, cumulatively resulting in the symptoms of depression. In addition, these changes overall inhibit the brain from making appropriate adaptive behaviours when faced with various environmental stimuli [28].

2.5 Genetics and the environment

Similar to migraine, depression develops as a result of the combination of one or more genetic predispositions and environmental factors, such as a stressful life event [29]. It is believed that about 50% of the risk of depression is due to genetic factors and the other 50% due to environmental factors [29]. Any of the genes or their associated modulatory factors such as miRNA, implicated in the mechanisms above, can lead to a genetic predisposition towards depression [29]. Therefore, similar to migraine, depression involves a complex interplay of genes and the environment.

2.6 Management of depression

As with migraine, the treatment of depression is usually multifaceted, with a combination of lifestyle factors, psychotherapeutic approaches, and medications. Lifestyle factors involve such things as regular exercise, a healthy well-balanced diet, and a supportive and caring social

network [30]. Psychotherapeutic approaches, which are broad in scope and practice, are provided to varying extents and as resources allow. These approaches include patient education, cognitive behavioural therapy and biofeedback [30]. These methods alone can be effective forms of therapy for some patients [30].

Medications for depression are generally used when patients experience moderate to severe symptoms that occur frequently and are disruptive for the patient. They include a wide range of agents that modulate levels of one or more of the mono-amines [30]. The most commonly used medications are the selective serotonin reuptake inhibitors or SSRIs. They comprise such medications as citalopram, fluoxetine and sertraline [31]. Selective serotonin-norepinephrine reuptake inhibitors, such as venlafaxine are also frequently used. The tricyclic antidepressants, such as amitriptyline, are used less frequently due to the need for higher doses and a higher potential for adverse effects [32]. Occasionally, monoamine oxidase inhibitors (MAOIs), such as selegiline, are also used [30]. However, MAOIs have high potential for adverse events and require strict dietary restrictions because combining them with certain dietary elements can lead to dangerously high blood pressure levels [32]. When patients have more than unipolar depression, i.e. bipolar disorder, where mood can fluctuate between a low to a high hypomanic or manic state, then other medications, such as lithium, valproate, and lamotrigine are commonly used as mood modulators [30].

2.7 Screening for depression

Classically, the term screening refers to the early detection of a disease before clinical symptoms or signs are apparent [33]. For example, screening is commonly performed for conditions such as breast or colon cancer, where the stakes of having undetected disease are high and earlier

treatment typically improves patient outcomes [33]. However, screening can also be performed in conditions where clinical symptoms or signs of a disease are subtle or not readily notable, mimic other conditions, take more time to tease out, or require expertise for diagnosis [34]. Sometimes in these cases, stakes of missing a diagnosis are not as high as fatal or rapidly progressive conditions such as cancer. In addition, falsely diagnosing patients with the condition may do more harm than good [35]. Depression can be an example of such a condition. In such cases, screening can be used in the clinical setting to rule out as many people as possible without the condition, i.e. a high true negative rate, with minimal false positives [34]. Therefore, in this context, an effective screening tool generally must have very high specificity. This is in contrast to case finding, which attempts to rule in as many people as possible with the condition, i.e. a high true positive rate, with minimal false negatives [34]. Therefore an effective case-finding tool has very high sensitivity.

Screening in depression remains controversial for a few reasons. First, even though screening may reduce false negatives, it is sometimes at the expense of more false positives [34]. Second, screening tools must be interpreted correctly with the clinician's judgment. Third, screening is only effective if allied with appropriate treatment and follow-up [34]. If a screening test is positive for depression, further assessment is required to determine if the patient does indeed have depression. This requires health-care resources. In Canada, there is high demand but limited mental health resources available [36]. As such, incorporating screening may be costly if it results in a high number of false positive cases, which may result in inappropriate treatment and referrals. Fourth, there are mixed recommendations regarding screening, as described below.

A Cochrane review performed in 2005 found that routine screening for depression had very little, if any effect on the management or outcomes of depression, when patients were followed for 6-12 months [37, 38]. The Cochrane review assessed 12 studies that involved primary care patients and described significant heterogeneity among the studies for recognition of depression and management of depression. The latter was reported only by a sub-sample of studies and only a few studies reported on the outcomes of depression at 6-12 months. Longer-term outcomes were not assessed. Some publication bias was also noted by the authors of this Cochrane review [37]. Some of the shortcomings of the studies in this review may have contributed to the lack of effect seen from screening. The Canadian Task Force for Preventative Healthcare (recommendations updated in 2013) also does not endorse routine screening in adults in the primary care setting [39]. However, this recommendation is weak and is based on very low quality evidence [39]. In addition, there is much less known about screening in specific disease conditions, such as migraine, where depression prevalence is usually higher than in primary care settings.

The US Preventive Service Task Force (updated in 2009) endorses screening for depression in primary care settings when staff-assisted care supports are available to ensure accurate diagnosis, management and follow-up (Grade B recommendation) [40]. They found that there was no significant difference between various validated screening tools used in the studies that they reviewed. These tools included the Center for Epidemiological Studies depression scale, Composite International Diagnostic Interview 2-item depression screen, Primary Care Evaluation of Mental Disorders Mood Module screening items, Geriatric Depression Rating Scale, the Hamilton Depression Rating Scale, the Montgomery-Asberg Depression Rating Scale, the Beck Depression Inventory, and the Patient Health Questionnaire-9

[40]. They recommend that the most practical tool for the given clinical setting be used, and that positive results on a screening tool be confirmed by a full diagnostic interview following the standard criteria from the DSM [40].

Several validated screening tools exist for depression. Popular ones include the Patient Health Questionnaire, PHQ-2 or PHQ-9, containing 2 or 9 items respectively, and the Hospital Anxiety and Depression Scale, HADS. These tools are commonly used because they are quick and easy to administer in clinical settings. The PHQ-2, PHQ-9 and HADS are all used in this study. These tools have been validated as depression screening tools in several large general populations [34, 38, 41-45], but have not been specifically studied in regards to their diagnostic accuracy in migraine patients. Hence, this forms an objective of the current research project; to determine which, if any of these tools, may be the most appropriate to use in migraine patients.

The PHQ-9 is a nine-question screening instrument for depression, but is not a traditional symptom rating scale [41, 46, 47] (Appendix C). Its items map directly to the DSM-IV major depression criteria [25]. The scale is scored in two ways: 1) by summation of item ratings, where each of the 9 questions is scored from 0 to 3, for a total range of scores of 0 to 27 [46, 47], or 2) by using a diagnostic algorithm [41]. The first method produces a total score whereby a decision must be made as to what score to use as a cut-off for determining if a patient has depression or not. Patients below the selected cut-off would be classified as not having depression and those above the selected cut-off would be classified as having depression. The second method uses an algorithm that determines if a patient has depression or not (Appendix D). The algorithm still uses the 9 questions from the PHQ-9 but certain criteria have to be met for the patient to be considered depressed. Both methods were used in this study to determine

which performs better in migraine patients. The PHQ-2, which is a shorter version of the PHQ-9, containing only the first two questions [43, 48], was also assessed in this study.

The sensitivity and specificity of both the PHQ-9 and PHQ-2 in the general population vary somewhat based on the study. In primary care settings assessing the general population, the PHQ-9 has between 61-74% sensitivity, and 91-94% specificity in adults [34, 38, 41]. The PHQ-2 has between 86-97% sensitivity, and 67-78% specificity in adults [34, 38]. In the general populations studied, neither the PHQ-2 nor PHQ-9 worked well in case-finding (because of low sensitivity), but the PHQ-9 was felt to be accurate for screening (because of high specificity). In addition, the algorithm scoring method in one study performed better than the linear cut-off scoring method [34].

In 2012, a meta-analysis was performed to determine the optimal cut-off score for diagnosing depression with the PHQ-9 [42]. Eighteen validation studies in various clinical settings were identified. The clinical settings included primary care, specialized clinics in brain injury, cardiology, stroke, and nephrology, and a few community samples. The pooled specificity (based on the studies that reported specific cut-offs) ranged from 0.73 to 0.96 for the cut-off scores of 7-15. For example, for a cut-off score of 7, the specificity was 0.73 (95% CI 0.63-0.82) and for a cut-off score of 15, the specificity was 0.96 (95% CI 0.94 to 0.97). The sensitivity was much more variable for cut-off scores of 7-15 [42]. In general, the PHQ-9 had acceptable diagnostic accuracy for the diagnosis of depression for cut-off scores of 8 to 11 (sensitivity 0.82-0.89 and specificity 0.83-0.89) [42].

The HADS (appendix not included for copyright reasons) is a tool that addresses both depression and anxiety and is designed specifically to avoid confounding depressive symptoms

with a patient's other illness-related symptoms [44]. The HADS is commonly used in clinical practice and has also been widely used in studies assessing the association of neurological conditions and depression [49, 50]. It has not been studied specifically in the context of migraine. The HADS was designed on the premise that depression in medical conditions may not have the same pattern of physical symptoms as depression diagnosed using the DSM criteria (DSM-III version at the time) [44]. In other words, if using the DSM criteria alone, physical symptoms experienced as a result of a medical condition itself may be falsely attributed to depression, creating false positives. Therefore, symptoms related to physical disorders, such as headache, lack of energy, dizziness, poor sleep, etc., are excluded [45].

The HADS consists of 14 questions, 7 of which assess anxiety and 7 of which assess depression. The odd numbered questions assess anxiety while the even numbered questions assess depression [45]. Each question is scored from 0-3 and the anxiety and depression scores are calculated separately. Scores for each category, anxiety or depression, will range from 0-21, with higher scores indicating a higher level of anxiety or depression [45]. The HADS is scored in a similar fashion to the PHQ-9 by summating the scores for each of the 7 questions in each category and determining a cut-off score whereby patients above that score would be considered anxious or depressed [45].

A literature review in 2002 assessing a number of papers published on the HADS concluded that an optimal balance between sensitivity and specificity was achieved most commonly at a cut-off score of 8+ for both the anxiety score (HADS-A) and the depression score (HADS-D) [45]. This produces a sensitivity and specificity of 0.8 for both subscales. The HADS has also shown high internal consistency across several studies, making it an appropriate tool to test in our study [45].

Screening tools require validation with the use of a gold standard test. The Structured Clinical Interview for DSM-IV (SCID-I for axis I or mood disorders) is an extensively validated and widely used semi-structured research diagnostic interview [51] (appendix not included for copyright reasons). It commonly serves as the gold standard in research studies [52]. For simplicity, the SCID-I will be referred to as the SCID, since no axis II disorders will be discussed in this study. The SCID has been shown to be reliable and valid in several studies (for both the DSM-III-R and DSM-IV) [52]. It is also user-friendly and makes diagnoses based on the DSM-IV [52]. Therefore, our study utilized the SCID as the gold standard test to determine the psychometric properties of the depression scales used in our study. It was also used to determine the prevalence of depression in our migraine sample.

Chapter 3: The Coexistence Of Migraine And Depression

3.1 Introduction

As mentioned in chapter 1, migraine is comorbid with a number of medical and psychiatric conditions. One of the most common of the psychiatric comorbidities is depression. Individuals with migraine are two to four times more likely to develop lifetime major depression as compared to persons without migraine [8, 9, 53-55]. This has been confirmed by a large number of population-based studies which have looked at the association between migraine and depression [8, 9, 11, 53-57]. Despite these studies however, the best tool to accurately identify depression in those with migraine remains elusive. In addition, measurement and sampling issues in some prior studies have led to variable prevalence estimates [8, 9, 11, 53-57]. Therefore, a prevalence estimate determined by a validated psychiatric interview such as the SCID should minimize measurement biases. As such, this study utilized the SCID to determine an estimate of the prevalence of depression in migraine patients attending the headache clinic in Calgary.

3.2 Epidemiology

A meta-analysis performed in 2011 assessed 12 original studies looking at the prevalence of depression in patients with and without migraine [58]. Prevalence estimates of depression were extremely variable, ranging from 3.4% to 24.4% in those without migraine. In patients with migraine, the prevalence estimates for depression ranged from 8.6% to 47.9%, depending on the study [58]. There was more consistency in the odds of depression in those with migraine as compared to those without migraine, ranging from 1.6 to 4.4. The overall OR of the 12

studies was 2.2 (95% CI 2.0 to 2.3), indicating that the odds of depression in a patient with migraine is two times the odds of depression in an individual without migraine [58]. The reasons for the high variability of prevalence estimates in the various studies in this meta-analysis may be due to a number of factors, including different populations studied, variable sample sizes, different study types (cross-sectional and case-control), as well as possible biases not accounted for in some of the studies. Regardless, one can conclude that a patient with migraine has 2-4 times the odds of someone without migraine of developing depression.

A Canadian population-based study using data from the Canadian Community Health Survey (CCHS) found similar results, with a much higher prevalence of depression in migraine patients as compared to those without migraine [59]. This study was not part of the meta-analysis above. The CCHS was a large cross-sectional study conducted by Statistics Canada, and assessed a number of medical conditions, including depression and migraine. Depression was diagnosed with the Composite International Diagnostic Interview Short Form for Major Depression and the diagnosis of migraine was by self-report [59]. The one-year prevalence of depression in the total study sample (n=131, 535) was 7.4% (95% CI 7.2-7.6), and the prevalence of reported migraine was 9.4% (95% CI 9.2-9.7). Patients with migraine were found to have a one-year prevalence of 17.6% (95% CI 16.6-18.6) for depression; much higher than that of the general population, and higher than other chronic medical conditions, which showed a prevalence of 7.8% (95% CI 7.5-8.0) for depression [59]. The association of migraine with depression remained consistent across various age groups. Similar to the meta-analysis, after controlling for age and gender, it was found that migraine patients had a 2.6 times greater prevalence of depression as compared to those without migraine [59]. This study indicates that compared to other chronic medical conditions, migraine headaches make a disproportionally

large contribution at the population level [59]. Although this study did not look specifically at suicide, it is important to recognize that the proportion of suicidal ideation and attempts are also greater in patients with both migraine with aura and depression, as opposed to having one of these conditions [60].

The studies above did not assess migraine subtypes, but other studies have looked at migraine with aura and migraine without aura. It has been found that the association with depression is stronger in patients who have migraine with aura [54, 56, 61-63]. A large Norwegian general population study assessing 9,000 participants found that women who had migraine with aura had a higher odds of depression (OR 2.24, 95% CI 1.57-3.18) than women who had migraine without aura (OR 1.30, 95% CI 1.06-1.61) [63]. Interestingly this association was not found in men; that is, no significant difference was seen in men who had migraine with aura as compared to men who had migraine without aura. A case-control study looking at 1250 adults with recurrent depression versus 851 controls found that participants with recurrent depression had a higher odds of having migraine with aura (OR 5.63, 95% CI 3.54-9.00) as opposed to migraine without aura (OR 3.67, 95% CI 2.20-6.14) [64].

In addition, several studies have shown that patients with chronic or daily headache have a higher prevalence of depression as compared to individuals with episodic headache [65]. One of the larger studies found that chronic daily headache sufferers were about four times as likely to report depression (OR 4.4, 95% CI 2.9-6.5) as opposed to episodic headache sufferers [66]. This study included chronic daily headache patients that had migraine and other types of headache. When migraine was considered separately, the effect was even more dramatic. In patients with frequent disabling migraines, the odds of depression was over 30 times higher (OR 31.8, 95% CI 12.9-78.5) as compared to patients with episodic headaches [66].

3.3 Risk factors for the coexistence of migraine and depression

Various risk factors for comorbid migraine and depression have been investigated. Migraine with aura appears to be the most prominent risk factor for psychiatric comorbidity, with several studies showing that patients who have migraine with aura have a higher risk of developing depression, as shown above [54, 56, 61-63]. As described above, patients who have chronic daily headache and frequent disabling migraines also have a higher odds of developing depression [65, 66]. A number of studies have also examined sociodemographic factors. One large population-based study found that patients who were widowed, separated, divorced, or were in a lower income category were more likely to have coexisting migraine and depression [10]. In addition to the above risk factors, childhood adversities, previous depression, and female gender all increase the likelihood of developing frequent or severe headaches as an adult [67].

3.4 The temporal association between migraine and depression

Studies assessing the temporal association between migraine and depression indicate that migraine precedes the onset of depression, but studies are inconsistent regarding the reverse association. In several longitudinal studies, it has been observed that the relationship between migraine and depression is a bidirectional one. This is well demonstrated by the population-based studies done by Breslau over the years [8, 53-55], showing that the likelihood of developing major depression was two to three times higher in patients with a prior history of migraine, as opposed to those without a prior migraine history. The reverse relationship was also shown. However, a recent retrospective cohort study using 12 years of follow-up data from the Canadian National Population Health Survey showed that migraine is associated with the later

development of major depressive episodes but an association in the other direction was not seen [68]. Long-term prospective studies are needed to clarify this association more definitively.

3.5 Pathophysiology

The exact biological mechanism by which migraineurs are susceptible to depression is not clear. However, there is evidence that several potential pathways may be involved. These include genetic factors, serotonergic dysfunction, ovarian hormone influences, and hypothalamic-pituitary adrenal (HPA) axis dysregulation [69]. Other findings include a dopamine D2 receptor genotype implicated in the pathophysiology of comorbid migraine and depression, and reduced tyramine conjugation, a marker of endogenous depression [58]. One study found migraine patients to have reduced tyramine conjugation when they had a lifetime history of depression, as compared to patients without migraine [58].

In the last several years, estrogen has received a great deal of attention in the pathophysiology of comorbid migraine and depression. Epidemiological, clinical and basic science studies have led support for this hypothesis [62]. Both migraine and mood disorders occur more commonly in women than men. Certain time periods increase the vulnerability to both conditions, such as puberty, perimenstrual time points, the post-partum period, and the perimenopausal period. Specifically, estrogen withdrawal seems to play a role [62]. Estrogens have been shown to modulate a number of factors, including neurotrophic factors and neuropeptides implicated in both migraine and depression. This includes modulation of the mono-amines, in particular serotonin. It appears that estrogen overall exerts an agonistic effect on the serotonergic system [62]. As described in chapters 1 and 2, serotonin plays an important role in the pathophysiology of both migraine and depression. Furthermore, estrogen and its

receptors are highly localized to regions of the brain implicated in both conditions, including the hypothalamus and limbic system. It is likely that estrogen has a modulatory effect on these structures and therefore can affect functions associated with the HPA axis and limbic structures as well as their widespread connections to other areas of the brain [62].

3.6 What is the outcome of patients with both migraine and depression?

Individuals who have both migraine and depression have poorer health outcomes than patients with either condition alone. A population-based study using data from the 2002 Canadian Community Health Survey assessed health-related outcomes in patients with migraine and various psychiatric conditions [10]. It found that those suffering from both migraine and a mental health disorder, including depression, had poorer health-related outcomes than those who had only one of these conditions. The health-related outcomes assessed in this study included 2-week disability, restriction of activities, poorer quality of life, and mental health care utilization. All of these factors were more likely in patients with comorbid migraine and a mental health disorder, as compared to one of these conditions alone [10].

In a different study assessing 1007 young adults, Breslau et al found similar results. Patients with both migraine and mental health disorders were more likely to report job absenteeism, use more mental health services, and report their general health as fair or poor [8]. Health-related quality of life has also been assessed in patients with both migraine and a mental health disorder, using validated scales such as the health-related quality of life scale (HRQoL). Studies in this area indicate that patients with both health conditions have lower mean scores in most categories of the HRQoL as compared to those without migraine [10].

In addition to poorer overall outcomes in patients with both migraine and depression, response to migraine treatment also appears to be worse in this category of patients as compared to patients with no psychiatric disease [19, 62]. One study used a population-based survey of adults to identify migraine sufferers among a large representative sample of the French general population [19]. Migraine was identified in the survey by using the ICHD-II diagnostic criteria (Appendix A). The study then performed a more in-depth data collection and analysis, in those defined as having active migraine (at least one migraine in previous 3 months). Data on depressive symptoms was also collected in these patients using the HADS [19]. The study showed that patients with both migraine and depression (as identified by the HADS) had less pain relief at 2 hours after acute migraine therapy, decreased tolerability to treatments, slower resumption of daily activities, and less satisfaction with their acute therapy, as opposed to patients with migraine but without any psychiatric disease [19].

3.7 Why is a study needed in this area?

Given the strong association between migraine and depression and the significant burden of disease, one can see that a validated screening tool for depression in migraine patients would potentially be useful. Very little research has been done in this area. It is not known which tool, of several available to screen for depression, may be appropriate for this patient population. A review of the literature demonstrates no studies comparing any of the depression screening tools in those with migraine. In addition, in the studies that have used depression screening tools in migraine, the sensitivity and specificity of the available tools have not been validated with a psychiatric interview, which remains the gold standard for the diagnosis of depression.

In this study, the psychometric properties of the PHQ-9, PHQ-2 and HADS questionnaires were assessed using the SCID as the gold standard. This study is the first of its kind performed in migraine patients. In addition to the advantages stated previously for the PHQ-9 and HADS, these scales were selected for this study because the PHQ-9 is the most widely used depression scale in North America and the HADS is the most widely used scale in Europe. Furthermore, the items for the PHQ-9 are based on the DSM-IV and those for the HADS are not, which also provides a nice contrast between the two scales. This study also estimated the prevalence of depression in migraine patients attending the Headache clinic in Calgary, using the gold standard psychiatric interview. This is helpful information as prevalence estimates have been variable in prior studies. Disability and quality of life were also assessed in our study given their importance in both migraine and depression.

3.8 Impact and significance of this project

Given that migraine and depression often go hand-in-hand, it is important for health-care professionals to have information about the performance of scales that they may use to screen migraine patients for depression. This is especially relevant when patients have risk factors for these comorbidities, since having both migraine and depression results in poorer health outcomes than having either condition alone. Therefore, with appropriate screening, patients can hopefully receive management of their multiple health conditions. In addition, having better tools to diagnose depression may reduce the significant burden of disease experienced by these patients, not only from a mental health perspective, but also from a migraine perspective. The findings should also guide future practice and increase awareness and understanding of these common comorbid conditions.

Chapter 4: Study Methodology

4.1 Overview of Study Objectives

1. Assess the sensitivity, specificity, and predictive values of two depression screening tools: PHQ-9 (Patient Health Questionnaire) and HADS (Hospital Anxiety and Depression Scale), in migraine patients that attend the headache clinic in Calgary using a gold standard semi-structured psychiatric interview (SCID).
2. Assess the prevalence of depression in migraine patients attending the headache clinic in Calgary, using the SCID.
3. Examine outcomes, specifically quality of life and disability, associated with depression in those with migraine attending the headache clinic in Calgary.

4.2 Study design

This study is cross-sectional in design. Data was collected for each patient by questionnaires and interviews. The site for this study was the Calgary Headache Assessment and Management Program (headache clinic), which is located within a tertiary care center (South Health Campus) in Calgary, Alberta, Canada. The patients seen at this clinic are referred by their primary care physician or another specialist for diagnosis and management of a wide spectrum of headache disorders. However, migraine is the most common type of headache seen at the clinic, representing >80% of patients seen in the clinic.

4.3 Sampling strategy

The sampling strategy for this study involved recruitment of consecutive patients who met all eligibility criteria for the study. Study recruitment began in January of 2013 and was completed by July 2013. Patients who gave consent to participate in the study were administered a number of questionnaires, which encompassed the first phase of the study. The second phase of the study involved psychiatric interviews for patients if they accepted to participate in the interview portion of the study.

