2018

Depression Management in Outpatient Settings: A Systematic Review of the Literature

Modupe Mary Okonofua
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral study by

Modupe Okonofua

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Barbara Gross, Committee Chairperson, Nursing Faculty
Dr. Janice Long, Committee Member, Nursing Faculty
Dr. Riyad Naser, University Reviewer, Nursing Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

Depression Management in Outpatient Settings: A Systematic Review of the Literature

by

Modupe Mary Okonofua

MSN, Walden University, 2014

MBA, Dallas Baptist University, 2001

Project Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Nursing Practice

Walden University

August 2018

Abstract

Depression is a mental illness that requires prompt identification and treatment due to grave consequences if untreated. Depression can affect a person's level of functioning, lead to worsening health conditions, comorbid substance abuse, and suicide. Despite these facts, the current state of nursing practice includes an inadequate diagnosis of patients with depression, lack of guidelines for the use of assessment tools and diagnostic tests to identify depression, and insufficient information concerning the accuracy of depression assessment tools. This systematic literature review examined 6 depression assessment tools in regard to their accuracy as identified by specificity, sensitivity, reliability, and validity. This project also examined the pros and cons, demographics, and healthcare settings that use these depression inventory tools. This project used the Orlando nursing process theory as a theoretical framework. Based on the review of 10 articles selected, evidence showed that the Hamilton depression rating scale has the highest sensitivity (93%) and specificity (97%) rates. The implications for positive social change include the opportunity for clinicians to use the findings of this project in their selection of depression assessment tools in healthcare settings. Other researchers can use this project as a valuable resource for management of major depressive disorders.

Depression Management in Outpatient Settings: A Systematic Review of the Literature

by

Modupe Mary Okonofua

MSN, Walden University, 2014

MBA, Dallas Baptist University, 2001

Project Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Nursing Practice

Walden University

August 2018

Dedication

This project is dedicated first to my parents, Engineer Bola and Felicia Adesiyan, without whom I would not be in existence. I would also like to dedicate this work to my supportive husband and children: Dr. Chryss Okonofua, Annette, Naomi, and Rebecca. Special thanks to the Almighty God for all He alone has done in my life, for the strength to daily pursue dreams, for the talents He sowed in me, for the support system He surrounded me with, and for being a special loving father. I say thank you Lord for the completion of an educational pursuit. I love you all!

Acknowledgments

My dissertation chair, Dr. Barbara Gross, has been very helpful in completing this project, answering my questions that seemed sometimes minor and yet professional to direct me along smoothly. Special thanks to my husband and children for their emotional support, assistance in proofreading essays, prayers, and commitment to ensuring I completed my doctoral pursuits. To all my brothers and sisters, most especially Pastor Oluyemi Adesiyan, Ayomiku Olufunmi, and my entire church family, RCCG Tower of Love, I say thank you.

Finally, I would like to thank my Father, the Almighty God, the only creator not created, the one that loves me more than I could ever imagine, the one that blessed me with special talents, the ability to pursue life's dreams, a wonderful family, and good health. Thank you, Lord, and I love you!

I am solely responsible for any errors or omissions within this doctorate of nursing practice project.

Table of Contents

List of Tables

List of Figures

Section 1: Nature of the Project

Introduction

The nature of this project was a systematic review of the literature concerning the use of depression inventory tools and diagnostic tests for management of patients with a depression diagnosis. In this project, I reviewed the pros and cons, characteristics, and indications of depression screening tools published in peer-reviewed literature between 2012 and 2017. Physicians and nurse practitioners could use the compiled information from this project to objectively select an appropriate tool to assess and monitor recovery of patients with the diagnosis of depression in an outpatient setting.

Primary care doctors and clinicians working in an outpatient setting frequently rely on patient responses to endorse items on self-report measures to develop a diagnosis (Bieling, McCabe, & Antony, 2004). The diagnosis of depression can be made with depressed mood or anhedonia with at least four or more of the accompanying symptoms in the previous 2 weeks, affecting a patient's level of functioning at work, school, or home (American Psychiatric Association [APA], 2010). The diagnostic criteria for depressive disorder, also known as clinical depression or depression, includes a depressed mood or loss of interest or pleasure in daily activities lasting at least 2 weeks (PsychCentral, 2017). The accompanying symptoms include increased or decreased sleep, decreased interest, a sense of guilt or worthlessness, decreased energy or easily fatigued, inability to concentrate or make decisions, decreased or increased appetite and weight, increased or decreased psychomotor activity, and suicidal ideation (Institute of Clinical Systems Improvement, 2017). In the first section of this project, I review the

background of the project, practice problem, project-focused question, and framework guiding the project. I also address any limitations and the significance of the project in advancing the quality of care provided by clinicians by using depression assessment inventory tools.