4.4 Study Population

The population for this study included follow-up patients from the headache clinic with acute or chronic migraine headaches. New patients were not included, as a diagnosis of migraine was not yet established for those patients. The proposed sample size was 300 patients.

4.4.1 Inclusion Criteria

1. Patients 18 years of age or older
2. Diagnosis of definite or probable episodic or chronic migraine as per the International Classification of Headache Disorders-Edition II (ICHD-II) criteria (Appendix A)
3. Able and willing to provide informed consent
4. Able to speak, read, write and understand English
5. Have an appropriate level of hearing to be able to complete the psychiatric interview on the phone
6. A resident of Alberta with a valid Provincial Health Number

4.4.2 Exclusion Criteria

1. Headache not meeting ICHD-II criteria for migraine
2. Major medical conditions that may interfere with the patient's ability to participate in the study, such as physician-diagnosed dementia, aphasia, or moderate or severe developmental delay

4.5 Sources or Methods of Recruitment

All consecutive follow-up migraine patients attending the headache clinic were invited to participate in the study, as they checked in for their regularly scheduled appointments. The study was not advertised elsewhere and no other methods of recruitment were utilized.

4.6 Sample Size Justification

Sample size for this study was determined by a precision-based calculation. Based on this calculation, if there is adequate precision once the 95% confidence intervals (CI) are calculated, then there is also an adequate sample size. The expected standard error (SE) of a proportion is: $SE = \sqrt{((P)(1-P) / n)}$, where P is the proportion of patients with depression and n is the proposed sample size. The literature estimates that approximately 25% of migraine patients suffer from depression [8, 9, 53, 55, 58]. The 95% CI's are then determined by multiplying by 1.96. For n=300, SE is 0.025 and 95% CI is 0.05. This was felt to be a reasonable degree of precision and therefore 300 patients were presented with the full consent in this study.

4.7 Data Collection

4.7.1 Phase I – Screening Questionnaires and chart review

Phase I of the study involved asking all eligible patients seen at the headache clinic if they would be willing to be approached by a research assistant regarding the study. If they agreed, they were then presented with the full study consent. Once the patient consented, they were then presented with the questionnaires to fill out, including the PHQ-9, HADS, SF-12, and the MIDAS questionnaires. The order in which the PHQ-9 and HADS questionnaires were given to the patient were randomized so as to reduce any possible influence of the patient's responses by answering one set of questions first followed by the other. These questionnaires were self-administered.

Self-reported demographic and medical information was also collected (Appendix E). This included marital status, highest educational level achieved, working at a paid job, primary source of income, alcohol consumption, smoking, illicit drug use, experiencing any medication side effects, diagnosis of depression by a health-care professional, and if depression was present, who was managing the depression (family physician, psychiatrist, another specialist, counsellor, psychologist, other, and/or no one managing depression). In addition, patients were asked if they were on antidepressants or receiving other forms of treatment for depression (medications, counselling, cognitive behavioural therapy, group therapy and/or other).

In phase I, data was collected by paper and pencil questionnaires. All questionnaires were reviewed for completeness and missing data. If one or more questions remained unanswered, the participant was asked if they missed the question or did not wish to answer it. If the question had been accidently missed, the patient was asked to answer it. If the patient did not wish to answer the question, then that question was marked as having missing data.

About four weeks following each patient's visit, a chart review was performed for the patient to collect additional information (Appendix F). These data were collected four weeks later to ensure that the consultation note from the physician seeing the patient was transcribed and available. A chart review was performed for all eligible patients during the study period, including those not completing the questionnaires. This was performed in order to capture some headache and depression-related data for all eligible patients at the headache clinic and for the assessment of selection bias. Patients completing the questionnaires and the SCID could then be compared to this larger group to ensure a representative sample of patients had been collected. The data abstracted from the chart review included the patient's date of birth, gender, migraine type (migraine with aura or without aura), episodic versus chronic migraine, current psychotropic medications, and current migraine preventative medications. Psychotropic medications included antipsychotics, antidepressants, benzodiazepines and related sedative/hypnotics, and mood stabilizers.

4.7.2 Phase II – Structured Clinical Interview for DSM

All patients consenting to the study were invited to undergo a semi-structured psychiatric interview. Phase II used the Structured Clinical Interview for DSM-IV (SCID) [25]. The SCID is the gold standard test in this study, providing definite evidence that a patient does or does not have depression based on DSM-IV criteria. These interviews were performed by trained senior psychology graduate students by phone within 2 weeks of the clinic visit and study enrolment. The interviewers were blinded to the results of the depression screening tools, so as to avoid any bias with asking the questions or interpreting the answers given by patients.

4.8 Data Management

This study was part of a larger study: The Neurological diseasE and Depression Study (NEEDS). As such, two study-related databases were used to manage the data. The first was the NEEDS manager, a secure online program developed by the Clinical Research Unit at the University of Calgary specifically for the NEEDS study. This program was used to hold a few key pieces of patient information, track each patient's involvement in the study, and schedule SCID interviews. The patient information entered into the NEEDS manager included the patient's name, regional health number, date of birth, clinic appointment day/time, physician's name, contact information (for the SCID), and study eligibility. Tracking patient involvement in the study included information on whether the patient gave consent to participate in the questionnaire and SCID parts of the study, and if the patient cancelled or was a no-show to the appointment. For scheduling SCID interviews, participants were able to view all possible day/times and were able to select a convenient time to be called.

The second database was the Idatafax, an online data management program which held the vital data collected for the study. Idatafax was also developed and maintained by the Clinical Research Unit at the University of Calgary. The questionnaires filled out by the patients, the data abstraction forms from the chart review, and the data from the SCID interviews were all initially recorded on hard copies compatible with the Idatafax program. Hard copies were then faxed to a specific number at the Clinical Research Unit and then a digitized replication of the hard copy was created in the database. The program was equipped with electronic data capture, allowing it to "read" the data filled in on the hard copies and populate them on the digital copy. However, a manual check of the data was still performed to ensure each piece of data had been captured accurately. Each sheet had an individual patient study ID number at the top, so once

the data had been finalized, it was stored in the Idatafax program under that specific patient study number. The data collected into the Idatafax program did not contain any personal identifying information for the participants of the study.

Any data collected for a patient was kept strictly confidential. Only permitted study researchers had access to this data. All data binders and research material were kept in a locked cabinet in the research office with access available only to research staff.

Once the data for all the patients had been collected into Idatafax, it was transferred into Excel. A different spreadsheet was created for each of the questionnaires, the chart review and the SCID. Variable names had been shortened or abbreviated and responses had been coded. Therefore, a data dictionary was created to identify each variable and its coding in detail, and the data was then ready for transfer into a statistical software. STATA statistical software version 12.0 was used for all data analysis in this study. In regards to analysis, the data has been reported in aggregate form and does not contain any associated personal identifying information.

Once the data was transferred into STATA, recoding was required to create variables that were compatible with STATA and had meaningful names and codes. The data was then assessed for accuracy and completeness, initially by visually scanning the variables, then by tabulating results or using a histogram to assess for any outliers or unlikely values.

4.9 General statistical Analysis

4.9.1 Descriptive statistics

The data collected was initially assessed descriptively. First, the number of patients seen at the Headache clinic and for each portion of the study was determined and displayed in a flow chart (Figure 5.1). Second, baseline characteristics and demographics for the group of patients

completing the chart review and all patients completing the questionnaires were tabulated. The specific variables collected are outlined in the data collection section above (Section *4.7.1*). Third, the group completing the chart review and those completing the questionnaires and SCID were compared in terms of key baseline features (age, gender, migraine type, migraine frequency, receiving antidepressants, and receiving migraine preventatives). This comparison was performed to assess whether or not the group completing the questionnaires and SCID was representative of the overall eligible study population. All variables were reported in proportions with the exception of age, which was reported as a mean, standard deviation and range for each group.

4.9.2 Scoring of scales and questionnaires

The PHQ-9, HADS, MIDAS and SF-12 questionnaires were scored according to published criteria (please see corresponding sections in Chapters 1 and 2). All questionnaires with the exception of the SF-12 were scored without any special computer programs. The SF-12 was scored into its physical and component scores (PCS and MCS) using the SF-12 analysis software. Means, medians and ranges were reported for all questionnaires and the distribution of scores was displayed by histograms. As thoughts of death or hurting oneself (question 9 of the PHQ-9) are considered quite important, this question was analyzed separately as well. Research has shown that responses to question 9 do predict suicide [70]. The number of patients and corresponding proportions were reported for the responses to this question. For the MIDAS, the proportion of patients in each disability category was also reported along with overall patient responses to the last three questions of the MIDAS (which are not calculated as part of the score).

4.9.3 General data analysis

As not all patients completing the questionnaires would go on to do the SCID, an important comparison would be between the patients completing the questionnaire only and those completing the questionnaires and SCID. For these groups of patients, the following comparisons were made to determine if any statistical differences existed between them: age, gender, depression by self-report, PHQ-9 scores, question 9 of the PHQ-9, PHQ-9 algorithm, HADS (anxiety and depression scores), MIDAS, and SF-12 (PCS and MCS). For the comparisons, the statistical test appropriate to the variable was used: a student's t-test was used for age; a two sample test of proportions for gender, depression by self-report, and the PHQ-9 algorithm; Fisher's exact test for question 9 of the PHQ-9 and MIDAS categories; and a Wilcoxon rank sum test for all the other variables. P values <0.05 were considered statistically significant.

4.10 Statistical analysis pertaining to each objective

4.10.1 Objective 1: Determination of cut-points, ROC analysis, and calculation of test psychometric properties for the depression screening scales

The first step of objective 1 involved the determination of cut-points for the PHQ-9, PHQ-2 and HADS. Patients with scores above and equal to a selected cut-point would be deemed to have depression and would be classified as test positive for that tool. Those with a score below the cut-point would be deemed not to have depression and would therefore be classified as test negative. The gold standard SCID interview would then establish the true status of these patients in regards to depression.

The cut-points for the PHQ-9 and HADS were assessed by Receiver-Operator Characteristic curves (ROC curves) to allow for an assessment of several cut-points. As the PHQ-2 was assessed by 2 cut-points only, an ROC curve was not necessary as part of the analysis. The ROC curve is based on a plot of sensitivity over (1 – specificity). The ROC curve is able to demonstrate the trade-off between sensitivity and specificity, i.e. the fact that as one increases, the other decreases [33]. In general, cut-off values in the upper left corner of an ROC curve produce a reasonable trade-off between sensitivity and specificity. However, if the goal is to minimize false positives, then a higher cut-off value with higher specificity is appropriate [33]. ROC curves also allow quantification of the area under the curve. A perfect test (100% sensitive and specific) would yield an area under the curve of 1.0. A test that is no better than chance would yield an area under the curve of 0.5 [33]. Therefore the higher the area under the curve, the better the overall performance of the test.

Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were then calculated for the most appropriate cut-point of each tool and several other cut-points for the PHQ-9 and HADS to allow for comparison. 95% confidence intervals were calculated around each of these measures to address the issue of random error. The percentage of patients correctly classified was also assessed for the PHQ-9 and HADS.

Sensitivity, the probability of testing positive given the presence of disease, was calculated as the number of patients who test positive for depression among those truly depressed over the total number with true depression (as determined by the SCID) [33]. Specificity, the probability of testing negative given the absence of disease, was calculated as the number who test negative for depression among those without true depression over the total number without true depression (as determined by the SCID) [33]. Positive predictive value, the probability of

disease given a positive test result, was calculated as the number of patients with true depression who test positive, over all patients who test positive [33]. Negative predictive value, the probability of no disease given a negative test result, was calculated as the number of patients without true depression who test negative, over all those who test negative [33]. These values were also determined for the PHQ-9 algorithm, which was assessed with a binary response; i.e. patient is depressed or patient is not depressed (Appendix D).

The rationale for the selection of cut-points is provided in the discussion section of this manuscript. However, the trade-off between sensitivity and specificity, the percentage of patients correctly classified, and the nature of misclassification was explored to make decisions regarding the most appropriate cut-points. Misclassification refers to errors that can occur in a study as a result of measuring the data [33]. Specifically, in our case, it refers to patients incorrectly classified as depressed or not depressed. Therefore, assessment of misclassification involved an examination of the number of false positive and false negative patients for the PHQ-9 and HADS. As a final step, the psychometric properties calculated above were tabulated to allow for a comparison of all the depression screening tools used in the study.

4.10.2 Objective 1: Stratified ROC analysis

Step 2 of objective 1 involved a stratified ROC analysis for the PHQ-9 and HADS to assess specific variables, including gender, age groups, episodic versus chronic migraine, and migraine type, to determine if these variables produce a different optimal cut point than the overall group of migraine patients. Sensitivities, specificities and their corresponding 95% confidence intervals were then calculated for each subgroup of patients.

4.10.3 Objective 2: Determining the prevalence of depression

The treated, untreated, and overall prevalence of depression in migraine patients was determined using the SCID (gold standard). Treatment was defined as all patients currently receiving antidepressants and mood stabilizers as per the chart review, and/or receiving any form of treatment for depression based on self-report (medication, counsellor, psychologist, etc.).

Point prevalence was calculated using the gold standard SCID results, and for comparison, also using the PHQ-9 and HADS, at the selected cut-points. Point prevalence was calculated as the total number of patients with depression (as per SCID results or questionnaire results) over the total number of participants. For patients who did only the questionnaires and not the SCID, the PHQ-9 and HADS were used to determine the prevalence of depression using the selected cut-points. This would allow one to determine whether or not this group of patients differed from the group that did the questionnaires and the SCID. 95% confidence intervals were calculated around the estimates to determine degree of random error.

4.10.4 Objective 2: Stratified prevalence estimates

Using the SCID results, a stratified analysis was also done to assess prevalence of depression for the following variables: gender, age groups, migraine type, migraine frequency, on migraine preventatives, on antidepressant medications, being treated by any method for depression and diagnosis of depression by a health professional. Comparisons were made by 2-sample tests of proportions, generating exact p-values, and 95% confidence intervals for the prevalence estimates were also calculated. The prevalence of untreated depression was then determined by the proportion of patients diagnosed as having depression by the SCID in the group that was receiving no depression treatment. In addition, depression variables, including

the proportion of patients on antidepressants and/or on depression treatment among the depressed patients were described in more detail. The sensitivity of the PHQ-9 and HADS was also calculated for the subset of migraine patients that were diagnosed as depressed by the SCID but not on any antidepressant medications.

4.10.5 Assessment of questionnaire scores in various patient groups

Another important factor to assess was to see if different groups of patients had statistically significant differences in their overall PHQ-9 and HADS scores. Factors assessed included gender, age groups, migraine type, migraine frequency (episodic vs. chronic), depression as diagnosed by the SCID, depression by self-report, treated for depression, on antidepressants, and on migraine preventatives. All comparisons were made using the Wilcoxon rank sum test.

4.10.6 Objective 3: Assessment of outcomes - quality of life and disability

The MIDAS scores and SF-12 physical and mental component scores were descriptively assessed in the first part of the statistical analysis section. Further analysis was done for objective 3 to assess differences between groups. Similar to the analysis of the PHQ-9 and HADS, the MIDAS and SF-12 PCS and MCS scores were compared for the following factors: gender, age groups, migraine type, migraine frequency (episodic vs. chronic), depression as diagnosed by the SCID, depression by self-report, treated for depression, on antidepressants, and on migraine preventatives. For the MIDAS, the number of patients in each disability category was assessed and comparisons between subgroups were made using Fisher's exact test. For the SF-12 PCS and MCS, median scores and their 95% confidence intervals were reported. All comparisons for these subgroups were made using the Wilcoxon rank sum test.

4.11 Further applying the overall results

As this study assesses migraine patients from a headache clinic, it cannot be directly generalized to the entire migraine population. However, Bayes theorem can be used to predict the positive predictive value (PPV) in other populations with a different prevalence rate of depression. The formula for this is: PPV = (Sens*P)/[(Sens*P) + (1 – Spec)*(1-P)], where Sens is sensitivity, P is prevalence and Spec is specificity. The most suitable estimate of sensitivity and specificity would be used based on results obtained from the analysis.

4.12 Ethical approval

The NEEDS study was submitted to the University of Calgary Conjoint Health and Research Ethics Board (CHREB) and received approval. Permission was then granted from CHREB to conduct this sub-study as a Master's thesis project. The study was not overburdening to patients, was not invasive or interventional and the knowledge gained from this research has the potential to benefit the general public, including the individuals involved in the study. It will also contribute to overall advancement of knowledge in the field.

Chapter 5: Results

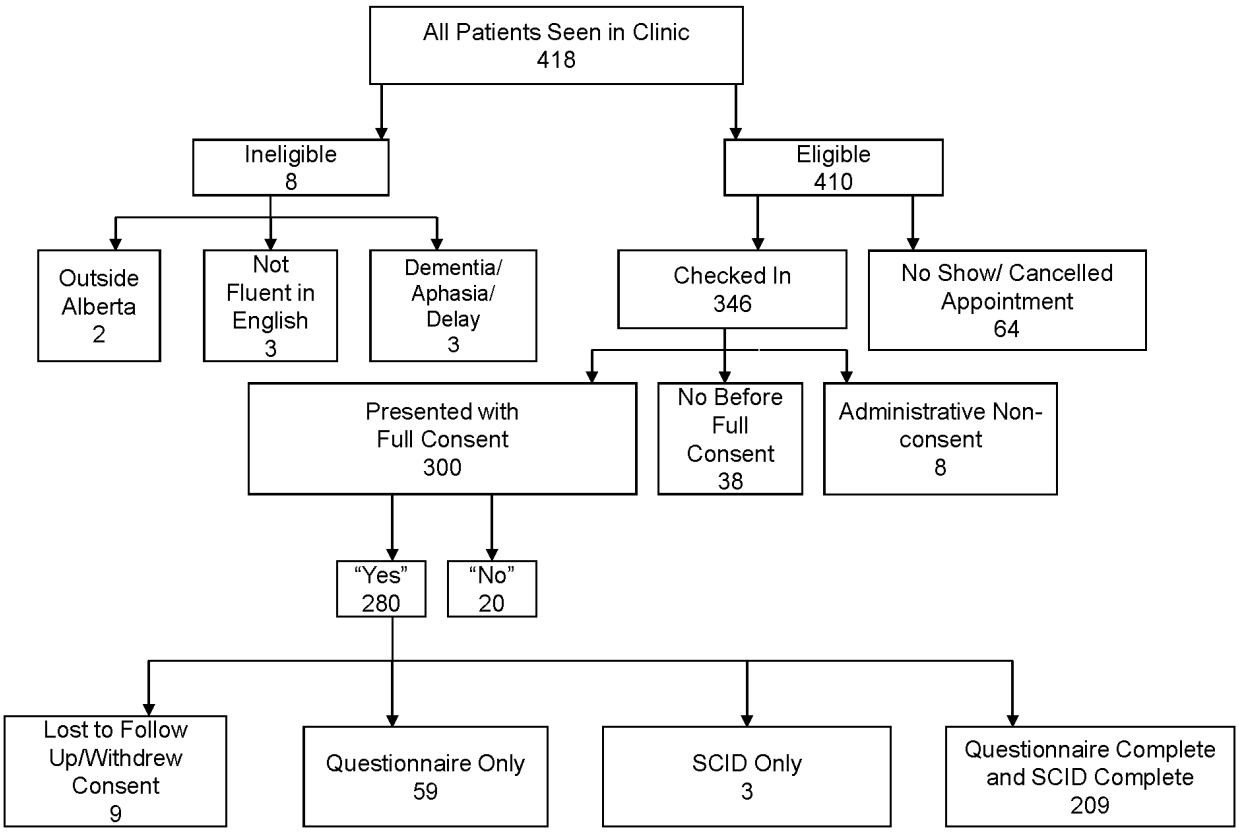
5.1 Overview

Figure 5.1 shows the number of patients seen in the Headache clinic during the time span of the study and the number of patients completing each portion of the study. A total of 418 patients were scheduled for visits to the clinic between late January 2013 and early July 2013. Of these patients, 410 met eligibility criteria for the study. A chart review was performed therefore on these 410 patients. Of these patients, 300 were then presented with the full consent. The rest either did not attend their appointment, did not consent to being approached for the study, or were not consented because they were called in for their appointment before consent could be obtained (administrative non-consent). Of the 300 patients presented with the full consent, 280 patients or 93.3% agreed to participate in the study. Of these patients, 209 (69.7% of the 300 patients, or 74.6% of the 280) completed both the questionnaires and the SCID, 59 patients completed only the questionnaires and 3 patients completed only the SCID. These latter 3 patients took the questionnaires with them to complete at home but unfortunately did not return the questionnaires despite multiple follow-up calls and reminders. Those who did not consent to the SCID did not feel comfortable giving out personal information over the phone or did not have time to complete the SCID.

The data collected for each patient was assessed for accuracy and completeness. Errors occurred infrequently and were corrected by going back to the source data if necessary. Each major variable was tabulated or graphed to assess for outliers. Only a few outliers were found in the data for the MIDAS and these were treated as missing data. Missing data occurred for both demographic information and for the questionnaires but was relatively infrequent. As a result,

the missing information pertaining to a specific variable was not used in the data analysis of that variable. No specific statistical methods were used to compensate for the missing data. The reason for this was that the missing data was sufficiently rare (< 3%) that removing it from the specific calculation or analysis would not alter the overall result. In addition, missing data for a specific patient did not allow proper calculation of total scores for a given questionnaire. As a result, some questionnaires had a slightly lower number of patients with a total score as compared to the numbers shown in figure 5.1.

Figure 5.1 Flow chart demonstrating the number of patients at each stage of the study.



5.2 Descriptive statistics

5.2.1 Baseline characteristics and demographics

Baseline characteristics and demographic factors are shown in Tables 5.1 and 5.2. Table 5.1 demonstrates data abstracted from the chart review (n=410). Table 5.2 demonstrates the characteristics of all patients completing the questionnaires (n=268). These tables are shown separately as some of the variables collected were unique in the two groups of patients. Table 5.3 compares key baseline features for the group of patients completing the chart review and the group completing the SCID and questionnaires (n=208), which forms the core group for data analysis. Later in section 5.3, Table 5.4 compares key features between the patients completing the questionnaires and SCID and those completing only the questionnaires (n=59).

Table 5.1 Characteristics of patients from chart review

Variable		**Distribution** Number of patients, unless otherwise indicated
Mean age (SD, range) (years)		42.6 (13.4, 18-87)
Female gender		330 (80.5%)
Migraine Type	Without aura	275 (67.1%)
	With aura	135 (32.9%)
Migraine Frequency	Episodic	248 (60.5%)
	Chronic	162 (39.5%)
Psychotropic use		22 (5.4%)
Antidepressant use		140 (34.1%)
Benzodiazepines/Sedative use		77 (18.8%)
Mood stabilizer use		9 (2.2%)
Migraine preventative use		381 (92.9%)

SD = standard deviation

As shown in Table 5.1, the mean age of patients was 42.6 years, 330 (80.5%) were female, 135 (one third) had migraine with aura and 248 (60.5%) had episodic migraine. Psychotropic medications were used infrequently, the most common being quetiapine and aripiprazol. Antidepressant use was most common among patients, with 140 patients (34.1%) on an antidepressant. The most common antidepressants used were duloxetine, venlafaxine, buproprion & citalopram. Benzodiazepines were used by about one fifth of patients, the most common being zopiclone and lorazepam. Few patients used mood stabilizers, the most common being lamotrigine. In terms of migraine preventatives, the majority of patients (92.9%) were

taking one or more of these medications. The most common preventatives included botulinum toxin, amitriptyline and topiramate.