Mental health disorders are a leading cause of disability in the United States (National Institute of Mental Health [NIMH], 2008). Recent statistics for depression among U.S adults are alarming. NIMH (2016) estimated that in 2015, 16.1 million adults aged 18 years or older in the United States had at least one major depressive episode, with this number representing 6.7% of all U.S. adults. The effects of depression can be self-limiting and devastating on patients, families, and the community as a whole. Depression is a mental health disorder ranked as a leading behavioral health disorder although it continues to be underrecognized and undermanaged in health care settings. Depression is projected to be the second largest cause of disability by the year 2020, but it is not often treated adequately (Maurer, 2012). Figure 1 shows depression prevalence among the U.S. adult population, whereas Figure 2 illustrates the prevalence of adult patients in the United States with severe impairment due to depression among different age groups and races.

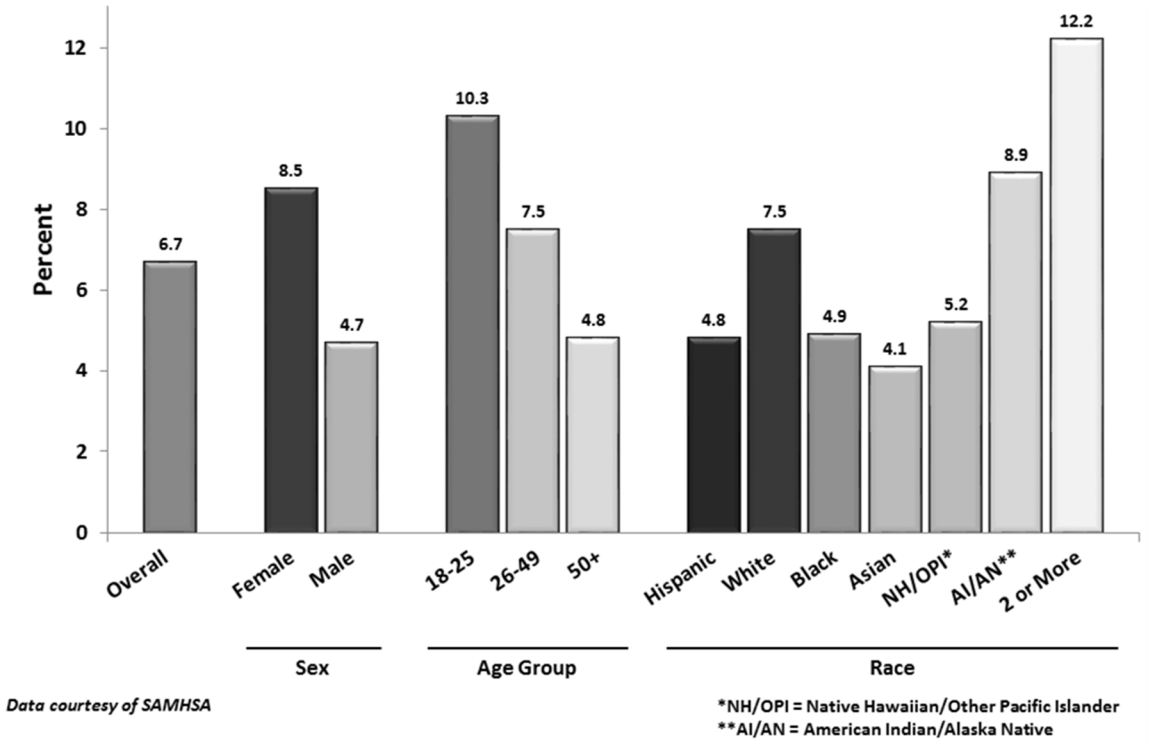

Figure 1. 12-month prevalence of depression among U.S. adults (2015). *Source*: NIMH (2016a). 12-month prevalence of major depressive episode among U.S. adults. *Notes*. Adapted directly from https://www.nimh.nih.gov/health/statistics/prevalence/major-depression-among-adults.shtml. No permission required to use the figure.

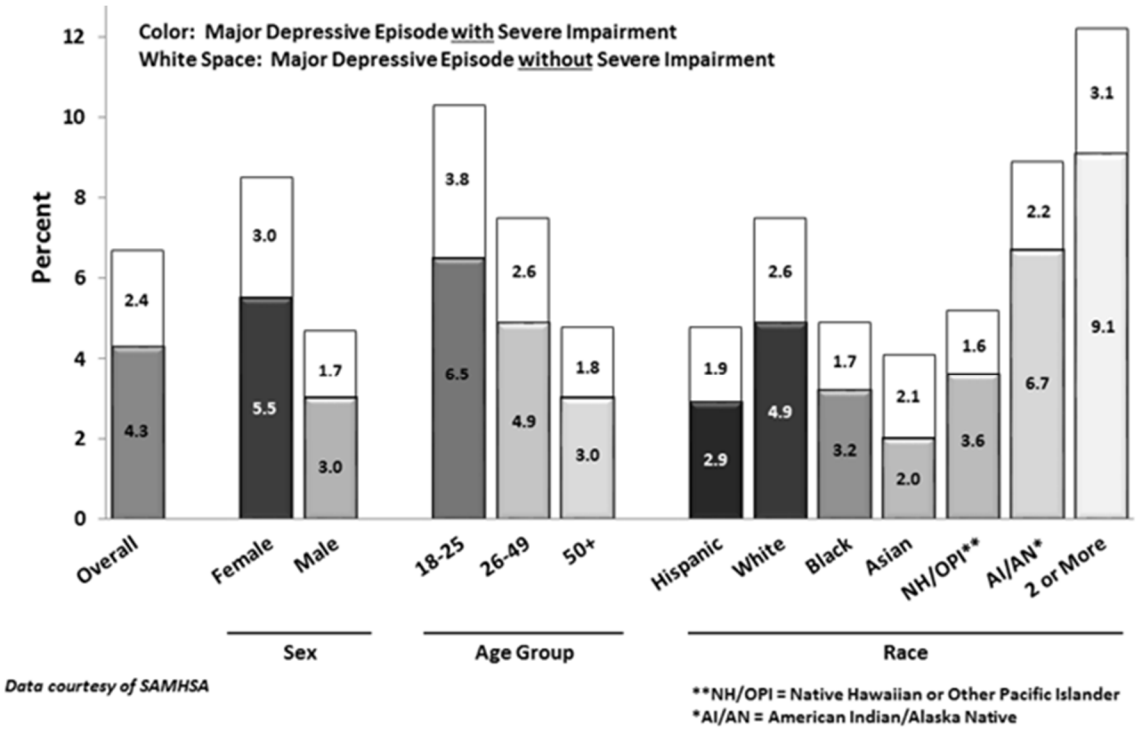

Figure 2. 12-month prevalence of major depressive episodes with severe impairment (2015). *Source*: NIMH (2016b). 12-month prevalence of depression with severe impairment. *Notes*. Adapted directly from https://www.nimh.nih.gov/health/statistics/prevalence/major-depression-with-severe-impairment-among-adults.shtml. No permission required to use the figure.

Although the effects of depression are of national concern, it is disheartening that rates of treatment remain low, and the treatment received is often inadequate (Pratt & Brody, 2014). My goals in this project included contributing to the existing literature concerning the management of depression from identifying the risk factors and present valuable information on assessment tools and diagnostic tests in a report that could be used by health care providers in clinical settings. Understanding the risk factors for depression is a step toward the ability to identify the causative factors and address the issues appropriately.