Table 5.2 Characteristics of patients completing questionnaires

Variable		**Distribution** Number of patients, unless otherwise indicated
Mean age (SD, range) (years)		42.4 (13.4, 18-76)
Female gender		215 (80.2%)
Marital Status	Married	150 (56.0%)
	Single	59 (22.0%)
	Common-law	32 (11.9%)
	Divorced, separated or widowed	27 (10.1%)
Education above grade 12		147 (54.9%)
Working in paid job		169 (63.1%)
Income Source = employment		155 (57.8%)
Alcohol use		169 (63.1%)
Smoking		35 (13.1%)
Illicit drug use		16 (6.0%)
Medication side effects		126 (47.0%)
Diagnosis of depression by a health professional		123 (45.9%)
Taking an antidepressant		86 (32.1%)
Other forms of therapy for depression		59 (22.0%)

SD = standard deviation

Table 5.2 shows a very similar mean age and gender distribution as compared to the patients from the chart review. Some key features to highlight are the following: 126 patients (47.0%) described having medication side effects and 123 patients (45.9%) stated that they had been diagnosed with depression by a health-care professional. Of the 123 patients, 83 patients (67.5%) were being managed by their family physician and 24 patients or close to 20% were being managed by a psychiatrist for their depression (not shown in table). The others were being managed by a psychologist, counselor, or no one was managing their depression. Eighty-six patients (32.1%) reported taking an antidepressant, which is very similar to the number found during the chart review (34.1% of patients). Overall, 59 patients (22.0%) described using other forms of therapy for depression.

Table 5.3 shows a comparison of the key baseline characteristics for patients participating in the chart review and those participating in the questionnaire and SCID.

Table 5.3 Comparison of key baseline features for patients participating in chart review and those participating in questionnaires and SCID

Baseline characteristic		**Chart review (n=410)**	**Questionnaires + SCID (n=208)**
Age (SD, range) (years)		42.6 (13.4, 18-87)	43.3 (13.4, 18-76)
Female gender		80.5%	80.2%
Migraine type	Without aura	67.1%	63.5%
	With aura	32.9%	36.5%
Migraine frequency	Episodic	60.5%	65.9%
	Chronic	39.5%	34.1%
On antidepressants		34.0%	32.7%
On migraine preventatives		92.9%	93.0%

SD = standard deviation

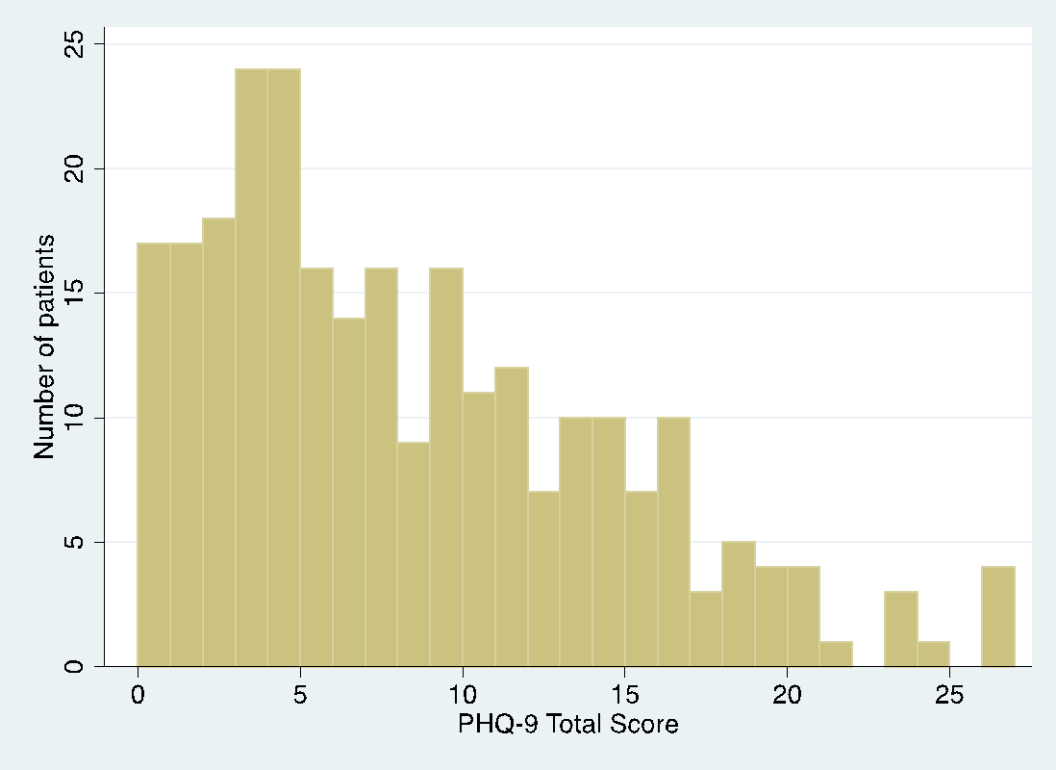
In comparison to the data from the chart review, no major differences were seen in any of the key baseline features, rendering statistical testing unnecessary in this case. The questionnaires and SCID group represents the most important group of patients from an analysis perspective, as this is the main group analyzed in regards to the objectives of the study. The data in Table 5.3 therefore indicates that these patients are representative of the overall eligible study population (i.e. the 208 patients are representative of the 410 deemed eligible).

5.2.2 Scales and Questionnaires

5.2.2.1 PHQ-9

Each of the 9 questions from the PHQ-9 is scored from 0 to 3. The total score is the sum of the scores for questions 1-9 with a range of 0 to 27. Five patients had missing information from questions 1-9. Therefore total scores would be inaccurate for those 5 patients. Therefore, those 5 patients were removed from the analysis, resulting in PHQ-9 scores being available for 263 patients, rather than 268 patients. The PHQ-9 total scores were right skewed and ranged from 0 to 27, with a mean score of 8, and a median score of 7. Figure 5.2 shows the distribution of the scores.

Figure 5.2 Distribution of PHQ-9 total scores.



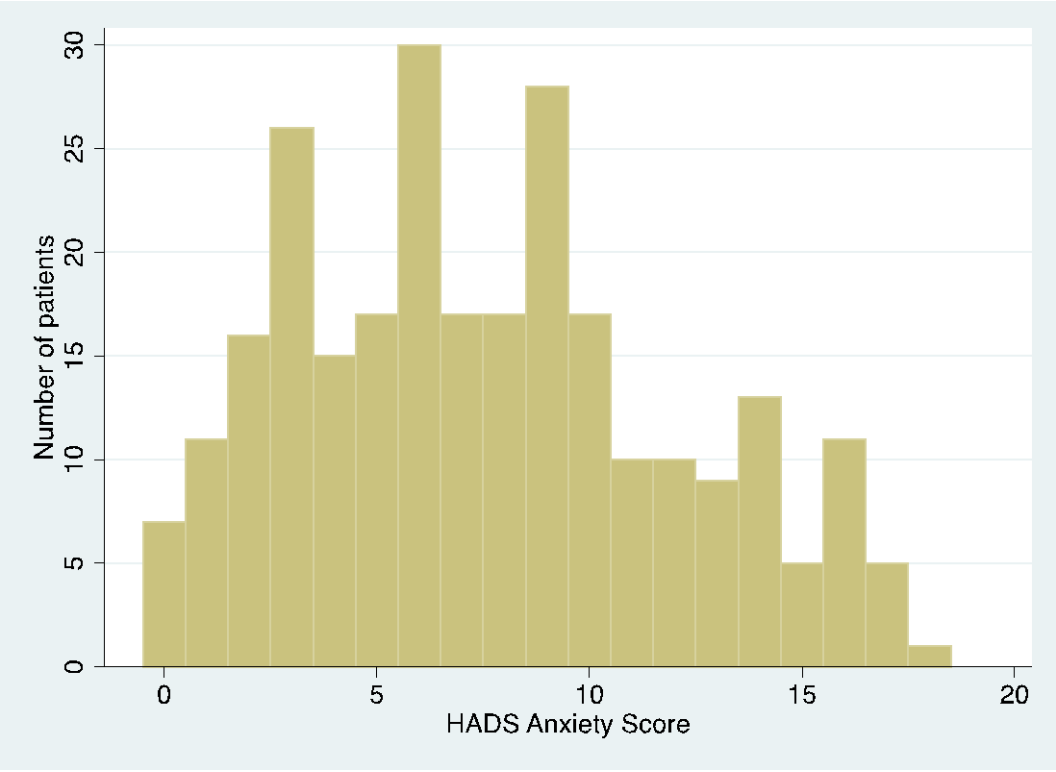
5.2.2.2 PHQ-9 Question 9

Question 9 of the PHQ-9 asks patients: do you have "Thoughts that you would be better off dead or of hurting yourself in some way". For this question, 29 patients (10.8%) responded that they had these thoughts on several days in the past two weeks, 7 patients (2.6%) said they had these thoughts more than half the days in the past two weeks and another 7 (2.6%) said they had these thoughts nearly every day.

5.2.2.3 HADS

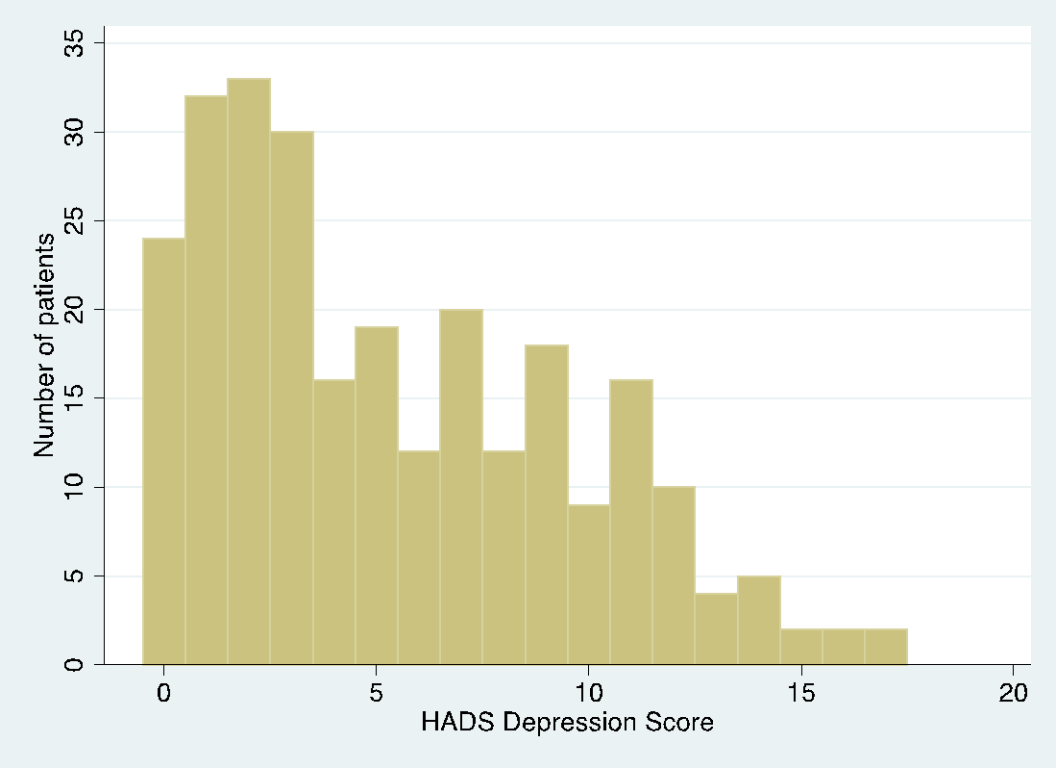
The HADS is composed of 14 questions, 7 of which contribute to the anxiety score (HADS-A) and 7 of which contribute to the depression score (HADS-D). The range of possible scores is 0-21 for each section. In the case of HADS-A, 3 patients had missing information and therefore were removed from the calculation of the total score, resulting in 265 patients. Similar to the PHQ-9, the HADS was right skewed. The mean score was 7.5, median score was 7.0, and range was 0-18. Figure 5.3 shows the distribution of scores for HADS-A.

Figure 5.3 Distribution of HADS anxiety scores



For the HADS depression scores, there were 2 patients with missing data, resulting in 266 patients. The mean score was 5.3, median score was 4.0, and range was 0-17. Figure 5.4 shows the distribution of scores for HADS-D.

Figure 5.4 Distribution of HADS Depression scores

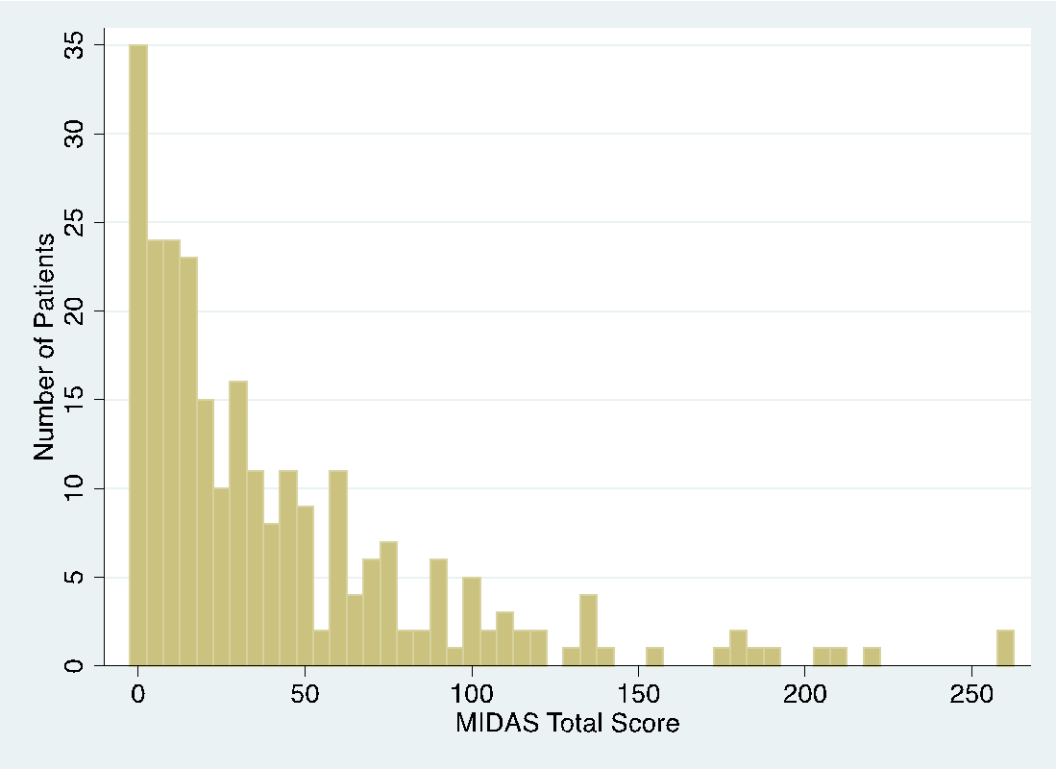


5.2.2.4 MIDAS

The MIDAS is scored by summation of the number of days from patients' responses for questions 1-5. The score can range from 0 to 270, as the max for each question is 90, but four questions of the five are paired into two sets of questions. Each pair of questions must add up to a maximum of 90 because one asks about missing work/school/etc. and one asks about reduced productivity at those activities. Therefore a patient must not count the days missed as part of the

response to the next question asking about days reduced in the same activities. Two patients had missing information for the MIDAS and therefore 266 patients took part in the total score. Of these, 8 patients had out of bounds values (scores above 270). The out of bounds values likely occurred because these patients counted the same days twice (both for missing those activities and reduced productivity at those activities). The data for these 8 patients were treated as missing data in order to preserve data quality. Therefore, a total of 258 patients were analyzed in this section. The MIDAS scores did not produce a normal distribution. Mean score was 42.2 and median score was 26.0. Range of scores was 0 to 261. The removal of the data for the patients with out of bounds values did not make a significant difference to the overall results, as a similar proportion of patients were in each MIDAS category. Figure 5.5 shows the distribution of scores for the MIDAS.

Figure 5.5 Distribution of MIDAS total scores



When categorized into degrees of disability, the following results were found: score of 0-5 (little or no disability): 18.6% of sample; score of 6-10 (mild disability): 9.7% of sample; score of 11-20 (moderate disability): 17.1% of sample; and score of 21+ (severe disability): 54.7% of sample. Therefore, more than half of patients had severe disability according to the MIDAS, reflecting the severe nature of disease in the migraine population seen at the Headache clinic in Calgary.

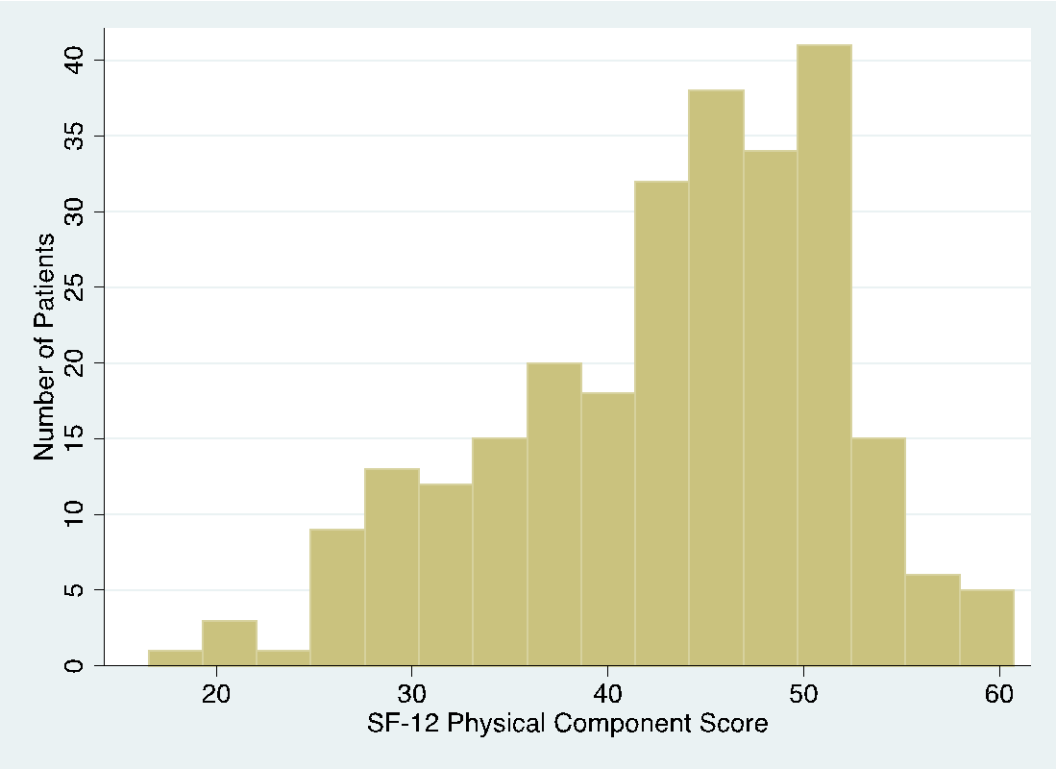
The MIDAS also has 3 other questions that are not counted towards the total score but yet provide useful information. The first question asks about the frequency of headache experienced by the patient in the last 3 months. In our sample, the mean number of headache days was 38.9 and median was 30, with a range of 0-90 days of headache. Of 261 patients who responded to

this portion, 97 patients (37%) had chronic migrane (≥ 15 days/month or about 45 or more days in past 3 months). Fifty-five patients (21%) had daily headache (90 days in last 3 months). The next question asks patients to rate the overall severity of their headaches over the last 3 months from a scale of 0-10. The mean headache intensity for the entire group of patients was 6.1, and median intensity was 6. An intensity level of 6 indicates moderate to severe pain that limits a patient's activities and some activities may be less of a priority [71]. The range was 0-10. The final question asks patients to rate the degree that headaches interfere with their life, from a scale of 0 to 10 (0=no interference, 10=severe interference). The range of scores was 0-10, mean was 5.9, and median was 6, i.e. a moderately severe level of interference to daily life.

5.2.2.5 SF-12

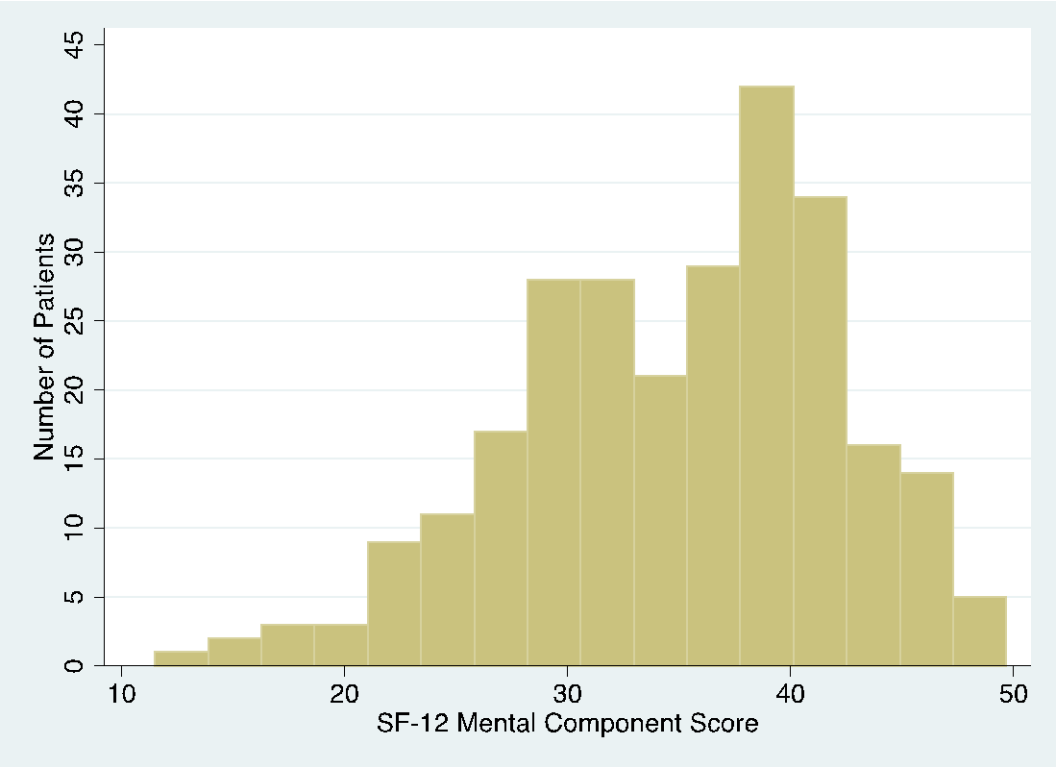
The Physical Component Score (PCS) and Mental Component Score (MCS) of the SF-12 are calculated based on the scores of the 12 questions, weighted as a utility scale [20]. Each has a range of 0 to 100. Zero indicates the lowest health-related quality of life and 100 indicates the highest health-related quality of life. The scores are calculated by a computer algorithm. Five patients had missing information and therefore a total score could not be calculated for those patients, resulting in 263 patients for this analysis. For the PCS, the scores ranged from 16.5 to 60.7. The mean was 43.1 and the median was 44.6. Figure 5.6 shows the distribution of scores for the PCS.