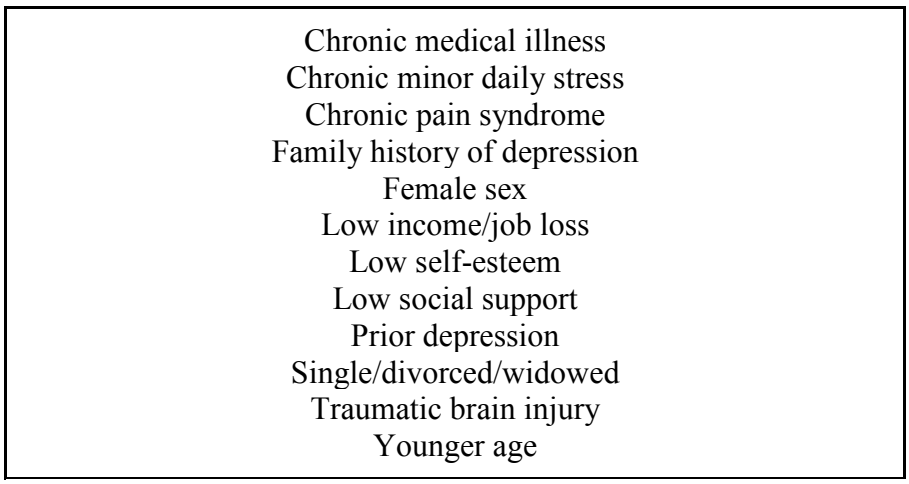

Figure 3. Risk factors for depression. Adapted from s*ource*: Maurer, D. M. (2012). Screening for depression. *Depression*, *100*, 23.

Maurer (2012) listed the risk factors of depression as chronic medical illness, chronic minor daily stress, chronic pain syndrome, family history of depression, female sex, low income/job loss, low self-esteem, low social support, prior depression, unmarried, divorced, or widowed, traumatic brain injury, and younger age. Therefore, being able to identify patients at high risk of developing this illness is crucial.

Depression can take different forms, but for the sake of this project, depression refers only to major depressive disorder (MDD) without psychosis. Therefore, this project excluded other forms of depression such as persistent depressive disorder, perinatal depression, bipolar depression, psychotic depression, and seasonal affective disorder. Differentiating MDD from other types of depression is significant regarding illness identification and applying appropriate treatment modalities.

The doctorate of nursing practice (DNP) nurse is often involved in the review of health care publications and functions as part of the team to develop evidence-based guidelines for his or her organization. This entails reviewing existing health care information for validity and reliability, which is one of the goals of this project. In short,

this project is a systematic review of the literature concerning the diagnosis of depression; this researcher studied the different assessment tools and diagnostic tests used in the management of a major depressive episode. Hence, this doctoral project was guided by the *Walden University Manual for Systematic Review of Literature*.

Multiple depression inventory tools can help diagnose depression in patients. No previous research has explored in detail whether one measurement or assessment tool has better predictability value compared with another or whether it functions in the company of other measures to increase diagnostic accuracy. This researcher reviewed existing evidence concerning assessment tools and diagnostic tests that could be used to improve health outcomes for patients with depression.

The need to accurately diagnose and manage patients with depression has never been more important than now. The effects of depression nationally and globally are significant. Maurer (2012) estimated the medical costs of depression to be $43 billion and put its price in low productivity at $17 billion annually. It is estimated to affect 5% to 13% of patients in primary care settings (Maurer, 2012). The effects of depression on patients, their families, and the communities have led to the need to screen patients accurately and manage their recovery objectively using appropriate measures. The development of a guideline using this systematic review of the literature improves diagnosis of patients with depression and makes a significant contribution to the existing knowledge base that could lead to positive social change.

Problem Statement

Louch (2008) indicated depression as a disease poses significant problems when not diagnosed and not treated effectively. The estimated medical cost of depression is $43 billion annually, and it costs economic productivity more than $17 billion annually (Maurer, 2012). Depression could lead to profound and overwhelming feelings of sadness that could negatively affect the patient's life, career, relationship with others, educational pursuits, and self-care habits. It can result in self-harm, irritability, anger, passing blame, and posing a threat to others. The feeling of depression can last a short period or for even a lifetime, and the economic implications on society can be enormous.

In terms of local relevance, the problem identified with typical practice in some health care settings includes subjective assessments with patients describing the level of depression using a numerical rating scale from 0 to 10, with 10 being severe depression, and patients describing accompanying signs, symptoms, and contributing factors to formulate the diagnosis and treatment plan (Okonofua, 2017). In practice, patients with a diagnosis of MDD or simply "depression" seen in primary care settings may not receive a diagnosis and treatment for depression. Therefore, having a guide for assessment and treatment of depression in the primary care setting is needed. For this project, the focus was on the identification of evidence-based information and evaluation tools for the diagnosis and management of depression in a primary care setting. I expected to address the difficulty in finding any single publication covering measures and diagnostic tests for MDD in a primary care setting.

Implications of depression can be severe and costly. According to the Centers for Disease Control (CDC; 2017), "In the United States, the economic burden of depression in 2010 was $210 billion, which includes the cost of workplace displacement, direct management, and treatment of depression and depression-related suicides or attempted suicides" (p. 1). Depression increases the risk of suicide-related morbidity and mortality, workplace absenteeism, lower productivity, anxiety, irritability, smoking, substance use, and eating disorders (CDC, 2017). The effects of depression on individuals, families, and the society as a whole are enormous, and accurate diagnosis and treatment are of utmost importance. Effects of depression have also been linked to self-destruct behaviors such as alcohol use, use of illicit substances, and misuse of prescribed medications. Davis, Uezato, Newell, and Frazier (2008) explained that nearly one-third of patients with MDD also have substance use disorders, which leads to a higher risk for suicide, greater social and personal impairment, and a higher rate for psychiatric conditions.

DNP nurses are at the forefront of using evidence-based resources to transform health care, design treatment protocols, and improve health outcomes for patients. Depression is an illness that, when undiagnosed or untreated, can have grave consequences for the patients, families, and the community as a whole. Depression is an illness that has implications at local, national, and international levels.

This doctoral project holds significance for the field of nursing practice. Depression demands long and costly treatment protocols. Mainly, treatment is inadequate; even when appropriate treatment is applied, there is incidence of recurrent episodes that are also expensive with attendant personal and societal socioeconomic

dislocations. Maurer (2012) stated depression is often not adequately treated, and even when treated appropriately, more than 75% of patients with depression have recurrent episodes, and 10% to 30% have residual symptoms.

The ability to quantify the level of depression or residual depression symptoms following treatments depends on the use of appropriate assessment tools. This doctoral project is significant to the field of nursing practice in more ways than one. DNP nurses play a significant role in health care research, health-related problem identification, proposing solutions, health care policy formation, and health care advocacy. Advanced practice nurses are leaders in health care practice who review problems within health care settings while formulating the appropriate solutions.

Purpose

In this DNP project, I systematically reviewed existing literature, compiled information on accuracy, and examined pros and cons of different depression inventory tools and diagnostic tests for management of depression with the goal of improving quality of care provided to patients by clinicians. Likewise, this doctoral project holds significance for the field of nursing practice in that there is a gap in the practice in the outpatient setting regarding the lack of use of assessment tools. I reviewed some depression scales that have been used to diagnose and treat patients with depressive disorders in a primary care setting. These assessment tools included the Beck depression inventory (BDI), Hamilton depression rating scale (HDRS), patient health questionnaire (PHQ-2 and PHQ-9), a major depression inventory (MDI), Zung self-rating depression scale (SDS), and Center for Epidemiologic Studies depression scale (CES-D).

I compiled current evidence concerning the use of depression assessment tools regarding their accuracy (identified by specificity, sensitivity, reliability, and validity), their advantages, disadvantages, possibly demographics, and health care settings from existing literature. I also reviewed basic lab tests that could be used to rule out medical illnesses that can present as depressed mood. The final product of the project was depression treatment resource material used in primary care settings.

The practice-focused question was: What are the characteristics and indications of the major depression screening tools/diagnostic tests published in peer-reviewed literature between 2012 and 2017 for adult patients?

The use of a population, intervention, comparison, and outcome (PICO) question for this project increases "the likelihood that the best evidence to inform practice will be found quickly and efficiently" (Stillwell, Fineout-Overholt, Melnyk, & Williamson, 2010, p. 58). The PICO for this project was as follows:

- P (population): adult patients in an outpatient clinic from ages 18 to 50 years,
- I (intervention): use of assessment tools and recommended diagnostic tests to diagnose depression,
- C (comparison): non-use of assessment tools and recommended diagnostic tests to diagnose depression, and
- O (outcome): improved health outcomes for patients with depression.

There is an existing gap-in-practice whereby management of depression lacks an easily accessible guide that compares the reliability of the depression inventory tools or diagnostic tests. The findings of this doctoral project contributed to filling this gap-in-

practice. The compilations of the depression assessment tools and diagnostic tests in a tabular form make the results of this project readily available for health care professionals, clinicians, patients, and family members to read and use for the management of depression.