Figure 5.6 Distribution of SF-12 Physical Component Scores



The MCS, among the 263 patients, ranged from 11.5 to 49.7, with a mean of 35.0, and median of 36.3. Figure 5.7 shows the distribution of scores. The MCS was lower than the PCS, indicating that these migraine patients had a lower mental health related quality of life, as compared to physical health related quality of life.

Figure 5.7 Distribution of SF-12 Mental Component Scores



5.3 General data analysis

One of the initial comparisons to make was to determine if any significant differences were present between the group of patients that completed both the questionnaires and SCID and the group that completed the questionnaires only. Table 5.4 shows a number of variables assessed in these patients. Depending on the variable and missing values, the exact number of patients was slightly less than 208 for the questionnaire and SCID group, and 59 for the questionnaire only group.

Table 5.4 A comparison of factors between patients completing both the questionnaires and SCID, and those completing the questionnaires only

Variable	Questionnaire & SCID (n=208)	Questionnaire Only (n=59)	P-value
Mean age (95% CI) (years)	43.3 (41.5-45.1)	39.2 (35.9-42.5)	0.02*
Gender: % female	80.2	80.0	0.48$
Depression by self-report: % of patients	48.1	40.0	0.13$
PHQ-9 median score (95% CI)	7 (6-8)	5 (4-8)	0.34
PHQ-9 question 9 score 0, 1, 2, 3 (n)	172, 24, 6, 6	53, 5, 1, 1	0.88@
PHQ-9 algorithm % depressed (95% CI)	17.8 (12.5-23.0)	15.0 (6.0-24.0)	0.31$
HADS-A median score (95% CI)	7 (6-8)	8 (6-9)	0.92
HADS-D median score (95% CI)	4 (3-5)	5 (3-6)	0.80
MIDAS categories: no, mild, mod, severe disability (n)	43, 20, 36, 105	5, 5, 8, 36	0.16@
SF-12 PCS median score (95% CI)	44.2 (42.6-45.4)	46.3 (43.3-48.5)	0.24
SF-12 MCS median score (95% CI)	36.1 (34.1-37.0)	37.2 (33.1-39.0)	0.28

P values determined by Wilcoxon rank sum test, unless otherwise indicated: *Student's t-test, $2-sample test of proportions, @Fisher's exact test.

Table 5.4 shows a difference between mean age in the 2 groups, with a p value of 0.02. This is statistically significant but the difference in mean age is small between the groups. All other variables assessed show no significant difference between the two groups as indicated by the high p-values. Overall, therefore, the two groups look very similar.

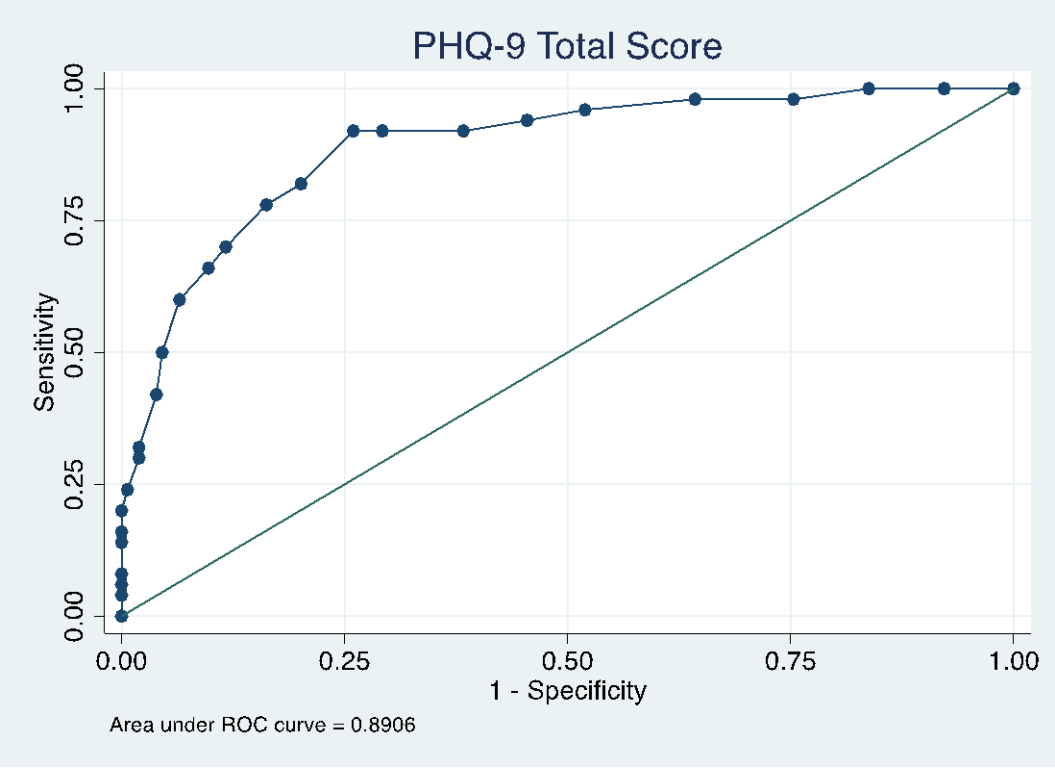
5.4 Data analysis specific to each objective

5.4.1 Objective 1: Determination of cut-points, ROC analysis and calculation of test psychometric properties

5.4.1.1 PHQ-9 total score

As the PHQ-9 total score had a right skewed distribution, a non-parametric method was used for the ROC analysis among the 204 patients who had data for this portion of the data analysis. Figure 5.8 shows the ROC curve. The area under the curve was 0.89 (95% CI 0.84-0.93).

Figure 5.8 ROC curve for the PHQ-9 total score



From the ROC curve, a detailed report of sensitivity and specificity for each cut-point can be generated, in addition to the percentage of patients correctly classified. This is shown in Table 5.5.

Table 5.5 Detailed report of sensitivity, specificity, and percentage of patients correctly classified for each cut-point of the PHQ-9 total score

Cutpoint	Sensitivity	Specificity	Correctly Classified
(>= 0)	100.0%	0.0%	24.5%
(>= 1)	100.0%	7.8%	30.4%
(>= 2)	100.0%	16.2%	36.8%
(>= 3)	98.0%	24.7%	42.7%
(>= 4)	98.0%	35.7%	51.0%
(>= 5)	96.0%	48.1%	59.8%
(>= 6)	94.0%	54.6%	64.2%
(>= 7)	92.0%	61.7%	69.1%
(>= 8)	92.0%	70.8%	76.0%
(>= 9)	92.0%	74.0%	78.4%
(>= 10)	82.0%	79.9%	80.4%
(>= 11)	78.0%	83.8%	82.4%
(>= 12)	70.0%	88.3%	83.8%
(>= 13)	66.0%	90.3%	84.3%
(>= 14)	60.0%	93.5%	85.3%
(>= 15)	50.0%	95.5%	84.3%
(>= 16)	42.0%	96.1%	82.8%
(>= 17)	32.0%	98.1%	81.9%
(>= 18)	30.0%	98.1%	81.4%
(>= 19)	24.0%	99.4%	80.9%
(>= 20)	20.0%	100.0%	80.4%
(>= 21)	16.0%	100.0%	79.4%
(>= 23)	14.0%	100.0%	78.9%
(>= 24)	8.0%	100.0%	77.5%
(>= 26)	6.0%	100.0%	77.0%
(>= 27)	4.0%	100.0%	76.5%
(> 27)	0.0%	100.0%	75.5%

Table 5.5 indicates that the highest percentage of correct classifications (85.3%) occurs at the 14 or greater cut-point. At this cut-point, the sensitivity is only 60.0% but the specificity is 93.5%. This is the cut-point felt to be the most appropriate for our sample. Please see the discussion section for the rationale in choosing this cut-point as the most suitable one. Essentially, the goal is to rule in depression and therefore a tool with higher specificity and PPV is important in this case.

Table 5.6 shows the two by two table for the PHQ-9 cut-point of 14 or greater with the SCID diagnosis of depression as gold standard (n=204), and Table 5.7 shows the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) associated with the PHQ-9 cut-point of 14 or greater. Fisher's exact test can be done to tell us if there is a significant difference between the proportions of patients diagnosed with depression using the SCID as compared to the PHQ-9. Fisher's exact test demonstrates a p-value of <0.0001 and therefore the difference between the proportion of depressed patients based on the SCID results as compared to the PHQ-9 is significant and extremely unlikely to be due to chance.

Table 5.6 Two by two table demonstrating the number of depressed and non-depressed patients according to the PHQ-9 cut-point of 14 or greater with a SCID diagnosis of depression

SCID	**PHQ-9 Not depressed**	**PHQ-9 Depressed**	**Total**
Not depressed	144	10	154
Depressed	20	30	50
Total	164	40	204

Table 5.7 Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) calculated from Table 5.6

Variable	Calculation	95% CI
Sensitivity	30/50 x 100 = 60.0%	45.2 – 73.6
Specificity	144/154 x 100 = 93.5%	88.4 – 96.8
PPV	30/40 x 100 = 75.0%	58.8 – 87.3
NPV	144/164 x 100 = 87.8%	81.8 – 92.4

The nature of misclassification can be explored further by analyzing the numbers in Table 5.6. We can see that 10 patients are false positive (scoring higher despite not being depressed) and 20 patients are false negative, scoring lower despite having depression. If one were to select a cut-point less than 14 for the PHQ-9, the sensitivity would be higher but specificity would be lower. This would result in a higher number of false positive patients. Selecting a higher cut-point would result in a higher specificity but very poor sensitivity, producing an unacceptable number of false negatives. Therefore, a cut-point of 14 or greater seems to produce the best balance.

5.4.1.2 PHQ-9 algorithm

The PHQ-9 algorithm produces a binary response, classifying patients as depressed or not depressed based on certain criteria (Appendix D). A two by two table with the SCID results as the gold standard can then be produced: Table 5.8 (n=208).

Table 5.8 Two by two table demonstrating the number of depressed and non-depressed patients according to the PHQ-9 algorithm with a SCID diagnosis of depression

SCID	PHQ-9 algorithm Not depressed	PHQ-9 algorithm Depressed	Total
Not depressed	148	8	156
Depressed	24	28	52
Total	172	36	208

Fisher's exact test generates a p value of <0.0001 and is therefore highly significant. This indicates that there is a significant difference between the number of patients diagnosed with depression for the SCID as compared to the PHQ-9 algorithm, and that this difference is extremely unlikely to be due to chance. Table 5.9 shows the sensitivity, specificity, PPV and NPV calculated from the two by two table shown in Table 5.8.

Table 5.9 Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) calculated from Table 5.8

Variable	Calculation	95% CI
Sensitivity	28/52 x 100 = 53.8%	39.5 - 67.8
Specificity	148/156 x 100 = 94.9%	90.1 – 97.8
PPV	28/36 x 100 = 77.7%	60.8 – 89.9
NPV	148/172 x 100 = 86.0%	80.0 – 90.9

5.4.1.3 PHQ-2

The PHQ-2 is the first 2 questions of the PHQ-9 with a score range of 0-6. Table 5.10 shows a two by two table with a cut-point of 2 or greater (n=207). Table 5.11 shows the

sensitivity, specificity, PPV and NPV with 95% confidence intervals for the two by two table shown in Table 5.10.

Table 5.10 Two by two table demonstrating the number of depressed and non-depressed patients according to the PHQ-2 cut-point of 2 or greater with a SCID diagnosis of depression

SCID	PHQ-2 Not depressed	PHQ-2 Depressed	Total
Not depressed	103	53	156
Depressed	6	45	51
Total	109	98	207

Table 5.11 Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) calculated from Table 5.10

Variable	Calculation	95% CI
Sensitivity	45/51 x 100 = 88.2%	76.1 – 95.6
Specificity	103/156 x 100 = 66.0%	58.0 – 73.4
PPV	45/98 x 100 = 45.9%	35.8 – 56.3
NPV	103/109 x 100 = 94.5%	88.4 – 98.0

Table 5.12 shows a two by two table with a cut-point of 3 or greater. Table 5.13 shows the sensitivity, specificity, PPV and NPV with 95% confidence intervals for the two by two table shown in Table 5.12.

Table 5.12 Two by two table demonstrating the number of depressed and non-depressed patients according to the PHQ-2 cut-point of 3 or greater with a SCID diagnosis of depression

SCID	PHQ-2 Not depressed	PHQ-2 Depressed	Total
Not depressed	140	16	156
Depressed	21	30	51
Total	161	46	207

Table 5.13 Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) calculated from Table 5.12

Variable	Calculation	95% CI
Sensitivity	30/51 x 100 = 58.8%	44.2 – 72.4
Specificity	140/156 x 100 = 89.7%	83.9 – 94.0
PPV	30/46 x 100 = 65.2%	49.8 – 78.6
NPV	140/161 x 100 = 87.0%	80.8 – 91.7

5.4.1.4 HADS

As the focus of this study is screening for depression, the HADS depression scores (HADS-D) were analyzed using ROC analysis and the anxiety scores were not assessed. As the HADS-D total score had a skewed distribution, a non-parametric method was used for the ROC analysis: Figure 5.9 (n=207) shows the ROC curve. The area under the curve was 0.92 (95% CI 0.88-0.96). Table 5.14 shows the details of sensitivity, specificity, and percentage of patients correctly classified for each cut-point of the HADS depression total score.

Figure 5.9 ROC curve for the HADS Depression total score

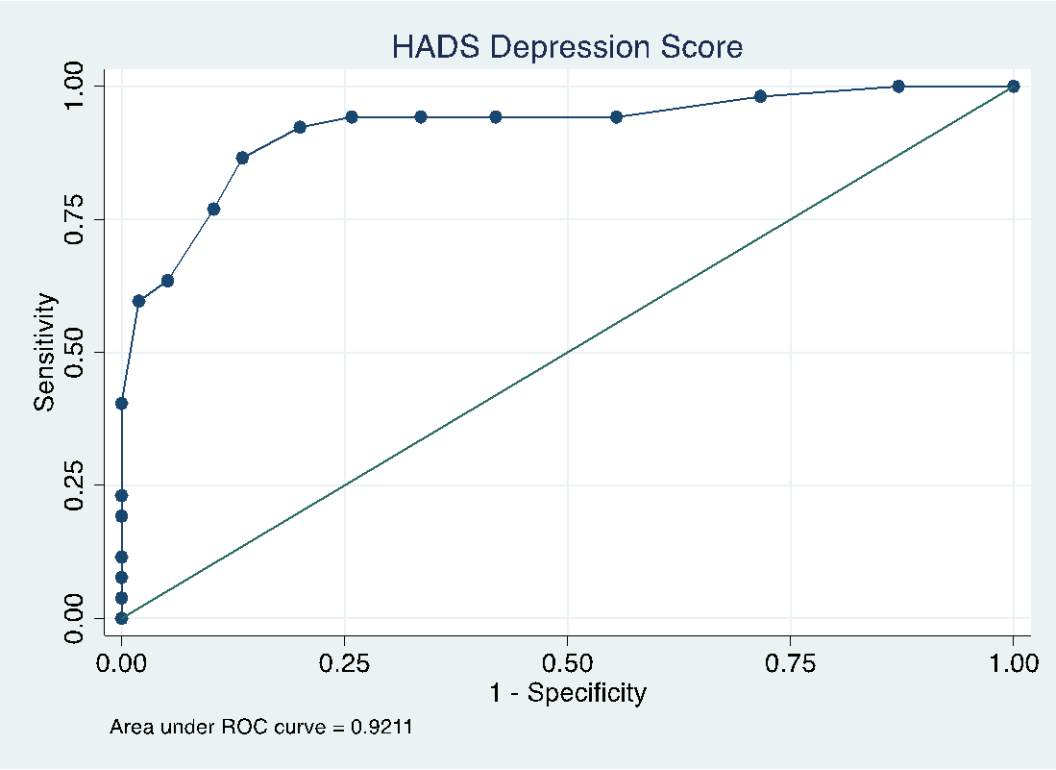


Table 5.14 Detailed report of sensitivity, specificity, and percentage of patients correctly classified for each cut-point of the HADS depression score

Cutpoint	Sensitivity	Specificity	Correctly Classified
(>=0)	100.0%	0.0%	25.1%
(>=1)	100.0%	12.9%	34.8%
(>=2)	98.1%	28.4%	45.9%
(>=3)	94.2%	44.5%	57.0%
(>=4)	94.2%	58.1%	67.2%
(>=5)	94.2%	66.5%	73.4%
(>=6)	94.2%	74.2%	79.2%
(>=7)	92.3%	80.0%	83.1%
(>=8)	86.5%	86.5%	86.5%
(>=9)	76.9%	89.7%	86.5%
(>=10)	63.5%	94.8%	87.0%
(>=11)	59.6%	98.1%	88.4%
(>=12)	40.4%	100.0%	85.0%
(>=13)	23.1%	100.0%	80.7%
(>=14)	19.2%	100.0%	79.7%
(>=15)	11.5%	100.0%	77.8%
(>=16)	7.7%	100.0%	76.8%
(>=17)	3.9%	100.0%	75.9%

For the same reasons described in the PHQ-9 section, a cut-point with higher specificity is important to rule in depression. However, sensitivity should not be extremely poor at the same time. So, for the HADS depression score, a cut-point of 11 or greater seems to have the best balance and does produce the highest percentage of correct classifications, with a sensitivity of 59.6% and a specificity of 98.1%. Table 5.15 shows the two by two table for the HADS-D cut-point of 11 or greater with the SCID diagnosis of depression as gold standard, and Table 5.16 shows the sensitivity, specificity, PPV and NPV associated with the HADS-D for a cut-point of

11 or greater. Fisher's exact test demonstrates a p-value of <0.0001 and therefore the difference between the proportion of depressed patients based on the SCID results as compared to the PHQ-9 is significant and extremely unlikely to be due to chance.

Table 5.15 Two by two table demonstrating the number of depressed and non-depressed patients according to the HADS-D cut-point of 11+ with a SCID diagnosis of depression

SCID	HADS Not Depressed	HADS Depressed	Total
Not Depressed	152	3	155
Depressed	21	31	52
Total	173	34	207

Table 5.16 Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) calculated from Table 5.15

Variable	Calculation	95% CI
Sensitivity	31/52 x 100 = 59.6%	45.1 - 73.0
Specificity	152/155 x 100 = 98.1%	94.4 – 99.6
PPV	31/34 x 100 = 91.2%	76.3 – 98.1
NPV	152/173 x 100 = 87.9%	82.0 – 92.3

The nature of misclassification can be explored for the HADS depression scores by taking a closer look at the numbers in Table 5.15. The results show 3 false positive patients (who scored 11) and 21 false negative patients. The 21 patients scored lower (scores were 1, 2, 6, 7, 8, 9, 10). If a cut-point of 10 or greater was used, the false negative rate would be slightly less (19 patients), but the false positive rate would be suddenly quite a bit higher (8 patients).

Therefore, a cut-point of 11 or greater for the HADS depressions score seems to provide the best balance.

5.4.1.5 Comparing results of the questionnaires

Table 5.17 provides a summary of the sensitivity, specificity, PPV and NPV of all of the questionnaires discussed above to allow for a comparison. The PHQ-2 cut-point of 3 or greater was used as it produced a much more appropriate specificity compared to a cut-point of 2 or greater. The PHQ-9 score and HADS-D score are also shown at other cut-points for comparison; the bold rows (PHQ-9 cut-point 14 and HADS-D cut-point 11) are the selected cut-points for our migraine sample.

Table 5.17 A comparison of the psychometric properties of the PHQ-9 score, PHQ-9 algorithm, PHQ-2, and HADS-D with the SCID diagnosis of depression

Test	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
PHQ-9 score Cut-point 10	82.0% (68.6-91.4)	79.9% (72.7-85.9)	56.9% (44.7-68.6)	93.2% (87.5-96.8)
PHQ-9 score Cut-point 11	78.0% (64.0-88.5)	83.8% (77.0-89.2)	60.9% (47.9-72.9)	92.1% (86.4-96.0)
PHQ-9 score Cut-point 12	70.0% (55.4-82.1)	88.3% (82.2-92.9)	66.0% (51.7-78.5)	90.1% (84.1-94.3)
PHQ-9 score Cut-point 13	66.0% (51.2-78.8)	90.3% (84.4-94.4)	68.8% (53.7-81.3)	89.1% (83.1-93.5)
PHQ-9 score Cut-point 14	**60.0% (45.2–73.6)**	**93.5% (88.4–96.8)**	**75.0% (58.8–87.3)**	**87.8% (81.8–92.4)**
PHQ-9 score Cut-point 15	50.0% (35.5-64.5)	95.5% (90.9-98.2)	78.1% (60.0-90.7)	85.5% (79.3-90.4)
PHQ-9 Algorithm	53.8% (39.5-67.8)	94.9% (90.1–97.8)	77.7% (60.8–89.9)	86.0% (80.0–90.9)
PHQ-2 score Cut-point 3	58.8% (44.2–72.4)	89.7% (83.9–94.0)	65.2% (49.8–78.6)	87.0% (80.8–91.7)
HADS-D score Cut-point 8	86.5% (74.2-94.4)	86.5% (80.0-91.4)	68.2% (55.6-79.1)	95.0% (90.0-98.0)
HADS-D score Cut-point 9	76.9% (63.2-87.5)	89.7% (83.8-94.0)	71.4% (57.8-82.7)	92.1% (86.5-95.8)
HADS-D score Cut-point 10	63.5% (49.0-76.4)	94.8% (90.1-97.7)	80.5% (65.1-91.2)	88.6% (82.7-93.0)
HADS-D score Cut-point 11	**59.6% (45.1-73.0)**	**98.1% (94.4 – 99.6)**	**91.2% (76.3–98.1)**	**87.9% (82.0–92.3)**

PPV = Positive predictive value, NPV = Negative predictive value
Bolded rows represent the selected cut-point for our migraine sample

As shown by the various cut-points of the PHQ-9 score and HADS-D score, as the cut-point increases, the sensitivity and NPV drop but the specificity and PPV rise. For the PHQ-9 cut-point of 15, the sensitivity becomes too low (50.0%) to be acceptable. Hence, for the PHQ-9, a cut-point of 14 provides the most optimal balance of the psychometric properties for our migraine sample. Similarly, for the HADS-D, a cut-point of 11 has the best specificity and PPV without a large compromise in sensitivity. Higher cut-points of the HADS-D are not shown in this table as the sensitivity becomes unacceptably low (<40%) as shown in Table 5.14.

When comparing the selected cut-points for the PHQ-9 and HADS-D with the PHQ-2 and PHQ-9 algorithm, the scales show similar sensitivity values except for the PHQ-9 algorithm, which shows a lower sensitivity. Among this group, the HADS-D shows the highest specificity, followed by the PHQ-9 algorithm and PHQ-9 score, which are similar. The PHQ-2 shows the lowest specificity. Positive predictive value is again highest for the HADS-D, lower for the PHQ-9 algorithm and score, and lowest for the PHQ-2 score. Negative predictive value is very similar for all scales in this group.

5.4.2 Objective 1: Stratified ROC analysis

A stratified ROC analysis was performed in this section to determine if the cut-points chosen above would be different for the PHQ-9 score and HADS depression score when subgroups of patients were assessed. The subgroups included males and females, age less than 43 years and equal to or above 43 years, episodic and chronic migraine, and migraine with aura and without aura. The age was stratified at 43 years because this was the mean age of the group and produced relatively equal numbers in each subgroup. The total number of patients analyzed

in this section was 208. The same rationale was used to determine the most appropriate cut-point in these subgroups as used for the PHQ-9 and HADS.

5.4.2.1 Gender

The group comprised of 80% females (n=167) and 20% males (n=41).

PHQ-9:

The most appropriate cut-point for males was the same as the overall sample at a score of 14 or above (92.7% correctly classified). However, males had a much better sensitivity at 85.7% and a slightly better specificity at 96.3% than the overall sample. These estimates were somewhat imprecise however with broad 95% confidence intervals (Table 5.18). For females, the most appropriate cut-point was again 14 (83.4% correctly classified). This produced a low sensitivity of only 50.0% and a specificity of 92.9%. Despite some imprecision in the values obtained, the PHQ-9 performed better in males than in females, having a higher sensitivity and specificity.

HADS:

The HADS depression score showed good test results when compared to the gold standard SCID for males. Here, a cut-point of 11 or greater had the highest percentage of correct classification (90.2%) with a sensitivity of 78.6% and specificity of 96.3%. For females, the test did not perform as well as males: a cut-point of 11 and greater had 88.0% correct classifications with only 52.6% sensitivity but very good specificity of 98.4%. Still, this cut-point was the most appropriate in females.

5.4.2.2 Age less than 43 years or 43 years and greater

About 51% of patients were less than 43 years (n=107) and 49% were 43 years or greater (n=101).

PHQ-9:

For age less than 43 years, a cut-point of 14 or greater appeared to have the best balance of sensitivity and specificity (83.0% correct classifications), with sensitivity of 57.7% and specificity of 91.3%. For age 43 or greater, cut-points 13 and 15 produced a similar percentage of correct classifications (88.8%) but cut-point 13 had higher sensitivity and cut-point 15 had higher specificity. Since specificity is considered more important (to rule in a diagnosis of depression), cut-point 15 was felt to be more appropriate for this group of patients: sensitivity 62.5%, and specificity 97.3%. The age 43 or greater group showed a higher sensitivity and specificity for the PHQ-9 score as compared to those less than 43 years old.

HADS:

For patients less than 43 years old, a cut-point of 11 or greater provided about 85% correct classifications, with a sensitivity of only 51.9% but a specificity of 96.3%. For patients 43 and older, a different cut-point appeared to produce good results. A cut-point of 8 or greater had the highest percentage of correct classifications (93.0%), with 92.0% sensitivity and 93.3% specificity. The results were overall better in the older age group as compared to the younger age group.

5.4.2.3 Episodic and Chronic migraine

Nearly 64% of patients had episodic migraine (n=133) and about 36% had chronic migraine (n=75).

PHQ-9:

For episodic migraine, it appeared that a number of cut-points had a similar degree of correct classifications (84.7%): 10, 15, 19, and 20. However the higher the cut-point, the lower the sensitivity. Therefore, in order to get a better than chance sensitivity, a cut-point that had at least over 50% sensitivity would be appropriate. A cut-point of 14 or greater was felt to be most optimal despite the poor sensitivity (52.0%), given a reasonable specificity (91.5%). For chronic migraine patients, the PHQ-9 showed better results, despite some imprecision in the estimates. A cut-point of 14 or greater showed a higher percentage of correct classifications (87.7%) than for episodic migraine patients. The sensitivity and specificity were also higher (68.0% and 97.9% respectively).

HADS:

For patients with episodic migraine, a cut-point of 11 or greater produced 91.0% correct classifications, with a poor sensitivity of 56.0% and excellent specificity of 99.1%. In chronic migraine patients, a cut-point of 11 or greater had the highest percentage of correct classifications (83.8%), with a sensitivity of 63.0% and a specificity of 95.7%. Although the cut-point was similar to that of episodic migraine, the test overall did not perform as well in chronic migraine patients (lower percentage of correct classifications and lower specificity).

5.4.2.4 Migraine with aura or without aura

About 36% of patients had migraine with aura (n=75) and the rest had migraine without aura (n=133).

PHQ-9:

For patients with migraine with aura, a cut-point of 13 or greater appeared to have the highest degree of correct classifications (86.3%). This cut-point produced a sensitivity of 66.7% and a specificity of 92.7%. In patients with migraine without aura, the results were similar to the overall sample. A cut-point of 14 or greater appeared best (85.5% correct classification), with a sensitivity of 59.4% and a specificity of 93.9%.

HADS:

In this case, in migraine with aura patients, the best cut-point seemed to be at 9 or greater, with 90.7% correctly classified and a sensitivity of 84.2% and specificity of 92.9%. For patients with migraine without aura, a cut-point of 11 or greater had the highest percentage of correct classifications (87.9%) with 60.61% sensitivity and 97.0% specificity. The cut-point for migraine without aura (11) was therefore different than that of migraine with aura (9).

5.4.2.5 Summary

Table 5.18 summarizes the sensitivities and specificities discussed above for the PHQ-9 and HADS depression scales. Both questionnaires performed much better in males than in females, and the PHQ-9 performed better than the HADS for the males. The cut-points felt to be most appropriate were the same as the overall sample. The age 43 or greater group showed a higher sensitivity and specificity for the PHQ-9 score as compared to those less than 43 years old. The cut-point was the same as the overall sample for those less than 43 years, but a cut-point of 15 or greater was felt to be more appropriate for those aged 43 or greater. For the HADS, the older age group performed better at a cut-point of 8 or greater and the younger group showed a similar cut-point to the overall sample. For the older age group, the HADS performed

better than the PHQ-9. For the PHQ-9, chronic migraine patients performed better than episodic ones, with cut-points being the same as the overall sample. In contrast, for the HADS, episodic migraine patients performed better than the chronic ones, with cut-points again being the same as the overall sample. Neither test was overall superior to the other. For patients with migraine with or without aura, no major differences were seen in the performance of the PHQ-9 for each subgroup. For patients with migraine with aura, a cut-point of 13 or greater gave better results. For migraine without aura patients, the cut-point was the same as the overall sample. For the HADS, patients who had migraine with aura performed better and the best cut-point was 9 or greater. The HADS performed better than the PHQ-9 for these patients. The migraine without aura group had the most appropriate cut-point at 11 or greater for the HADS, similar to the overall sample.

Table 5.18 Sensitivities and specificities for the stratified analysis of PHQ-9 and HADS depression scales

Stratified Variable	PHQ-9 Sensitivity (95% CI)	PHQ-9 Specificity (95% CI)	HADS Sensitivity (95% CI)	HADS Specificity (95% CI)
Male	85.7% (57.2-98.2)	96.3% (81.0-99.9)	78.6% (49.2-95.3)	96.3% (81.0-99.9)
Female	50.0% (33.4-66.6)	92.9% (87.2-96.8)	52.6% (35.8-69.0)	98.4% (94.5-99.8)
Age < 43 years	57.7% (38.8-77.6)	91.3% (82.8-96.4)	51.9% (31.9-71.3)	96.3% (89.4-99.2)
Age ≥ 43 years	62.5% [15]* (42.5-82.0)	97.3% [15]* (90.8-99.7)	92.0% [8]* (74.0-99.0)	93.3% [8]* (85.3-97.8)
Episodic migraine	52.0% (31.3-72.2)	91.5% (84.8-96.1)	56.0% (34.9-75.6)	99.1% (94.9-100.0)
Chronic migraine	68.0% (46.0-83.5)	97.9% (93.5-99.8)	63.0% (42.4-80.6)	95.7% (89.5-98.5)
Migraine with aura	66.7% [13]* (43.4-87.4)	92.7% [13]* (82.7-98.0)	84.2% [9]* (60.4-96.6)	92.9% [9]* (82.7-98.0)
Migraine without aura	59.4% (42.1-77.1)	93.9% (87.4-97.8)	60.6% (42.1-77.1)	97.0% (91.5-99.4)

*Numbers in these [brackets] represent the cut-point deemed to be most appropriate. Otherwise, the cut-point is the same as the overall sample (14 for PHQ-9, 11 for HADS).

Overall therefore, the questionnaires performed better in males, those 43 years of age or older, chronic migraine patients and migraine with aura patients, than their counterparts. Males performed better with the PHQ-9 at a cut-point of 14 or greater. The older age group performed better with the HADS, at a cut-point of 8 or greater. Chronic migraine patients performed better

on the PHQ-9 at a cut-point of 14 or greater. Migraine with aura patients performed better with the HADS, at a cut-point of 9 or greater. Therefore, cut-points were different for only two groups, those aged 43 or older, and patients who had migraine with aura. All other cut-points were the same as the overall sample of patients.

5.4.3 Objective 2: Prevalence of depression

A total of 212 patients completed the SCID. Fifty-three of 212 patients or 25.0% met criteria for the diagnosis of major depressive disorder according to the gold standard SCID results. From the 212 patients, 70 patients (33%) were identified by the SCID to have had a past major depressive episode, with age of onset varying from 5 years of age to 62 years of age. Of these 70 patients, 29 (nearly 38%) had one episode in the past. Of the 53 currently depressed patients, 39 (74%) had previous episodes of major depression, 14 of them (36%) having one prior episode.

For patients who did not complete the SCID and only completed the questionnaires (n=59), the point prevalence of depression was determined using the PHQ-9 score at a cut-point of 14 or greater and the HADS depression score at a cut-point of 11 or greater (same cut-points as overall sample). For comparison, the group who had completed the SCID and questionnaires was also assessed using the two questionnaires.

For the PHQ-9, the point prevalence of depression for those only completing the questionnaires was 18.6% (95% CI 9.7-30.9). This estimate is slightly lower than the SCID, and the 95% CI's are very wide. The group completing the SCID and questionnaires had a similar point prevalence: 20.1% (95% CI 14.8-26.3). The 95% CI's are wide, but narrower than the questionnaire-only group, given a larger sample size and therefore a smaller random error.

For the HADS test, the patients completing the questionnaires only, showed a depression prevalence of 10.2% (95% CI 3.8-20.8), and those completing both the questionnaires and SCID showed a prevalence of 16.9% (95% CI 12.1-22.7). Again, 95% CI's are wider for the first group (lower sample size, lower precision). Although visually the two estimates look different, there is quite a bit of overlap of the 95% CI's. The prevalence estimates with HADS were lower than the PHQ-9 and the SCID.

5.4.4 Objective 2: Stratified prevalence estimates using SCID

In this section, a number of variables were assessed to determine if prevalence of depression would be different for various groups of patients. Stratification of the following baseline variables was performed: gender, age groups, migraine type, migraine frequency, and on migraine preventatives. Stratification of depression-related variables was also assessed, including whether or not the patient was: on antidepressant medications, being treated by any method for depression, and diagnosed with depression by a professional. Comparisons of subgroups were made by 2-sample test of proportions. Tables 5.19 and 5.20 show the stratified variables, the prevalence estimates, and the 95% CI's for each subgroup of patients. The p-value for assessment of statistical significance is also shown, among a total of 208 patients analyzed.

Table 5.19 Point prevalence estimates of depression using the results of the SCID, stratified by baseline characteristics

Stratified variable	Point prevalence	95% CI	P-value
Male (n=41)	34.1%	19.6 - 48.7	0.066
Female (n=167)	22.8%	16.4 – 29.1	
Age < 43 years (n=107)	25.2%	17.0 - 33.5	0.47
Age ≥ 43 years (n=101)	24.8%	16.3 – 33.2	
Migraine with aura (n=75)	25.3%	15.5 – 35.2	0.47
Migraine without aura (n=133)	24.8%	17.5 – 32.2	
Episodic migraine (n=133)	18.8%	12.2 – 25.4	**0.003**
Chronic migraine (n=75)	36.0%	25.1 – 46.9	
On migraine preventative (n=194)	25.8%	19.6 – 31.9	0.17
Not on migraine preventative (n=14)	14.3%	0.0 – 32.6	

As shown in Table 5.19, gender, age, migraine type, and being on a migraine preventative did not significantly influence prevalence estimates for depression. The mean age for depressed patients (43.8 years, 95% CI 41.6-46.0) and non-depressed patients (40.7 years, 95% CI 37.4-44.0) were also compared and showed no significant difference (p=0.074). However, having chronic migraine produced a significantly different result for the prevalence of depression as compared to episodic migraine. The episodic group had a depression prevalence of 18.8% and the chronic migraine group had a prevalence of 36.0%. This was statistically significant with a p value of 0.003. In regards to the use of migraine preventatives, only 14 of 208 patients were not on a preventative, i.e. 6.7%. The majority (101 patients or 49%) were

taking one preventative, while 94 patients (45%) were on 2 or more preventatives. No significant difference was seen in the prevalence of depression however. This may have been because of the very small number of patients in the group not using preventatives, as seen by the very wide 95% CI's.

Table 5.20 Point prevalence estimates of depression using the results of the SCID, stratified by depression variables

Stratified variable	Point prevalence	95% CI	P-value
On antidepressants (n=59)	42.4%	29.8 – 55.0	**0.0001**
Not on antidepressants (n=149)	18.1%	11.9 – 24.3	
Treatment for depression (n=67)	41.8%	30.0 – 53.6	**0.0001**
No treatment for depression (n=141)	17.0%	10.8 – 23.2	
Diagnosed with depression by health professional (n=98)	39.8%	30.1 – 49.5	**<0.00001**
Not diagnosed with depression by health professional (n=109)	11.0%	5.1 – 16.9	

Treatment includes medications, counselling, cognitive behavioural therapy, group therapy & other therapies

Of 208 patients, 59 (28.4%) were on antidepressant medications. Fourteen patients (6.7%) were on two agents and two patients (about 1%) were on three different agents. The prevalence of depression was 18.1% in the group not taking antidepressants, and nearly 42% in those taking antidepressants. The difference was highly significant (p=0.0001). Similar depression prevalence estimates and significance were found for the group of patients treated for depression by any modality and not treated by any modality, respectively. In the group not

treated for depression, 24 patients (17.0%) were diagnosed with depression. Therefore, the prevalence of untreated depression was 17.0% in this study sample. Finally, the diagnosis of depression by a health professional also made a difference. Those that were previously diagnosed by a health professional had a current depression prevalence of nearly 40%, while those not previously diagnosed by a health professional had a depression prevalence of 11%. The difference was significant ($p<0.00001$).

Another way to look at the depression variables is to assess the number of patients on antidepressants, on any depression treatment, and diagnosed with depression by a professional, among those diagnosed with depression. Of the 208 patients, 52 (25%) had a SCID diagnosis of depression. Of these, 25 patients (about 48%) were on antidepressants, indicating that over half of these depressed patients remained pharmacologically untreated. For any modality of depression treatment (including antidepressants), 28 patients (about 54%) were on treatment. Therefore, the proportion of patients with depression on no treatment was 46% (100 – 54%) in this study sample. In regards to a diagnosis of depression by a professional, 51 of the 52 SCID depressed patients had a response available. Of the 51 patients, 39 (76%) mentioned being diagnosed with depression by a professional.

It is worthwhile taking a closer look at the patients diagnosed with depression using the SCID who were not taking antidepressants, i.e. the 18.1% or 27 patients from the group of 149 not on an antidepressant. One can look at the sensitivity of the PHQ-9 and HADS in these patients to determine how many would have been recognized as depressed. Of the 27 patients, 26 had PHQ-9 and HADS depression scores and given the low number, an ROC curve or specificity could not be generated. Using a cut-point of 14 or greater for the PHQ-9, the sensitivity was 46.2% (12/26 x 100, 95% CI 26.6-66.6). The 95% CI was quite wide, but

allowed one to identify some of the patients who were depressed and were untreated. Using a cut-point of 11 or greater for the HADS, the sensitivity was 48.1% (13/27 x 100, 95% CI 28.7-68.1). The results were quite similar with the HADS.

5.4.5 Comparison of questionnaire scores in various patient groups

Table 5.21 shows the median scores for the PHQ-9 and HADS depression scales among various groups of patients. Median scores were reported as the PHQ-9 and HADS depression scores did not have the classical normal distribution but still had a bell-shaped curve and a central tendency. Comparisons for each pair of variables were made using the Wilcoxon rank sum test. For this section, the PHQ-9 included 204 patients and the HADS included 207 patients.

Table 5.21 PHQ-9 and HADS depression median scores and associated p-values for various groups of patients

Stratified Variable (n for PHQ-9, n for HADS-D)	PHQ-9 median score (95% CI)	p-value	HADS-D median score (95% CI)	p-value
Male (n=41, 41)	9.0 (5.0-13.0)	0.12	7.0 (3.0-9.0)	0.077
Female (n=163, 166)	7.0 (5.0-8.0)		4.0 (3.0-5.0)	
Age < 43 years (n=106, 107)	7.5 (7.0-9.6)	0.13	5.0 (4.0-7.0)	0.20
Age ≥ 43 years (n=98, 100)	5.0 (4.0-7.0)		3.0 (2.7-5.0)	
Migraine with aura (n=73, 75)	7.0 (5.6-9.0)	0.53	4.0 (3.0-6.0)	0.66
Migraine without aura (n=131, 132)	6.0 (4.0-8.2)		5.0 (3.0-6.0)	
Episodic migraine (n=131, 133)	5.0 (4.0-7.0)	**0.0025**	3.0 (3.0-4.0)	**0.001**
Chronic migraine (n=73, 74)	9.0 (7.6-11.0)		7.0 (5.1-9.0)	
SCID – Depressed (n=50, 52)	14.5 (13.0-16.0)	**<0.0001**	11.0 (9.4-12.0)	**<0.0001**
SCID – Not depressed (n=154, 155)	5.0 (4.0-6.0)		3.0 (2.0-3.2)	
Depression by self-report (n=97, 98)	10.0 (8.3-12.0)	**<0.0001**	7.0 (6.0-9.0)	**<0.0001**
No depression by self-report (n=106, 108)	4.0 (4.0-6.0)		2.0 (2.0-3.0)	
Treated for depression (n=66, 67)	11.0 (9.0-13.0)	**<0.0001**	8.0 (6.0-9.0)	**<0.0001**
Not treated for depression (n=138, 140)	5.5 (4.0-6.0)		3.0 (2.0-4.0)	
On antidepressants (n=58, 59)	11.0 (9.0-13.9)	**<0.0001**	9.0 (6.0-9.6)	**<0.0001**
Not on antidepressants (n=146, 148)	5.5 (4.0-6.4)		3.0 (2.0-4.0)	
On migraine preventatives (n=190, 193)	7.0 (5.0-9.0)	0.46	4.0 (3.0-6.0)	0.49
Not on migraine preventatives (n=14, 14)	6.0 (1.8-9.7)		3.0 (1.8-7.2)	

Table 5.21 indicates that the PHQ-9 showed similar results to the HADS, both indicating that those with chronic migraine, SCID diagnosis of depression, reporting depression, on any treatment for depression, and on antidepressants, had higher scores on the depression scales, and these were statistically significant compared to their counterparts.

5.4.6 Objective 3: Disability and quality of life

Please see the sub-section of scales and questionnaires under the descriptive analysis section above for the distribution of MIDAS and SF-12 scores. Further analysis was done in this part to determine if any differences exist between the scores for various groups of patients. These are the same patient groups as the last portion of objective two above and the data is shown in Tables 5.22 and 5.23. Table 5.22 shows data related to the MIDAS among 204 patients. One patient had a score that was out of bounds (score of 300 when max score is 270) and was therefore not used in the analysis to maintain data quality. As the MIDAS produced a non-normal distribution of scores and had no central tendency, the number of patients in each category was assessed and comparisons were made using Fisher's exact test. Table 5.23 shows data related to the SF-12 (n=205). The physical and mental component scores are shown separately. SF-12 scores did not have a classical normal distribution but did demonstrate a bell-shaped curve and a central tendency. Therefore median scores were reported for each variable and comparisons were made using the Wilcoxon rank sum test.

Table 5.22 Number of patients in each MIDAS category and associated p-value for various groups of patients

Stratified Variable	MIDAS disability category (counts)				p-value
	No	Mild	Mod	Severe	
Male (n=41)	12	1	9	19	0.17
Female (n=163)	32	19	28	84	
Age < 43 years (n=105)	8	10	24	63	**<0.0001**
Age ≥ 43 years (n=99)	36	10	13	40	
Migraine with aura (n=73)	13	10	12	38	0.44
Migraine without aura (n=131)	31	10	25	65	
Episodic migraine (n=131)	29	14	29	59	0.13
Chronic migraine (n=73)	15	6	8	44	
SCID – Depressed (n=50)	9	0	5	36	**0.001**
SCID – Not depressed (n=154)	35	20	32	67	
Depression by self-report (n=96)	10	7	18	61	**0.001**
No depression by self-report (n=107)	33	13	19	42	
Treated for depression (n=65)	5	5	10	45	**0.001**
Not treated for depression (n=139)	39	15	27	58	
On antidepressants (n=57)	5	5	8	39	**0.006**
Not on antidepressants (n=147)	39	15	29	64	
On migraine preventatives (n=190)	39	19	35	97	0.64
Not on migraine preventatives (n=14)	5	1	2	6	

The MIDAS demonstrated statistically significant differences in the number of patients in each category of disability for the following variables: age less than 43 years, diagnosis of

depression, depression by self-report, treatment for depression, and taking an antidepressant. These groups had a higher proportion of patients in the moderate and severe disability categories as compared to their counterparts. Surprisingly, chronic migraine patients showed no difference in their disability categories as compared to episodic migraine patients.