Nature of the Doctoral Project

This project was a systematic review of the literature examining published articles on depression assessment and treatment in primary care settings from the years 2012 to 2017. Studies involved patients of mixed race, males, and females between ages 18 to 50 years. Sources of evidence included Walden University Library databases, online scholarly articles, and search engines such as Google Scholar. Once I completed the systematic review of the literature, I provided the findings to the clinical setting as an educational session on depression assessment tools and diagnostic tests for the management of depression diagnosis.

I tabulated all diagnostic and assessment tools identified and reviewed using a Microsoft Excel spreadsheet. The report included several headings including depression assessment tools, the number of articles reviewed, the accuracy of the tools regarding sensitivity, specificity, reliability, validity, age groups, or demographics for using the assessment tool, type of health care setting, and advantages and disadvantages of the assessment tools as identified from existing literature.

Significance of the Project

The stakeholders affected by this project included patients, families, health care providers, and the community as a whole. I sought to review assessment tools that would

accurately, effectively, and efficiently diagnose patients with depressive disorders on specificity, sensitivity, and objectivity. The identification of accurate assessment tools assists in the management of diagnosis and treatment of depression in primary care settings. This study contributes to the body of evidence-based resources for the care of patients with depressed mood.

For the patients, this project produced a compilation of assessment tools for the diagnosis of depression best fitted for certain age groups or demographics. For the health care providers and patients, the project helped identify the advantages and disadvantages of tools used for depression diagnosis, and the project helped answer the question of whether the assessment tools accurately measured depression among patients.

The project is valuable to health care professionals, businesses, and associations at local, national, and global levels. Depression is an illness identified as a global concern. It is the fourth leading cause of disability worldwide, and by 2020, it will be the second leading cause according to World Health Organization projections (WHO, 2008).

This project is of importance to nursing practice and applies to other health care professions. The potential of transferability of this project to related practice sectors such as psychology and public health fields is high. There are also potential implications of positive social change using the findings of this doctoral project. The use of results of this project could lead to a change in behavior and a change in practice modalities of health care professionals or clinicians. Furthermore, these objective criteria could serve as a valuable tool for patients to understand how to manage their illness accurately and more objectively.

Summary

Effective diagnosis and management of depression can be complicated. Identifying accurate, effective, and efficient assessment tools and diagnostic tests assists in the proper identification and management of depression in patients. This project is a compilation of depression assessment tools with a view to identifying their accuracy regarding their specificity, sensitivity, reliability, validity, advantages, and disadvantages. This project concludes with objective criteria for making a correct diagnosis of depression in a primary health care setting.

Section 1 of this project covered the introduction, problem statement, research questions, relevance of addressing this topic locally, and the project's relevance to the field of nursing practice. I also identified the meaningful gap-in-practice and how I hoped to fill the gap, the sources of evidence, the plan on how to compile findings, and a statement on how the project could be of significance to other similar practice areas.

In the next section, I address the background and context of the doctoral project, concepts, models or theories, and the relevance of the doctoral topic to nursing practice, and I cite existing scholarship.

Section 2: Background and Context

Introduction

Depression is a serious mental illness that is treatable, yet, can involve disabling life events or can lead to suicide. Among patients diagnosed with depression, the rate of suicide is approximately 10% to 15% (Lichtblau, 2011). The need to address this silent epidemic has never been as urgent as now, demanding the attention of health care providers, health care organizations, the government, patients, and families. The absence of guidelines for assessment and management of patients diagnosed with depression complicates this practice problem because few studies have been conducted to ascertain the adequacy of assessment tools used for depression.

For this project, the practice-focused question was: What are the characteristics and indications of the major depression screening tools published in peer-reviewed literature between 2012 and 2017?

The purpose of this study was to examine the literature systematically to determine the accuracy, pros, and cons of assessment tools used for depression. The project also included basic lab test reports that could be used to rule out medical illnesses, which can present as depressed mood. To this extent, this project was a systematic review of the literature of published depression assessment tools and diagnostic tests for the treatment of depression in a primary care setting. The depression assessment tools were BDI, HDRS, and PHQ-2 and PHQ-9. Other depression assessment tools to be covered included MDI, SDS, and CES-D.

The management of depression lacks an easily accessible guide that can be used to rate the level of depression and rule out medical illnesses that resemble depression. Up to now, "Current psychiatric nursing practice remains grounded in tradition, unsystematic trial and error, and authority; although some of the wisdom that has been passed down over time is questionable, it continues to influence nursing practice today" (Zauszniewski, Bekhet, & Haberlein, 2012, p. 3).