Table 5.23 SF-12 PCS and MCS median scores and associated p-values for various groups of patients

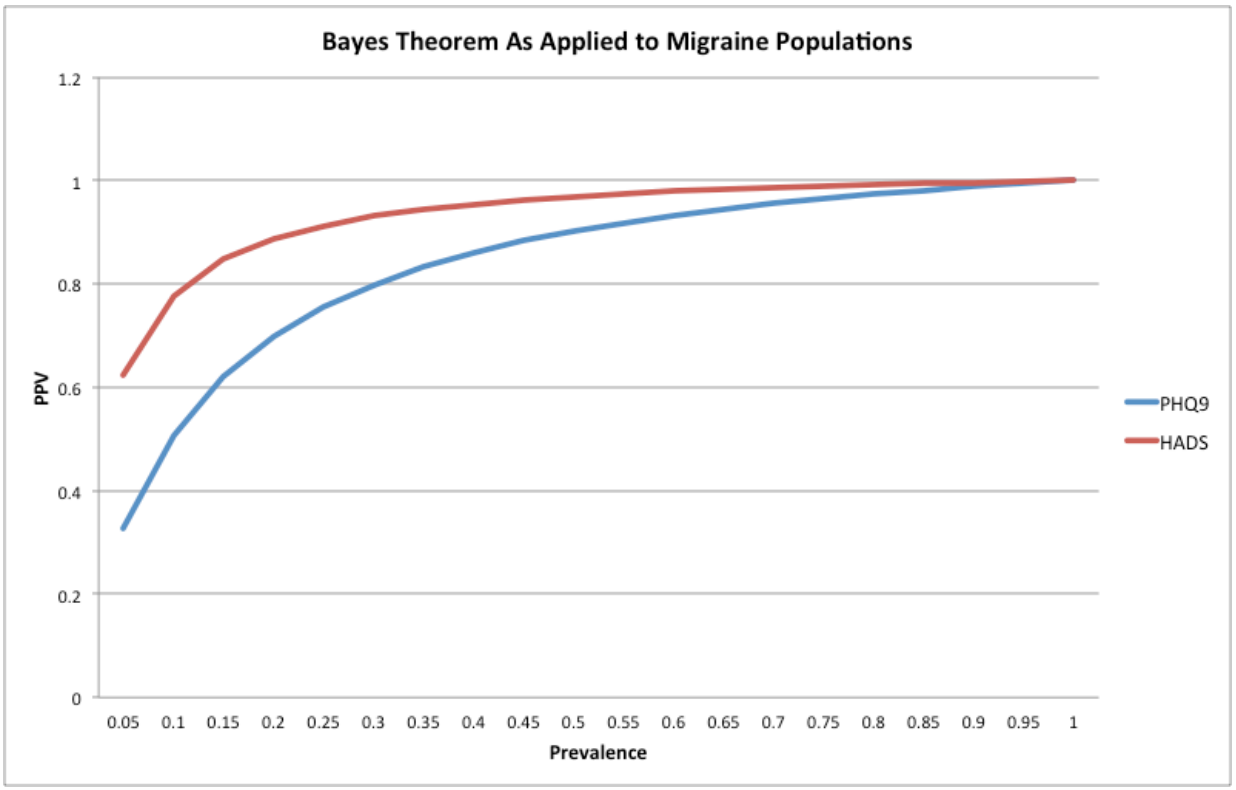
Stratified Variable	PCS median score (95% CI)	p-value	MCS median score (95% CI)	p-value
Male (n=40)	46.2 (42.1-49.9)	0.29	33.0 (29.7-37.9)	0.17
Female (n=165)	44.2 (42.6-45.3)		36.6 (35.1-37.7)	
Age < 43 years (n=106)	44.3 (42.9-45.9)	0.57	34.4 (32.1-37.0)	0.11
Age ≥ 43 years (n=99)	44.6 (42.0-46.9)		36.6 (35.5-39.2)	
Migraine with aura (n=74)	43.8 (41.8-46.0)	0.71	35.4 (32.1-37.8)	0.66
Migraine without aura (n=131)	44.9 (42.6-46.1)		36.6 (34.3-37.8)	
Episodic migraine (n=131)	45.4 (44.1-47.0)	**0.0001**	36.6 (34.4-37.9)	0.26
Chronic migraine (n=74)	40.0 (37.0-44.6)		35.2 (32.4-37.0)	
SCID – Depressed (n=50)	35.2 (31.8-37.9)	**<0.00001**	28.7 (26.4-30.3)	**<0.00001**
SCID – Not depressed (n=155)	46.0 (44.9-47.5)		37.8 (36.4-39.0)	
Depression by self-report (n=95)	42.6 (38.0-44.1)	**<0.00001**	32.9 (30.3-35.3)	**<0.00001**
No depression by self-report (n=109)	46.8 (44.6-48.0)		38.6 (37.1-39.7)	
Treated for depression (n=64)	41.9 (36.8-43.9)	**0.0001**	33.2 (30.3-35.7)	**0.002**
Not treated for depression (n=141)	45.5 (44.2-47.3)		37.5 (35.8-38.9)	
On antidepressants (n=56)	39.8 (36.0-43.9)	**0.0001**	33.2 (30.3-35.6)	**0.004**
Not on antidepressants (n=149)	45.5 (44.2-47.0)		37.0 (35.8-38.4)	
On migraine preventatives (n=191)	44.2 (42.5-45.4)	0.083	36.0 (33.8-37.5)	0.99
Not on migraine preventatives (n=14)	49.4 (43.3-52.4)		36.6 (31.5-38.6)	

In regards to the SF-12, statistically significant differences in the scores for PCS and MCS were seen for a SCID diagnosis of depression, depression by self-report, treatment for depression by any modality, and treatment for depression with an antidepressant. These groups showed a lower quality of life both physically and mentally compared to their counterparts. Interestingly, the PCS also showed a significant difference in scores for patients with episodic and chronic migraine, with the latter group having lower scores and therefore a lower physical quality of life. This difference was not seen for the mental component score, indicating no difference for mental quality of life for episodic versus chronic migraine patients.

5.5 Further applying the results

Bayes theorem can be used to predict the positive predictive value (PPV) in other populations with a different prevalence rate of depression. The formula is: PPV = (Sens*P)/[(Sens*P) + (1 – Spec)*(1-P)], where Sens is sensitivity, P is prevalence and Spec is specificity. In this case, a PPV can be calculated for the PHQ-9 and HADS depression scales, at various prevalence estimates. Based on the prevalence, once can then calculate a PPV and generate a curve for that tool, as shown in Figure 5.10. The selected cut-points of 14 or greater and 11 or greater for the PHQ-9 and HADS-D were used respectively. The sensitivity used for the formula therefore was 60.0% and specificity was 93.5% for the PHQ-9. The sensitivity was 59.6% and specificity was 98.1% for the HADS-D.

Figure 5.10 Application of Bayes Theorem to migraine populations with variable prevalence estimates of depression and resulting positive predictive values (PPV)



As shown in Figure 5.10, PPV rises as the prevalence of depression rises but starts to plateau at about 0.5 or 50% prevalence. The PPV rises most dramatically between 0.05 and 0.25 for the prevalence. Overall, the HADS (upper line) demonstrates a better PPV for the same prevalence estimates as compared to the PHQ-9. However, both curves converge at higher prevalence estimates.

Chapter 6: Discussion And Conclusions

6.1 Summary of key study findings

The purpose of this cross-sectional study was to assess how well the PHQ-9 and HADS depression screening tools perform in migraine patients, determine the prevalence of depression in this patient population using a gold standard semi-structured psychiatric interview, and examine outcomes for these patients, namely disability and quality of life. The three key findings of this study were:

1. With the intent of ruling in the diagnosis of depression, the PHQ-9 at a cut-point of 14 and the HADS depression scale at a cut-point of 11 produced the greatest proportion of correct classifications and an optimal balance of psychometric properties for the studied migraine population.
2. The point prevalence of depression in this study was found to be 25.0% (95% CI 19.0-31.0), and the prevalence of untreated depression was found to be 17.0% (95% CI 10.8 – 23.2).
3. Patients with both migraine and depression had significantly higher degrees of disability and a poorer quality of life as compared to patients without depression.

6.2 Introduction

Four hundred and ten patients were eligible for the study and a chart review was performed on these patients. Three hundred patients were then presented with the consent form and 268 (93.3%) completed questionnaires and 209 (69.7%) completed both the questionnaires and SCID. The participation rate was therefore very good for the questionnaires and fair for the questionnaires and SCID.

The degree of missing data or out of bounds data was minimal in this study and therefore unlikely to influence overall results. In addition, the baseline features of the patients participating in the chart review were comparable to the group completing the questionnaires and SCID. Therefore, the sample of migraine patients forming the core group for study analysis (with very little missing data) was felt to be representative of the eligible study population. Comparisons of this core group and those completing the questionnaires only (n=59) also revealed similar characteristics both demographically, as well as on responses to questionnaires.

6.3 Objective 1: Determination of cut-points, ROC analysis and calculation of test psychometric properties

The ROC curve for the PHQ-9 demonstrated an area under the curve of 0.89 (95% CI 0.84-0.93). This indicates quite a good overall performance of the PHQ-9 as a screening tool for depression in our migraine sample. The ROC analysis for the PHQ-9 total score demonstrated a sensitivity of 60.0% and a specificity of 93.5% at a cut-point of 14 or greater. This was felt to be the most appropriate cut-point for the overall migraine population for a number of reasons. First, in a screening tool for a chronic condition such as depression, one would want to rule in a diagnosis of depression, making specificity very important [34]. A tool with higher sensitivity

but lower specificity would produce more false positive cases, and thus would likely be less appropriate for screening in this population.

Second, there is substantial cost associated with inappropriate referrals to psychiatry and mental health services in the current Canadian health climate. The health system in Canada is already strained for resources and it can take months for patients to access specialty care. In addition, placing inappropriate patients on the wait list for psychiatric assessment only makes the problem worse by delaying access to those who truly need it. This goes back therefore to the concept of reducing the number of patients who are false positive on screening tests. A test with higher specificity is better suited to the current situation, even if it may come at the cost of sensitivity.

Third, there is significant stigma with the diagnosis of a mental health disorder, not to mention anxiety and distress to the patient personally. As such, patients falsely diagnosed with depression on screening tests may experience these emotions and the stigma unnecessarily. Furthermore, these patients may be prescribed antidepressants inappropriately. In such a case, physicians employing non-specific measurement strategies are at risk of breaking one of the basic ethical principles of medicine: primum non nocere (first, do no harm).

Finally, the cut-point of 14 for the PHQ-9 is the one that has the highest percentage of correct classifications, and therefore the highest accuracy. Although accuracy alone can be criticized for the determination of a cut-point, in this case it coincides with our goal of the study, i.e. having an appropriate balance between minimizing the number of false positive patients and false negative patients. In other words, we explored the nature of misclassification and deemed this cut-point to be the most appropriate. In addition to maximizing specificity, we would aim for a sensitivity that performs better than chance and therefore above 50%.

In studies assessing the general population, optimal cut-points for the PHQ-9 are 8-11, and traditionally 10 [42]. If a cut-point of 10 were used for our migraine sample, then the sensitivity would be better at 82.0% but specificity would be much lower at 79.9%, as compared to a cut-point of 14 (sensitivity of 60.0% and a specificity of 93.5%). Positive predictive value would also be quite poor at 56.9% (compared to 75.0% for PHQ-9 cut-point 14). Therefore, our data indicate that the cut-points recommended for the general population would not be suitable for a migraine population attending a headache clinic, for the reasons given above.

The PHQ-9 algorithm performed similarly to the PHQ-9 total score with the exception of a lower sensitivity at 53.8%. Given that the algorithm is more complicated to use for physicians in an outpatient setting and has a lower sensitivity (with a comparable specificity), the algorithm would not be recommended for routine depression screening in a migraine population attending a headache clinic.

The PHQ-2 score at a cut-point of 3 also showed slightly lower sensitivity and specificity compared to the PHQ-9 score. A cut-point of 2 did not perform as well overall. Even though the PHQ-2 is simple and quick for patients to complete, the poorer performance makes it a less desirable screening tool.

In regards to the HADS depression scale (HADS-D), the area under the ROC curve was 0.92 (95% CI 0.88-0.96). Similar to the PHQ-9, this indicates an overall good performance of the test as a screening tool in our migraine population. For the same reasons as described for the PHQ-9, a cut-point of 11 or greater was chosen as the most appropriate for the HADS-D. This led to a sensitivity of 59.6% and a specificity of 98.1%. Again, this cut-point happened to coincide with the highest level of correct patient classifications and accuracy. The HADS-D therefore showed a higher specificity than the PHQ-9 score and also showed a better positive

predictive value (91.2% as compared to 75.0% for the PHQ-9). Negative predictive value was similar.

For the overall migraine patients in this study therefore, the HADS-D had the best performance followed by the PHQ-9 score. One consideration however is the cost of the questionnaires. The HADS-D must be purchased, while the PHQ-9 is available free of charge. The HADS costs approximately fifty Canadian cents per test administration. One must balance this cost against the cost of the PHQ-9 performing somewhat poorer than the HADS-D. Having a closer look at the misclassification of the two tests will allow a better sense of this. In approximately 200 patients, the PHQ-9 questionnaire led to 10 false positive patients and 20 false negative patients, whereas the HADS-D led to 3 false positive patients and 21 false negative patients. Therefore, there were 7 extra false positive patients with the PHQ-9 and a comparable number of false negatives. If one considers the cost and stigma associated with 7 extra false positives for every 200 or so patients assessed, this can be substantial and may justify the cost of purchasing the HADS. If the results of our study are confirmed by other studies, it may be worth the extra cost to purchase the HADS, given the savings it provides in regards to a lower false positive rate.

If the goal of the study had been to identify the majority of patients with depression, even at the cost of higher false positive cases, then a lower cut-point for the PHQ-9 and HADS depression scales would have been appropriate. For example, the PHQ-9 at a cut-point of 9 has a sensitivity of 92.0% and a specificity of 74.0%, resulting in 78.4% of patients correctly classified. The HADS depression scale at a cut-point of 7 has a sensitivity of 92.3% and a specificity of 80.0%, resulting in 83.1% of patients correctly classified. The HADS would still be recommended over the PHQ-9 given the better performance. If the goal of the study had been

to identify a reasonable number of patients with depression but at the same time avoid a high number of false positive cases, then a cut-point with relatively equal sensitivity and specificity would have been more appropriate (e.g. PHQ-9 cut point 10 with sensitivity of 82.0% and specificity of 79.9% and HADS cut-point 8 with sensitivity and specificity of 86.5%). Therefore, the cut-points chosen are reflective of the goals of the screening.

So, is routine screening for depression worthwhile in an outpatient migraine population attending a tertiary care clinic? This is a question that requires further study to answer, as there are many factors to consider. The first factor is internal validity, i.e. how reflective are the results obtained in our study to the truth [72]. When one critically appraises diagnostic or screening tests for internal validity, three primary questions need to be answered. First, was there an independent blind comparison with a reference gold standard of diagnosis [72]? In our case, the answer is yes. The gold standard was the SCID and the SCID interviewers were blinded to the results of the screening questionnaires. This is an important element in reducing observer bias [72]. Observer bias refers to bias that can occur as a result of interpreting the reference or gold standard test with knowledge of the results of test under study. This bias can lead to an overestimation of a test's accuracy, especially if the gold standard test is open to subjective interpretation [73].

Second, was the diagnostic test evaluated in an appropriate spectrum of patients, similar to those who would be tested in clinical practice [72]? The answer to this would again be yes in our study. The population of interest is a migraine population attending a headache clinic and we were able to capture patients with varying degrees of headache frequency and disability, like those that would be seen at other headache centers. The wide variety of patients reduces spectrum bias [72]. Spectrum bias refers to bias that can occur as a result of performing the test

of interest in a select group of patients, for example in a group of people already known to have the disease and a separate group of normal patients. In such a case, diagnostic accuracy would be overestimated [73].

Third, was the reference gold standard applied to all patients regardless of the results of the screening tests [72]? The answer to this is yes as well. All patients were offered the SCID and completed the interview before the results of their screening tests were known. This is important in minimizing verification bias [72]. Verification bias can occur when the decision to perform the gold standard test is based on the results of the test under study. As such, some patients will have verification of their true disease status and others may not [73]. Therefore, we can conclude that our study met the criteria for internal validity. A second major factor to consider however is external validity, i.e. can the results obtained be applied to migraine populations in other headache centers [72]? This requires validation in a second independent group of patients [72]. Hence, more studies are necessary to prove external validity.

Once validation of the screening tests has occurred, the practical implications and impact have to be considered. We know that the PHQ-9 and HADS are quick and simple to administer, but will they be taken up for use by headache clinics? If taken up, will they improve patient care and patient outcomes? As mentioned in the introduction, in the primary care setting, some differences of opinion exist in the literature regarding routine screening. A Cochrane review in 2005 found that routine depression screening had very little effect on the management or outcomes of depression [37, 38]. The Canadian Task Force for Preventative Healthcare also does not currently recommend routine screening [39]. However, the US Preventive Services Task Force has endorsed screening for depression in primary care settings when appropriate staff-assisted care is available [34, 40]. In this case, we are dealing with a specific population

with a higher prevalence of depression than in a primary care setting [59]. Thus, studies need to be done looking specifically at migraine patients attending headache clinics.

Why may there be little or no effect on patient outcomes when screening for depression? A recent article tried to answer this question by assessing a number of systematic reviews and meta-analyses [26]. One of the larger systematic reviews published in 2009 assessed the diagnostic ability of family physicians to identify depression without any screening tools or aids [34]. They found that the diagnostic sensitivity across 41 studies was about 47.3%. In a smaller number of studies that reported both sensitivity and specificity, the sensitivity was 50.1% and specificity was about 81.3% [34]. It was felt that a great deal of the under-identification of depression stems from a lack of physician's judgment about the severity of symptoms. Therefore, the clinician recognizes symptoms of depression but judges them to be insignificant clinically [34]. It also appears that once depression is identified, treatment occurs in less than 60% of patients [26]. The exact reasons for this are unclear. In addition, once treatment is initiated, the care provided to patients is inadequate in up to 50% of cases. It appears that lack of close follow-up of patients is the primary reason for inadequate care [26]. So, in future studies, similar issues need to be considered in the migraine population when considering the practicalities of depression screening and its impact on the care of patients.

In our study, given that the false negative rate was high for the HADS-D, a number of patients, about 10% (20 or so in the group of 200) would not be diagnosed with depression. However, those that are diagnosed are more likely to be correctly diagnosed given the high specificity of the test. Because of the relatively high prevalence of depression in this population (about 25%), high positive and negative predictive values were seen as well. These have important clinical implications because a patient who tests positive or negative is likely to be

correctly diagnosed. If future studies do confirm the results of this study and do indeed show a positive impact of screening, the HADS-D may be an appropriate screening tool for such a population when resources allow. Screening of patients could occur after their initial visit with a headache specialist and perhaps on a yearly basis. Patients who are positive on the test can then be further assessed and referred to the appropriate resources.

When a stratified analysis of the depression scales was performed in our study, the scales did appear to perform better in certain groups of patients than others. Both scales performed better in males than females. In males, the PHQ-9 performed better than the HADS. Both scales also performed better in those aged 43 and over than those younger than 43 years, but at different cut-points than the overall sample (15 for PHQ-9 and 8 for the HADS-D). The PHQ-9 also performed better in patients with chronic migraine and in those who had migraine with aura. A different cut-point was felt to be best for migraine with aura patients (13 for PHQ-9 and 9 for HADS-D). However, 95% CI's were quite wide for a number of these estimates, indicating that the sensitivities and specificities were imprecise.

Should different tests and different cut-points be used for subsets of patients? If the screening tools are further validated and screening is felt to be worthwhile in future studies, a much larger study specifically designed to assess subgroups of patients would help answer this question. If a larger study confirms that the test performs better in certain patient groups, then ideally the test and cut-point that is most suitable for each group should be used. That would lead to less false negative and false positive cases. However, in a routine outpatient setting, this would likely be too complicated to do. This may be possible in a case where an electronic system, such as a computer program or other aid is available to bring up the best test for the patient based on a few key baseline questions. The computer program would then also be

programmed to know the best cut-point and provide a conclusion. Such a system would be costly and would require more research to determine if it would be worthwhile.

6.4 Objective 2: Prevalence of depression

The point prevalence of depression in this migraine population was 25.0% (95% CI 19.0-31.0%). Although this estimate is somewhat broad, it still provides an acceptable degree of precision. For patients who did not complete the SCID and only did the questionnaires, the point prevalence of depression was 18.6% (95% CI 9.7-30.9) using the PHQ-9 (cut-point 14) and 10.2% (95% CI 3.8-20.8) using the HADS-D (cut-point 11).

Stratified prevalence estimates were performed for the following baseline variables: gender, age groups, migraine type, migraine frequency, and being on migraine preventatives (Table 5.19). Of these, only chronic migraine appeared to be associated with a higher prevalence of depression. Those with chronic migraine had a depression prevalence of 36.0% (95% CI 25.1-46.9), whereas those with episodic migraine had a depression prevalence of 18.8% (95% CI 12.2-25.4) ($p=0.003$). This correlates with other studies reporting a higher prevalence of depression in those with more frequent migraine [66]. Females and patients who have migraine with aura also have a higher prevalence of depression in a number of studies [54, 56, 61-64], but this was not found in our study. The reasons for this are unclear, but smaller sample sizes for some of the subgroups in our study may not have allowed differences to be significant. In addition, mean age for depressed and non-depressed patients was also assessed in our study and did not significantly differ between the two groups.

Besides small sample sizes in the subgroups of patients discussed above, one can speculate why differences were not seen in depression prevalence estimates for gender and age

as previously reported in the general population. One reason may be that these differences are not as noticeable in patients who have migraine, because migraine itself serves as a powerful risk factor for depression. Perhaps migraine overpowers the other factors and thus makes them less prominent. In patients who do not have migraine, age and gender may be more prominent risk factors for depression. Rothmans' causal pie model may help to explain this [74]. Rothman's model divides the causes of disease into their component causes (pieces of the pie) and states that each component cause has an essential part in causing the disease [74]. Therefore, the disease is multifactorial with several factors contributing to its manifestation. Using this model, one can therefore speculate that age and gender are perhaps smaller component causes as opposed to migraine, which may be a much larger component cause of depression.

No significant difference was seen in the prevalence of depression between patients on migraine preventatives and those on no migraine preventatives. The reason for this may have been that the group not taking any preventatives was very small (n=14) and had quite an imprecise point prevalence estimate of depression (14.3%, 95% CI 0.0-32.6). One would expect a lower prevalence of depression in patients not on migraine preventatives because these patients are likely to have episodic, less frequent, and less disabling headaches.

Point prevalence estimates were also assessed with some depression variables (Table 5.20). Patients on antidepressants, those being treated by any means for depression and those diagnosed with depression by a health professional had much higher point prevalence estimates (around 40%) as compared to those not an antidepressants, not on any treatment for depression and not diagnosed with depression by a health professional. These differences were statistically significant. These results are not surprising. However, this indicates that about 40% of patients who are on antidepressants or are being treated for depression still meet criteria for major

depressive disorder. This may be for a number of reasons, such as inadequate dose of medication, recent initiation of treatment, poor response to treatment, or poor compliance with treatment. In addition, 27 (18.1%) of those patients not on antidepressants, 24 (17.0%) of those not on any treatment, and 12 (11.0%) of those not diagnosed with depression by a professional, met criteria for depression using the SCID. This indicates that quite a few patients that have not been treated for depression or diagnosed with depression do indeed have depression. The prevalence of untreated depression in this study was 17.0% (95% CI 10.8 – 23.2). Therefore, efforts to both better identify depression and manage it more effectively are important.

Among patients diagnosed with depression using the SCID (52 of 208 patients or 25.0%), 27 patients (51.9%) were not on any antidepressants, and 24 patients (46.1%) were not on any treatment for depression. This is a substantial number of patients, indicating that nearly half of patients diagnosed with depression were not receiving any treatment. There may be a number of reasons for this, including failure to diagnose or recognize the symptoms of depression as clinically significant, hesitation on the part of the patient or physician to initiate treatment, or no current treatment. Perhaps some of these patients were previously treated and did not tolerate the medications or did not have good response to treatment. Regardless, a high prevalence of untreated depression should alert the physician to determine the reasons for the lack of treatment and offer appropriate management if warranted.