In Section 2 of this doctoral project, I address the background/context of the doctoral project, concepts, models, and theories, as well as the relevance of the doctoral topic to nursing practice, and I cite existing scholarship. The final report findings are a compilation of assessment tools and diagnostic tests for depression.

Concepts, Models, and Theories

In this DNP project, I used an appropriate nursing theory guided by the *Walden University Manual for Systematic Review*. The goal of this systematic review of the use of assessment inventory tools and diagnostic tests in the management of depression was to generate guidelines for best clinical practice. The identified problem was the lack of standardized instruments in some outpatient practices when caring for patients with the depression diagnosis. The goal was for clinicians to be able to use the product of this project as a guide to inform treatment and choice of depression inventory tools and diagnostic tests. Hence, in this project, I systematically reviewed existing literature and synthesized the findings concerning treatment and management of depression.

I used Orlando's nursing process theory (ONPT). ONPT was formulated by Ida Orlando, having monitored the interaction between nurses and patients. ONPT serves as a

guide for the nurse to use in the nursing process (i.e., assessment, diagnosis, planning, intervention, and evaluation) when caring for patients. Faust (2002) identified the purpose of the ONPT to include "encouraging the use of the deliberative process, improving the interaction between the nurse and patient, perception validation, and the use of the nursing process to produce positive outcomes or patient improvement" (p. 14).

The primary element of this theory is that nurses should be able to provide "immediate explorations of patient's perceptions, thoughts and feelings when ill" (Orlando, 1987, p. 405). This theory propagates that patients' cries for help indicate a level of distress that can be analyzed and properly addressed by the nurses' appropriate interventions. The patient's presenting behavior might be a cry for help, yet, the patient's needs may not be what they appear to be (Orlando, 1987). Because of this, nurses must use their perception, thoughts about perception, or the feeling engendered from their thoughts to explore the meanings of the patient's behavior (Orlando, 1987). This process helps nurses determine the nature of the patient's distress and provide the help he or she needs. The theory conceptualizes professional nursing functions, the process of nursing discipline, reactions that follow immediately, and improvements to the status quo (Orlando, 1987).

Nursing process deals with the necessary steps to care for a patient starting with assessment, followed by diagnosis, planning, intervention, and evaluation. In the 1950s, nursing introduced a three-step process including observing, measuring and gathering data, and analyzing the findings (Doenges & Moorhouse, 2012). The first step of the nursing process starts with a complete, thorough, and accurate evaluation of patient's

concerns and health issues. The need for an accurate assessment is significant. An inaccurate assessment certainly results in incorrect overall nursing processes and adverse clinical outcomes. The major dimension of ONPT is that the nurse should inquire about patient concerns using appropriate tools or guidelines. The need for accurate assessment as identified by this theory is a reflection of the goal of this project. To conduct an accurate assessment, nurses need to use valid tools and their perceptions from interacting with patients and intuitions. The use of relevant tools improves communication between patients as well as providers, and it ensures accurate diagnosis and assessment of depression severity.

Theories Related to Depression

Concerning causes of depression, there are several theories proposed, and these relate to underlying causes of depression. These theories relate to neurotransmitter deficiencies such as serotonin (5-HT), norepinephrine (NE), and brain-derived neurotrophic factor, as well as stress. According to Lichtblau (2011), "Depression is due to a deficiency in some aspects of either NE or 5-HT activity in the brain" (p. 22). This theory also assumes stress can trigger interference with the gene that regulates a neurotrophic factor derived in the brain. The neurotrophic factor is a critical neurochemical that helps keep the neurons viable. It has also been described as "fertilizer for the brain cell" (Lichtblau, 2011, p. 22). In short, these theories are relevant to this topic, as the project addresses management of depression.

Key terms

Key terms used in this dissertation include the following:

- *assessment tools*: resources used to evaluate the incidence, occurrence, or severity of an illness;
- *depression*: the state of being downcast or a state of severe sadness that can impair the patient's level of functioning at work, school, or home;
- *diagnosis*: the ability to make an assertion of an illness presented by a patient's symptoms;
- *neurotransmitters*: chemical messengers produced by neurons that allow carrying messages or signals from one neuron to another;
- *sensitivity*: the ability of a tool to accurately test positive for an illness; sensitivity tests are used to rule out other diseases (i.e., the ability of the test to correctly identify those with the disease [true positive rate]); and
- *specificity:* how often the test will be negative when the patient does not have the disease; specificity, in short, rules out those without the disease (true negative rate).