As an example, the performance of the PHQ-9 score and HADS were assessed in the 27 patients diagnosed with depression using the SCID but not on antidepressants. These patients comprised the 18.1% of patients from the 149 not on antidepressants (Table 5.20). Scores were available for 26 patients for the PHQ-9; 12 patients were identified as having depression using this scale (cut-point of 14+, sensitivity 46.2%, 95% CI 26.6-66.6). The HADS identified 13 of

27 patients (sensitivity 48.1%, 95% CI 28.7-68.1). So although the 95% CI are quite wide for both scales, they still allowed the identification of depression in some of these patients which may have otherwise remained undiagnosed.

The PHQ-9 and HADS depression scores were also compared among several subgroups of patients (Table 5.21). Patients with chronic migraine, those diagnosed with depression using the SCID, those self-reporting depression, those treated for depression, and those patients on antidepressants all had higher median scores on the PHQ-9 and HADS as compared to their counterparts. These associations were all statistically significant. These results are reassuring as they indicate that the PHQ-9 and HADS scores correlate appropriately with the symptoms of depression and were higher in the groups where they were expected to be higher.

6.5 Objective 3: Disability and quality of life

6.5.1 Migraine disability as assessed by the MIDAS

In scoring the MIDAS, there were 10 patients with out of bounds data. However, the inclusion or exclusion of such data did not alter the overall results obtained. In this migraine population, more than half of patients had severe disability and about 17% had moderate disability according to the MIDAS. This is reflective of the more severe disease experienced by patients seen at a specialty headache clinic. A similar population was assessed in a prior study [1]. The population (n=864) included headache patients referred to one of five neurology clinics in Canada with expertise in headache diagnosis and management. The data for these patients was obtained from the Canadian Headache Outpatient Registry and Database (CHORD), which prospectively collected data on new patients between September 2001 to January 2004 [1]. Although some patients had other headache diagnoses, the majority of patients had migraine.

Baseline features were very similar to our study sample. The authors found that 57% of patients were categorized as having severe disability according to the MIDAS. This is very similar to our finding of nearly 55%.

In our study, further data collected through the MIDAS revealed that 37% of patients had chronic migraine and 21% had daily headache. This compares similarly to the population in the study described above [1]. Furthermore, patients in our study felt that their average pain intensity was moderate to severe and enough to limit some activities, especially those of lower priority. Patients also felt that their headaches interfered at a moderately severe level with their daily life. Again, this data indicates that these patients are affected to a significant extent by their headaches. This is similar to other reports in the literature [1, 17, 75].

When stratified by gender, age, migraine type, migraine frequency, a SCID diagnosis of depression, depression by self-report, treatment for depression, on antidepressants, and on migraine preventatives, some interesting differences were found for MIDAS scores (Table 5.22). Patients who were younger than 43 years of age, who were diagnosed with depression using the SCID, who self-reported depression, who were on treatment for depression, and who were taking antidepressants had more severe disability according to the MIDAS as compared to their counterparts. These associations were all statistically significant. Patients with chronic migraine did not appear to have more severe disability than the episodic migraine patients. This may have occurred because many patients with episodic migraine may still have fallen into the moderate and severe MIDAS categories even if their absolute score was lower than the chronic group. Given the non-normal distribution of the MIDAS, a mean or median would not be accurate for comparing these groups. No differences were seen in patients on migraine preventatives as

compared to those not on migraine preventatives. Similar to before, this may have occurred because of the low number of patients in the group without preventatives (n=14).

6.5.2 Quality of life as assessed by the SF-12

For the Physical Component Score (PCS) of the SF-12, the mean was 43.1 and median was 44.6, where zero is the lowest health-related quality of life and 100 is the highest. For the Mental Component Score (MCS), the mean was 35.0 and median was 36.3. Therefore, the MCS was lower than the PCS, indicating a lower mental health-related quality of life in this population of migraine patients. This may have occurred because some migraine patients may have experienced cognitive and mental difficulties to a larger extent than physical limitations during a migraine. A population-based study [76] assessed SF-12 scores in 389 migraine patients identified by a validated telephone interview and reported a mean score of 45.2 for the PCS and a score of 43.4 for the MCS. As compared to our study sample, the PCS was similar but the MCS was quite a bit higher in the population-based study. This is likely because our study sample consisted of more severely affected migraine patients as compared to migraine patients in the general population.

The same stratification of variables performed for the MIDAS was done for the SF-12 PCS and MCS (Table 5.23). Patients with a SCID diagnosis of depression, depression by self-report, treatment for depression by any modality, and treatment for depression with an antidepressant had lower quality of life scores both physically and mentally compared to their counterparts. These associations were statistically significant. The population-based study described above also showed significantly lower scores for patients with comorbid migraine and depression for both the PCS and MCS [76]. In our study, patients with chronic migraine also

had lower scores on their PCS (median score 40.0), as compared to patients with episodic migraine (median score 45.4). This difference was not seen for the MCS (medians scores 35.2 and 36.6 for chronic and episodic migraine respectively). A recent study [75] assessed 123 episodic and 123 chronic migraine patients diagnosed by a specialist from the Jefferson Headache Center in Philadelphia and found significantly lower scores for chronic migraine patients for both the PCS (median score of 37 for chronic migraine and 43 for episodic migraine) and MCS (median score 37 for chronic migraine and 49 for episodic migraine). It is unclear why our study showed a difference in PCS scores but not MCS scores. However, the median scores are comparable with the Jefferson Headache Center study, other than the lower score in our study for episodic migraine patients on the MCS.

6.5.3 Disability and Quality of life

Overall therefore, disability and quality of life are both affected to a large extent in the migraine population but certainly more affected in patients who also have depression. Even patients that are being treated for depression experience greater disability and a poorer quality of life. These results are in keeping with a number of other studies [8, 10, 19, 62, 75, 76], and demonstrate the importance of recognition and appropriate management of this group of patients.

6.6 Projection to other migraine populations

This study included a migraine population that attends a specialty headache clinic at a tertiary care center. Therefore, the patients are likely more severely affected by their migraines than the general migraine population. Thus, the prevalence estimate of 25% for depression and the positive predictive value (PPV) obtained from the tools in this study may not be reflective of

other migraine populations. Bayes Theorem allows the prediction of the PPV with different prevalence estimates. Sensitivity and specificity, unlike the PPV or NPV, do not depend on the prevalence estimate. In this case, the formula therefore uses the sensitivity and specificity of the tools in this study to predict PPV (Figure 5.10).

When Bayes Theorem is applied using the PHQ-9 and the HADS depression scale at the selected cut-points (14 for PHQ-9 and 11 for HADS), it shows that PPV rises as prevalence of depression rises. The PPV rises more rapidly when prevalence estimates increase from 5% to 25% and starts to plateau when prevalence estimates reach 50% and higher. In this case, the HADS performs better, demonstrating a higher PPV for a given prevalence estimate as compared to the PHQ-9. Once the prevalence estimates reach very high levels (above 70% or so), both tools perform very well, converge, and approach 100%. Physicians can therefore use the tools in this study, as well as future studies, to predict the PPV for their own migraine population. However, the accuracy of such projections depends on the equivalence of sensitivity and specificity in different populations.

6.7 Study Strengths

This study has several strengths. It is the first of its kind to formally assess two commonly used depression screening tools in migraine patients, and begin the process of validation of the tools in this population using a gold standard semi-structured psychiatric interview. As described above, this study meets the criteria for internal validity of diagnostic/screening tests, using critical appraisal criteria set out to evaluate such studies. As a result of being internally valid, this study significantly reduces concerns with observer bias, spectrum bias, and verification bias. Second, the order of the PHQ-9 and HADS questionnaires

was randomized to reduce any possible influence from responses to the first questionnaire on the patient's responses to the second questionnaire. Third, the questionnaires and SCID were not very time-consuming or burdensome for patients to complete and did not interfere with medical appointments. Fourth, there was minimal missing data or out of bounds values in the study which preserved the high quality of the data. Fifth, this study provided an estimate of the prevalence of depression using the gold standard psychiatric interview, rather than by self-report, or through the use of questionnaires or surveys. This is a major factor in the study minimizing bias. Previous prevalence studies have utilized surveys or other techniques besides a psychiatric interview, and as a result have been afflicted by various forms of bias, particularly selection and misclassification bias [58, 59]. Selection bias refers to systematic error as a result of the manner in which participants are selected for a study. This type of bias can occur when the characteristics of the participants selected for the study are systematically different from those in the source population [74]. Misclassification bias is systematic error as a result of measurement flaws in a study, leading to misclassification of participants with regard to exposure or outcome status [74]. Sixth, this study used consecutive patients deemed eligible for the study, which further reduces selection bias. Seventh, participation rate was fair in this study and when the study sample was compared to the eligible study population, no major differences were found. Therefore, the study sample was representative of the eligible study population. This is yet another factor reducing selection bias. Eighth, this study sheds more light on the fact that patients with both migraine and depression have higher degrees of disability and poorer quality of life. Ninth, this study adds further to the body of knowledge regarding migraine and depression and illustrates the need for further studies to fully validate the depression screening tools and assess the impact of screening in the migraine population.

6.8 Study Limitations

There are some limitations to this study. First, although the MIDAS has been previously validated in migraine patients, there is still room for error in interpreting some of the questions. In this study, the MIDAS was misinterpreted by a few patients, leading to some out of bounds values. This reduced the effective sample size slightly for that portion of the analysis. However, the small reduction of the sample size did not impact the overall results to any significant extent. The MIDAS was chosen over other validated headache disability scales because it gives clinically relevant and easily interpretable results [1]. Second, the Headache clinic does tend to see migraine patients who have more moderate to severe conditions. In addition, by seeing follow-up patients, it is possible that some new patients that had milder disease and did not require further follow-up in the clinic were not included in our sample. As a result, the prevalence of depression may be higher in this population than in the general migraine population. However, we would not be generalizing our prevalence estimate to the general migraine population, but rather to patients attending a headache clinic. As described above, Bayes theorem is also a method to apply the data from this study to other populations. Third, the prevalence estimate was somewhat imprecise with moderately wide 95% CI's (19-31%). This was related to a smaller sample size than we had hoped for. However, despite the broad range for the prevalence estimate, it still allows one to see the high prevalence of depression in this migraine population. Fourth, some of the subgroups in the study had small to moderate sample sizes. This led to some imprecise values for the psychometric properties of the depression screening tools and for some of the prevalence estimates. However, this was not one of the main goals of the study and the subgroup analysis was done mainly in an exploratory manner. Future larger studies could be designed specifically to look at subgroup differences.

6.9 Knowledge Translation and Dissemination

The results of this study have the potential to impact many people, including healthcare professionals in the fields of neurology, psychiatry and family medicine, patients and their families, and policy-makers. Therefore, the knowledge gained in this study needs to be disseminated to these individuals. Dissemination of information to health-care professionals would take the form of publication of manuscripts, presentation at neurology, psychiatry, and family medicine grand rounds, and national and international conferences, such as the Canadian Headache Society meeting. In addition, until future studies are done, the Bayes Theorem formula or graph could be used by clinicians to determine the positive predictive value of a patient having depression using the PHQ-9 or HADS. This would help guide clinicians on management; i.e. referral to a psychiatrist or mental health professional, referral to other community resources, etc.

For patients and families, a brief discussion could take place during visits with their physician. In addition, handouts and brochures summarizing the information can be made available. In regards to policy-makers, the information can be presented to them directly, such as through letters outlining the impact of the study results and resources needed to improve patient care and for future studies. Indirectly, the information can also be conveyed to them through the media.

6.10 Future studies

Future studies are required to further validate the PHQ-9 and HADS depression scales. This would ideally involve two to three different migraine populations attending various headache centers and larger sample sizes. Larger sample sizes would then lead to more precise

psychometric values on the depression screening tools and allow for subgroup analysis. Studies looking at the direct and indirect costs for populations using screening tools and those not using screening tools would also be invaluable. Following that, if screening is felt to be worthwhile and cost-efficient, the impact of screening needs to be fully assessed (preferably through a randomized controlled trial) to allow firm conclusions to be made regarding screening in the migraine population.

Bibliography

1. Sauro, K.M., et al., *HIT-6 and MIDAS as measures of headache disability in a headache referral population.* Headache, 2010. **50**(3): p. 383-95.
2. Bigal, M.E. and R.B. Lipton, *The epidemiology, burden, and comorbidities of migraine.* Neurol Clin, 2009. **27**(2): p. 321-34.
3. Vos, T., et al., *Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010.* Lancet, 2012. **380**(9859): p. 2163-96.
4. Levin, M., *Classification of Primary Headaches: Concepts and Controversies.* . Continuum - Headache: American Academy of Neurology, 2006. **12**: p. 32-51.
5. Headache Classification Subcommittee of the International Headache, S., *The International Classification of Headache Disorders: 2nd edition.* Cephalalgia, 2004. **24 Suppl 1**: p. 9-160.
6. Ward, T.N., *Migraine diagnosis and pathophysiology.* Continuum (Minneap Minn), 2012. **18**(4): p. 753-63.
7. Abu-Arefeh, I. and G. Russell, *Prevalence of headache and migraine in schoolchildren.* BMJ, 1994. **309**(6957): p. 765-9.
8. Breslau, N. and G.C. Davis, *Migraine, physical health and psychiatric disorder: a prospective epidemiologic study in young adults.* J Psychiatr Res, 1993. **27**(2): p. 211-21.
9. Breslau, N., et al., *Joint 1994 Wolff Award Presentation. Migraine and major depression: a longitudinal study.* Headache, 1994. **34**(7): p. 387-93.
10. Jette, N., et al., *Comorbidity of migraine and psychiatric disorders--a national population-based study.* Headache, 2008. **48**(4): p. 501-16.
11. Kalaydjian, A. and K. Merikangas, *Physical and mental comorbidity of headache in a nationally representative sample of US adults.* Psychosom Med, 2008. **70**(7): p. 773-80.
12. Lipton, R.B., et al., *Prevalence and burden of migraine in the United States: data from the American Migraine Study II.* Headache, 2001. **41**(7): p. 646-57.
13. Kurth, T., H. Chabriat, and M.G. Bousser, *Migraine and stroke: a complex association with clinical implications.* Lancet Neurol, 2012. **11**(1): p. 92-100.

14. Burch, R., *Headache Currents Commentary.* Headache, 2012.
15. Irene Worthington1, T.P., Marek J. Gawel1,8,9, Jonathan Gladstone1,2, Paul Cooper4, Esma Dilli5, Michel Aube6, Elizabeth Leroux7, Werner J. Becker3 and o.b.o.t.C.H.S.A.M.T.G.D. Group, *Canadian Headache Society Guideline Acute Drug Therapy for Migraine.* Canadian Journal of Neurological Sciences, 2013. **40**(Sept, Supp 3).
16. Tamara Pringsheim1, W.J.D., Gordon Mackie2, Irene Worthington3, Michel Aubé4, Suzanne N. Christie5, Jonathan Gladstone6, Werner J. Becker1 on behalf of the Canadian Headache Society Prophylactic Guidelines Development Group, *Canadian Headache Society Guideline for Migraine Prophylaxis.* Canadian Journal of Neurological Sciences, 2012. **39**(Number 2; Supp 2).
17. Stewart, W.F., et al., *Reliability of the migraine disability assessment score in a population-based sample of headache sufferers.* Cephalalgia, 1999. **19**(2): p. 107-14; discussion 74.
18. Busija, L., et al., *Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL).* Arthritis Care Res (Hoboken), 2011. **63 Suppl 11**: p. S383-412.
19. Lanteri-Minet, M., et al., *Anxiety and depression associated with migraine: influence on migraine subjects' disability and quality of life, and acute migraine management.* Pain, 2005. **118**(3): p. 319-26.
20. Ware, J., Jr., M. Kosinski, and S.D. Keller, *A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity.* Med Care, 1996. **34**(3): p. 220-33.
21. Feinstein, A.R., *The Basic Elements of Clinical Science.* J Chronic Dis, 1963. **16**: p. 1125-33.

22. Whiteford, H.e.a., *Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010.* Lancet, 2013. **382**(Nov 9): p. 1575-86.
23. Murray, C.J., et al., *Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010.* Lancet, 2012. **380**(9859): p. 2197-223.
24. Organization, W.H. *Depression fact sheet.* Available from: http://www.who.int/mediacentre/factsheets/fs369/en/.
25. Association, A.P., *Diagnostic and Statistical Manual of Mental Disorders* Fourth Edition ed. 2000, Washington.
26. Craven, M.A. and R. Bland, *Depression in primary care: current and future challenges.* Can J Psychiatry, 2013. **58**(8): p. 442-8.
27. Patten, S.B., et al., *Descriptive epidemiology of major depression in Canada.* Can J Psychiatry, 2006. **51**(2): p. 84-90.
28. Russo, S.J. and E.J. Nestler, *The brain reward circuitry in mood disorders.* Nat Rev Neurosci, 2013. **14**(9): p. 609-25.
29. Villanueva, R., *Neurobiology of major depressive disorder.* Neural Plast, 2013. **2013**: p. 873278.
30. Ramasubbu, R., et al., *The Canadian Network for Mood and Anxiety Treatments (CANMAT) task force recommendations for the management of patients with mood disorders and select comorbid medical conditions.* Ann Clin Psychiatry, 2012. **24**(1): p. 91-109.
31. Magni, L.R., et al., *Fluoxetine versus other types of pharmacotherapy for depression.* Cochrane Database Syst Rev, 2013. **7**: p. CD004185.
32. Blackwell, B., *Adverse effects of antidepressant drugs. Part 1: monoamine oxidase inhibitors and tricyclics.* Drugs, 1981. **21**(3): p. 201-19.
33. Kestenbaum, B., *EPIDEMIOLOGY AND BIOSTASTICS: AN INTRODUCTION TO CLINICAL RESEARCH*. 2009, University of Washington.
34. Mitchell, A.J., *Clinical utility of screening for clinical depression and bipolar disorder.* Curr Opin Psychiatry, 2012. **25**(1): p. 24-31.

35. Kanchanaraksa, S., *Evaluation of Diagnostic and Screening Tests: Validity and Reliability*. 2008, Johns Hopkins University: Johns Hopkins Bloomberg School of Public Health.
36. Latimer, E.A., G.R. Bond, and R.E. Drake, *Economic approaches to improving access to evidence-based and recovery-oriented services for people with severe mental illness.* Can J Psychiatry, 2011. **56**(9): p. 523-9.
37. Gilbody, S., A.O. House, and T.A. Sheldon, *Screening and case finding instruments for depression.* Cochrane Database Syst Rev, 2005(4): p. CD002792.
38. Maurer, D.M., *Screening for depression.* Am Fam Physician, 2012. **85**(2): p. 139-44.
39. Canadian Task Force on Preventive Health, C., et al., *Recommendations on screening for depression in adults.* CMAJ, 2013. **185**(9): p. 775-82.
40. Force, U.S.P.S.T., *Screening for depression in adults: U.S. preventive services task force recommendation statement.* Ann Intern Med, 2009. **151**(11): p. 784-92.
41. Kroenke, K., R.L. Spitzer, and J.B. Williams, *The PHQ-9: validity of a brief depression severity measure.* J Gen Intern Med, 2001. **16**(9): p. 606-13.
42. Manea, L., S. Gilbody, and D. McMillan, *Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis.* CMAJ, 2012. **184**(3): p. E191-6.
43. Arroll, B., N. Khin, and N. Kerse, *Screening for depression in primary care with two verbally asked questions: cross sectional study.* BMJ, 2003. **327**(7424): p. 1144-6.
44. Zigmond, A.S. and R.P. Snaith, *The hospital anxiety and depression scale.* Acta Psychiatr Scand, 1983. **67**(6): p. 361-70.
45. Bjelland, I., et al., *The validity of the Hospital Anxiety and Depression Scale. An updated literature review.* J Psychosom Res, 2002. **52**(2): p. 69-77.
46. Kroenke, K., et al., *The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review.* Gen Hosp Psychiatry, 2010. **32**(4): p. 345-59.
47. Lowe, B., et al., *Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9).* J Affect Disord, 2004. **81**(1): p. 61-6.

48. Whooley, M.A., et al., *Case-finding instruments for depression. Two questions are as good as many.* J Gen Intern Med, 1997. **12**(7): p. 439-45.
49. Schiffer, R.B., *Depression in neurological practice: diagnosis, treatment, implications.* Semin Neurol, 2009. **29**(3): p. 220-33.
50. Feinstein, A., *The neuropsychiatry of multiple sclerosis.* Can J Psychiatry, 2004. **49**(3): p. 157-63.
51. First M, S.R., Gibbon M, Williams J, *Structured clinical interview for DSM-IV Axis I disorders-patient edition (SCID-I/P, Version 2.0).* 1995: New York: New York State Psychiatric Institute.
52. Sanchez-Villegas, A., et al., *Validity of a self-reported diagnosis of depression among participants in a cohort study using the Structured Clinical Interview for DSM-IV (SCID-I).* BMC Psychiatry, 2008. **8**: p. 43.
53. Breslau, N., et al., *Comorbidity of migraine and depression: investigating potential etiology and prognosis.* Neurology, 2003. **60**(8): p. 1308-12.
54. Breslau, N., K. Merikangas, and C.L. Bowden, *Comorbidity of migraine and major affective disorders.* Neurology, 1994. **44**(10 Suppl 7): p. S17-22.
55. Breslau, N., et al., *Headache and major depression: is the association specific to migraine?* Neurology, 2000. **54**(2): p. 308-13.
56. Kececi, H., S. Dener, and E. Analan, *Co-morbidity of migraine and major depression in the Turkish population.* Cephalalgia, 2003. **23**(4): p. 271-5.
57. Ratcliffe, G.E., et al., *The relationship between migraine and mental disorders in a population-based sample.* Gen Hosp Psychiatry, 2009. **31**(1): p. 14-9.
58. Antonaci, F., et al., *Migraine and psychiatric comorbidity: a review of clinical findings.* J Headache Pain, 2011. **12**(2): p. 115-25.
59. Molgat, C.V. and S.B. Patten, *Comorbidity of major depression and migraine--a Canadian population-based study.* Can J Psychiatry, 2005. **50**(13): p. 832-7.
60. Breslau, N., *Migraine, suicidal ideation, and suicide attempts.* Neurology, 1992. **42**(2): p. 392-5.
61. Merikangas, K.R., D.E. Stevens, and J. Angst, *Psychopathology and headache syndromes in the community.* Headache, 1994. **34**(8): p. S17-22.

62. Peterlin, B.L., M.J. Katsnelson, and A.H. Calhoun, *The associations between migraine, unipolar psychiatric comorbidities, and stress-related disorders and the role of estrogen.* Curr Pain Headache Rep, 2009. **13**(5): p. 404-12.
63. Oedegaard, K.J., et al., *Migraine with and without aura: association with depression and anxiety disorder in a population-based study. The HUNT Study.* Cephalalgia, 2006. **26**(1): p. 1-6.
64. Samaan, Z., et al., *Migraine in recurrent depression: case-control study.* Br J Psychiatry, 2009. **194**(4): p. 350-4.
65. Buse, D.C., et al., *Psychiatric comorbidities of episodic and chronic migraine.* J Neurol, 2013. **260**(8): p. 1960-9.
66. Tietjen, G.E., et al., *High prevalence of somatic symptoms and depression in women with disabling chronic headache.* Neurology, 2007. **68**(2): p. 134-40.
67. Lee, S., et al., *Association of headache with childhood adversity and mental disorder: cross-national study.* Br J Psychiatry, 2009. **194**(2): p. 111-6.
68. Modgill, G., et al., *A population-based longitudinal community study of major depression and migraine.* Headache, 2012. **52**(3): p. 422-32.
69. Baskin, S.M. and T.A. Smitherman, *Migraine and psychiatric disorders: comorbidities, mechanisms, and clinical applications.* Neurol Sci, 2009. **30 Suppl 1**: p. S61-5.
70. Simon GE, R.C., Peterson D, Oliver M, Whiteside U, Operskalski B, Ludman EJ., *Does response on the PHQ-9 Depression Questionnaire predict subsequent suicide attempt or suicide death?* Psychiatr Serv, 2013. **64**(12): p. 1195-202.
71. Canada, H.N.; Available from: http://headachenetwork.ca/publications/.
72. Jaeschke, R., G. Guyatt, and D.L. Sackett, *Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group.* JAMA, 1994. **271**(5): p. 389-91.
73. Lijmer, J.G., et al., *Empirical evidence of design-related bias in studies of diagnostic tests.* JAMA, 1999. **282**(11): p. 1061-6.
74. Oleckno, W.A., *Epidemiology - Concepts and Methods*. 2008, Long Grove, Illinois: Waveland Press, Inc.
75. Young, W.B., et al., *The stigma of migraine.* PLoS One, 2013. **8**(1): p. e54074.

76. Lipton, R.B., et al., *Migraine, quality of life, and depression: a population-based case-control study*. Neurology, 2000. **55**(5): p. 629-35.

Appendix A: ICHD diagnostic criteria for Migraine headache without aura and with aura*

Migraine without aura:

Recurrent headache disorder manifesting in attacks lasting 4-72 hours. Typical characteristics of the headache are unilateral location, pulsating quality, moderate or severe intensity, aggravation by routine physical activity and association with nausea and/or photophobia and phonophobia.

Diagnostic criteria:

A. At least 5 attacks fulfilling criteria B-D
B. Headache attacks lasting 4-72 hours (untreated or unsuccessfully treated)
C. Headache has at least two of the following characteristics:
   1. unilateral location
   2. pulsating quality
   3. moderate or severe pain intensity
   4. aggravation by or causing avoidance of routine physical activity (eg, walking or climbing stairs)
D. During headache at least one of the following:
   1. nausea and/or vomiting
   2. photophobia and phonophobia
E. Not attributed to another disorder

Migraine with aura:

Typical aura consisting of visual and/or sensory and/or speech symptoms. Gradual development, duration no longer than one hour, a mix of positive and negative features and complete reversibility characterize the aura which is associated with a headache fulfilling criteria for 1.1 **Migraine without aura.**

Diagnostic criteria:

A. At least 2 attacks fulfilling criteria B-D

B. Aura consisting of at least one of the following, but no motor weakness:
   a) fully reversible visual symptoms including positive features (eg, flickering lights, spots or lines) and/or negative features (ie, loss of vision)
   b) fully reversible sensory symptoms including positive features (ie, pins and needles) and/or negative features (ie, numbness)
   c) fully reversible dysphasic speech disturbance

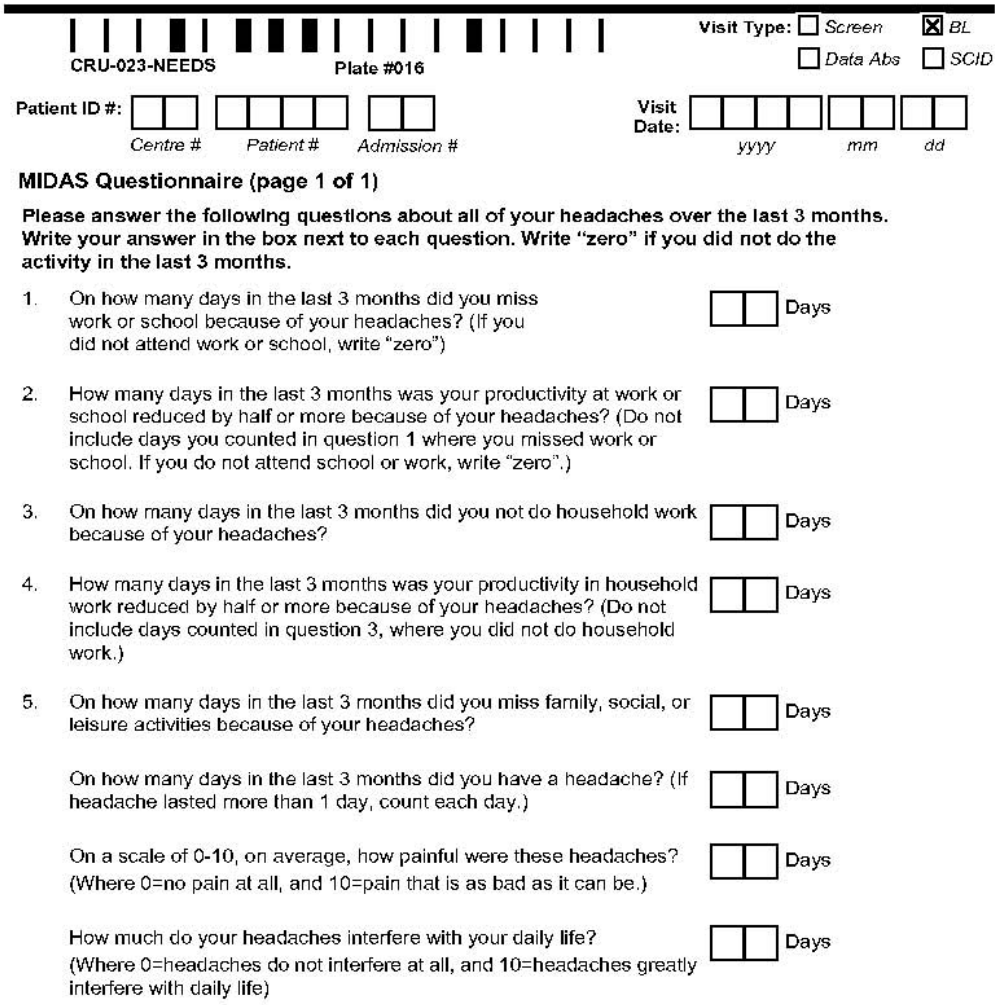
C. At least two of the following:
   a) homonymous visual symptoms and/or unilateral sensory symptoms
   b) at least one aura symptom develops gradually over ≥5 minutes and/or different aura symptoms occur in succession over ≥5 minutes
   c) each symptom lasts ≥5 and ≤60 minutes

D. Headache fulfilling criteria B-D for 1.1 **Migraine without aura** begins during the aura or follows aura within 60 minutes.

E. Not attributed to another disorder

* Headache Classification Subcommittee of the International Headache, S., The International Classification of Headache Disorders: 2nd edition. Cephalalgia, 2004. 24 Suppl 1: p. 9-160.

Appendix B: MIDAS**



CRU-023-NEEDS Plate #016

Visit Type: ☐ Screen ☒ BL ☐ Data Abs ☐ SCID

Patient ID #: ☐☐ ☐☐☐☐ ☐☐ (Centre # / Patient # / Admission #)

Visit Date: ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

MIDAS Questionnaire (page 1 of 1)

Please answer the following questions about all of your headaches over the last 3 months. Write your answer in the box next to each question. Write "zero" if you did not do the activity in the last 3 months.

1. On how many days in the last 3 months did you miss work or school because of your headaches? (If you did not attend work or school, write "zero") ☐☐ Days

2. How many days in the last 3 months was your productivity at work or school reduced by half or more because of your headaches? (Do not include days you counted in question 1 where you missed work or school. If you do not attend school or work, write "zero".) ☐☐ Days

3. On how many days in the last 3 months did you not do household work because of your headaches? ☐☐ Days

4. How many days in the last 3 months was your productivity in household work reduced by half or more because of your headaches? (Do not include days counted in question 3, where you did not do household work.) ☐☐ Days

5. On how many days in the last 3 months did you miss family, social, or leisure activities because of your headaches? ☐☐ Days

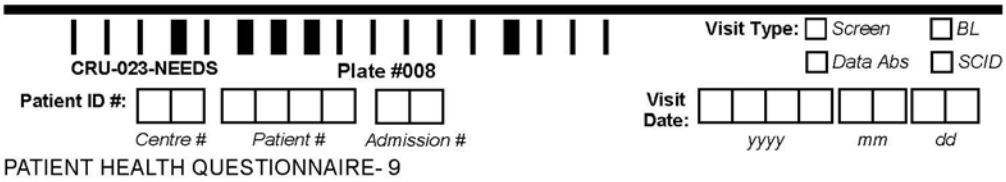
On how many days in the last 3 months did you have a headache? (If headache lasted more than 1 day, count each day.) ☐☐ Days

On a scale of 0-10, on average, how painful were these headaches? (Where 0=no pain at all, and 10=pain that is as bad as it can be.) ☐☐ Days

How much do your headaches interfere with your daily life? (Where 0=headaches do not interfere at all, and 10=headaches greatly interfere with daily life) ☐☐ Days

** Stewart, W.F., et al., Reliability of the migraine disability assessment score in a population-based sample of headache sufferers. Cephalalgia, 1999. 19(2): p. 107-14; discussion 74.

Appendix C: PHQ-9***



PATIENT HEALTH QUESTIONNAIRE- 9

Over the *last 2 weeks*, how often have you been bothered by any of the following problems? (Check the response that most closely applies to you)

	Not At All	Several Days	More Than Half the Days	Nearly Every Day
1. Little interest or pleasure in doing things	☐	☐	☐	☐
2. Feeling down, depressed, or hopeless	☐	☐	☐	☐
3. Trouble falling or staying asleep, or sleeping too much	☐	☐	☐	☐
4. Feeling tired or having little energy	☐	☐	☐	☐
5. Poor appetite or overeating	☐	☐	☐	☐
6. Feeling bad about yourself- or that you are a failure or have let yourself or your family down	☐	☐	☐	☐
7. Trouble concentrating on things, such as reading the newspaper or watching television	☐	☐	☐	☐
8. Moving or speaking so slowly that other people could have noticed. Or, being so fidgety or restless that you have been moving around a lot more than usual	☐	☐	☐	☐
9. Thoughts that you would be better off dead or of hurting yourself in some way	☐	☐	☐	☐

10. If you checked off *any* problems (questions 1-9), how *difficult* have these problems made it for you to do your work, take care of things at home, or get along with other people?

☐ Not difficult at all

☐ Somewhat difficult

☐ Very difficult

☐ Extremely difficult

*** Kroenke, K., R.L. Spitzer, and J.B. Williams, The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med, 2001. 16(9): p. 606-13.

Appendix D: PHQ-9 Algorithm***

The algorithm is calculated based on the following criteria:

A. Of the 9 questions in the PHQ-9, the patient must score 2 or greater on at least 5 of the 9 questions, one of which must be question 1 or 2.
B. Question 9 (suicidal ideation) can be counted towards one of the 5 questions if the score is ≥ 1.

*** Kroenke, K., R.L. Spitzer, and J.B. Williams, The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med, 2001. 16(9): p. 606-13.

Appendix E: Demographics Questionnaire

CRU-023-NEEDS Plate #006

Visit Type: ☐ Screen ☐ BL ☐ Data Abs ☐ SCID

Patient ID #: ☐☐ ☐☐☐☐ ☐☐ (Centre # / Patient # / Admission #)

Visit Date: ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

DEMOGRAPHICS (page 1 of 2)

Thank you for taking the time to complete this questionnaire. Due to the nature of the study, some of the questions below will seem very similar. Please answer each question based on its own merits.

1. Which of the following best describes your current marital status? (check one)

☐ Married ☐ Separated ☐ Divorced ☐ Widowed

☐ Common-law (living together) ☐ Single, never married

2. What is the highest grade, degree, certificate or diploma that you have ever completed?

☐ Grade 8 or lower

☐ Grade 9-11

☐ Grade 12-13

☐ Bachelor's degree

☐ Trade certificate or diploma from a vocational school or apprenticeship training

☐ Non-university certificate or diploma from a community college, CEGEP, school or nursing

☐ University certificate below bachelor's level

☐ University degree or certificate above bachelor's level

3. Are you currently working in a paid job? Please include part-time jobs, seasonal work, contract work, self-employment, baby-sitting and any other paid work, regardless of number of hours worked.

☐ Yes ☐ No

4. What is your primary source of income? (check one)

☐ Employment ☐ Social Services ☐ Spouse/Family Support ☐ Investment/Savings

☐ AISH ☐ Retirement Pension ☐ Disability Insurance

5. Do you ever drink alcohol (e.g. beer, wine or spirits)? (check one)

☐ Yes ☐ No

6. Do you currently smoke cigarettes? (check one)

☐ Yes ☐ No

7. Do you use illicit drugs such as marijuana or cocaine? (check one)

☐ Yes ☐ No

version date: July 24, 2012

CRU-023-NEEDS Plate #007

Visit Type: ☐ Screen ☐ BL ☐ Data Abs ☐ SCID

Patient ID #: ☐☐ ☐☐☐☐ ☐☐ (Centre # / Patient # / Admission #)

Visit Date: ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

DEMOGRAPHICS (page 2 of 2)

CLINICAL BACKGROUND

8. Do you have side-effects to any of your current medications?

☐ Yes ☐ No

9. Have you been diagnosed with depression by a health professional? (check one)

☐ Yes- please answer questions 10-12 ☐ No- go to next page

10. Who is currently managing (treating) your depression? (select all that apply)

☐ My family physician is managing my depression ☐ A psychiatrist is managing my depression

☐ Another specialist is managing my depression. Please specify specialty (e.g. Neurologist):

_____

☐ A counselor is managing my depression. Please specify the type of counselor (e.g. Social Worker):

_____

☐ Another person is managing my depression. Please specify that person's profession (e.g. a Nurse)

_____

☐ A psychologist is managing my depression

☐ I do not currently have anyone managing my depression

11. Are you currently taking any anti-depressant medications (e.g. Celexa, Effexor, etc.)?

☐ Yes ☐ No

12. Are you currently receiving other forms of treatment for your depression? Please select all that apply.

☐ Medications

☐ Counseling

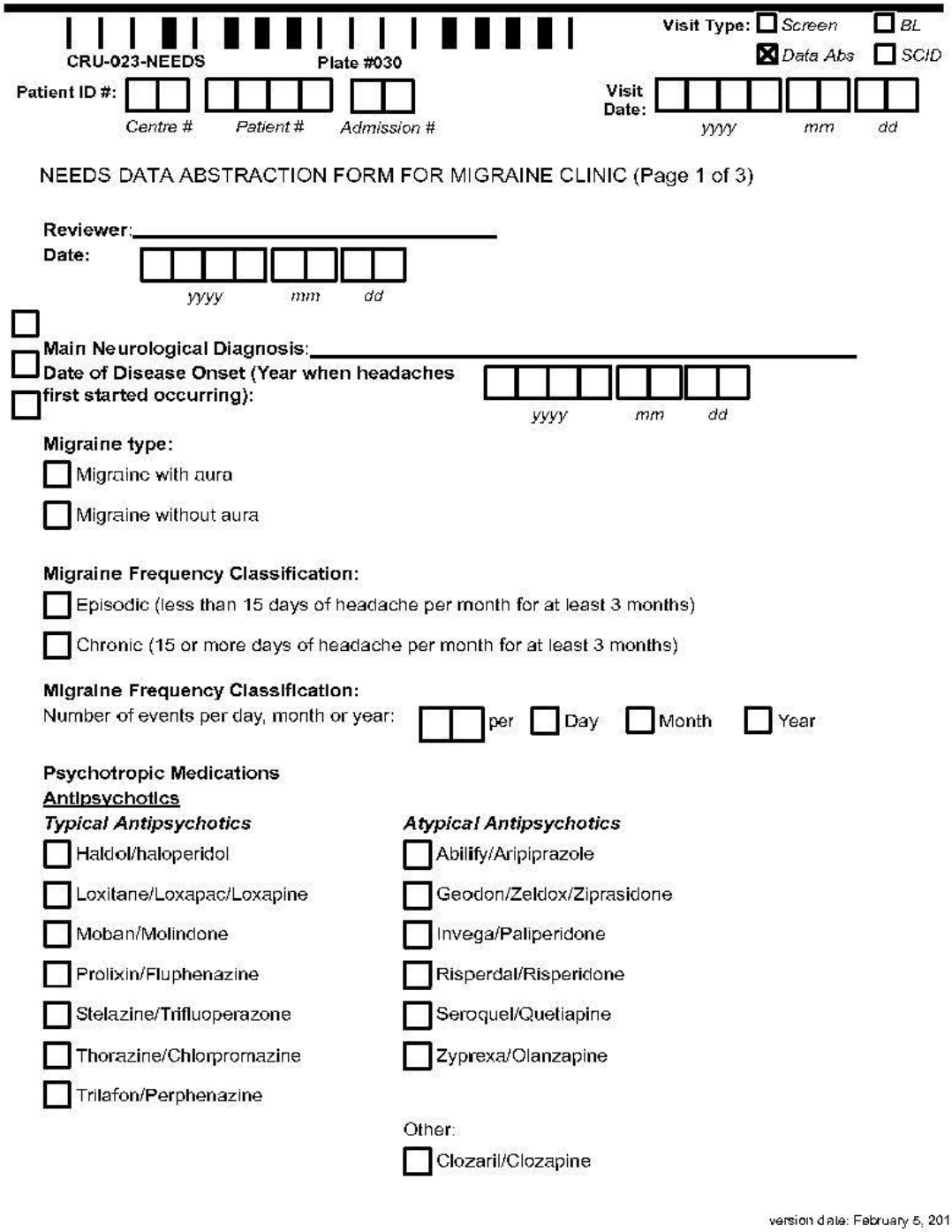
☐ Talk Therapy (cognitive behavioral therapy)

☐ Group Therapy

☐ Other- please specify:_____

version date: July 24, 2012

Appendix F: Data Abstraction Form



CRU-023-NEEDS Plate #030

Visit Type: ☐ Screen ☐ BL ☒ Data Abs ☐ SCID

Patient ID #: ☐☐ ☐☐☐☐ ☐☐ (Centre # / Patient # / Admission #)

Visit Date: ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

NEEDS DATA ABSTRACTION FORM FOR MIGRAINE CLINIC (Page 1 of 3)

Reviewer: _____

Date: ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

☐ **Main Neurological Diagnosis:** _____

☐ **Date of Disease Onset (Year when headaches first started occurring):** ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

Migraine type:

☐ Migraine with aura

☐ Migraine without aura

Migraine Frequency Classification:

☐ Episodic (less than 15 days of headache per month for at least 3 months)

☐ Chronic (15 or more days of headache per month for at least 3 months)

Migraine Frequency Classification:

Number of events per day, month or year: ☐☐ per ☐ Day ☐ Month ☐ Year

Psychotropic Medications

Antipsychotics

Typical Antipsychotics

☐ Haldol/haloperidol

☐ Loxitane/Loxapac/Loxapine

☐ Moban/Molindone

☐ Prolixin/Fluphenazine

☐ Stelazine/Trifluoperazone

☐ Thorazine/Chlorpromazine

☐ Trilafon/Perphenazine

Atypical Antipsychotics

☐ Abilify/Aripiprazole

☐ Geodon/Zeldox/Ziprasidone

☐ Invega/Paliperidone

☐ Risperdal/Risperidone

☐ Seroquel/Quetiapine

☐ Zyprexa/Olanzapine

Other:

☐ Clozaril/Clozapine

version date: February 5, 2013

CRU-023-NEEDS Plate #031

Visit Type: ☐ Screen ☐ BL ☒ Data Abs ☐ SCID

Patient ID #: ☐☐ ☐☐☐☐ ☐☐ (Centre # / Patient # / Admission #)

Visit Date: ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

NEEDS DATA ABSTRACTION FORM FOR MIGRAINE CLINIC (Page 2 of 3)

Antidepressants

Tricyclic Antidepressants

- ☐ Anafranil/Clomipramine
- ☐ Elavil/Amitriptyline
- ☐ Norpramin/Desipramine
- ☐ Pamelor/Notriptyline
- ☐ Sinequan/Doxepine

SSRIs

- ☐ Celexa/Citalopram
- ☐ Lexapro/Cipralex/Escitalopram
- ☐ Luvox/Fluvoxamine
- ☐ Paxil/Paroxetine
- ☐ Prozac/Fluoxetine
- ☐ Zoloft/Sertraline

SNRIs

- ☐ Cymbalta/Duloxetine
- ☐ Effexor/Venlafaxine
- ☐ Pristiq/Desvenlafaxine

Buproprion

- ☐ Wellbutrin/Zyban/Bupropion

Other

- ☐ Desyrel/Trazodone
- ☐ Remeron/Mirtazapine

Benzodiazepines &Related Sedative-Hypnotics

- ☐ Ativan/Lorazepam
- ☐ Imovane/Lunesta/Zopiclone
- ☐ Klonopin/Rivotrol/Clonazepam
- ☐ Librium/Chlordiazepoxide
- ☐ Restoril/Temazepam
- ☐ Sonata/Starnoc/Zaleplon
- ☐ Serax/Oxazepam
- ☐ Tranxene/Clorazepate
- ☐ Valium/Diazepam
- ☐ Xanax/Alprazolam

Mood Stabilizers

- ☐ Epival/Depakote/Valproic Acid
- ☐ Lamictal/Lamotrigine
- ☐ Lithium
- ☐ Tegretol/Carbamazepin

version date: February 5, 2013

CRU-023-NEEDS Plate #032

Visit Type: ☐ Screen ☐ BL ☒ Data Abs ☐ SCID

Patient ID #: ☐☐ ☐☐☐☐ ☐☐ (Centre # / Patient # / Admission #)

Visit Date: ☐☐☐☐ ☐☐ ☐☐ (yyyy / mm / dd)

NEEDS DATA ABSTRACTION FORM FOR MIGRAINE CLINIC (Page 3 of 3)

Current Disease Specific Medications (please specify dose)

Preventatives

☐ Atacand/Candesartan	_____	☐ Flunarizine/Sibelium	_____
☐ Botox/Botulinum toxin	_____	☐ Inderal; Propranolol	_____
☐ Calan; Verelan/Verapamil	_____	☐ Indocid/Indomethacin	_____
☐ Cymbalta/Duloxetine	_____	☐ Lamictal/Lamotrigine	_____
☐ Effexor/Venlafaxine	_____	☐ Petadolex (Butterbur)	_____
☐ Lithium	_____	☐ Tegretol/Carbamazepine	_____
☐ Lyrica/Pregabalin	_____	☐ Tenormin/Atenolol	_____
☐ Magnesium	_____	☐ Trileptal/Oxcarbazepine	_____
☐ Neurontin/Gabapentin	_____	☐ Topamax/Topiramate	_____
☐ Pamelor/Nortriptyline	_____	☐ Riboflavin	_____
☐ Elavil/Amitriptyline	_____	☐ Other _____	_____
☐ Epival; Depakote/Valproic Acid/ Divalproex Sodium	_____		

Comments: _____

version date: February 5, 2013