Relevance to Nursing Practice

In nursing practice, the need for accurate assessment when caring for patients is crucial. It demands emphasis and utmost attention because a misdiagnosis can commence from inaccurate assessment. Multiple health care organizations have in place recommendations concerning the use of evaluation tools in depressed patients. According to APA (2010), "there must be an assessment tool used for the initial diagnosis and follow-up visit of these patients to determine progress or effectiveness of medications" (p. 4).

The ability to manage any illness, medical or psychiatric, is vital. Evidence-based guidelines used for management of any diseases are available and most can be found online or in print form. The current state of nursing practice related to this issue is an inadequate diagnosis of patients with depression and lack of guidelines for the use of assessment tools and diagnostic tests in patients with depressed mood. According to Zauszniewski et al. (2012), "Psychiatric nursing is still very much tradition-based, with its accompanying trial and error and authority methods and yet this tradition continues in spite of the questions it has generated over the years" (p. 1).

The tools necessary for an adequate diagnostic assessment of psychiatric illness include questionnaires, rating scales, standardized data forms, and structured interviews (APA, 2010). These instruments can be tools for diagnosis, parameters for measuring social or vocational functions, or for monitoring changes in severity of symptoms and the side effects of treatments. These instruments could also assist psychiatrists to ask the relevant questions during diagnostic interviews. The current state of nursing practice shows the use of assessment tools is not widespread in primary care settings (Zauszniewski et al., 2012). This gap-in-practice can be addressed with the use of appropriate tools and guidelines when managing patients with depression.

Several national associations have advocated for the use of assessment tools in the management of depression. Current research data on new practices or tools for addressing the problem as identified in this project is minimal, and most literature still focuses on the use of assessment tools such as BDS and PHQ-9 assessment instruments. However, the latest recommendation by the U.S. Preventive Services Task Force (USPSTF; 2016) is in

line with the APA's 2010 recommendation. The USPSTF (2016) recommends "screening for depression in the general adult population, including pregnant and postpartum women. Screening should be implemented with adequate systems in place to ensure accurate diagnosis, effective treatment, and appropriate follow-up" (p. 1).

Depression can be a disabling illness, affecting patients irrespective of sex, race, ethnicity, financial circumstances, occupation, religious beliefs, or age. According to Bromet et al. (2011), "major depression is a severe, recurrent disorder linked to diminished role functioning and quality of life, medical morbidity, and mortality" (para. 1). People dealing with depression often use descriptors such as:

> stressed out, unhappy, sad, melancholic, miserable, sorrowful, woeful, gloomy, despondent, hopeless, upset, tearful, in the dumps and having the doldrums. Other terms used for depression include low spirits, having a heavy heart, in despair, desolate, hopeless, upset, tearful, in the dumps, having the blues or in a funk. (Lichtblau, 2011, p. 22)

With multiple words that can be used to describe this mental health issue and available treatment options, depression continues to be a major problem seen in primary care and is the fourth most common complaint encountered in such settings, which translates to approximately 1 in every 10 patients seen (Lichtblau, 2011).

According to Uman (2011), "with an ever-increasing plethora of studies being published in the health studies, it is challenging if not impossible for busy clinicians and researchers alike to keep up with the literature" (p. 57). This doctoral project advances nursing practice and fills the gap in existing literature concerning an easy to use set of

guidelines that shows the accuracy, pros, and cons of assessment tools and diagnostic tests in the management of depression. It is a resource tool to improve the standard of care or act as a guide in planning depression treatment protocol in the outpatient setting. According to Uman (2011), "reviews summarizing the outcomes of various interventions trials are an extremely efficient method for obtaining 'the bottom line' about what works and what does not" (p. 57).

Depression Assessment Tools

There is a summary of the assessment tools as appeared in this project. The depression assessment tools reviewed for this project included the following: BDI, HDRS, PHQ-2 and PHQ-9, MDI, SDS, and CES-D.

Beck depression inventory. The BDI is a widely used scale created by Aaron T. Beck. Applicable to both clinical and non-clinical settings, BDI is known to be one of the most extensively researched and utilized self-report depression inventory scales (Campbell, Maynard, Roberti, & Emmanuel, 2012). A more modernized version, the BDI-II, is a 21-item self-rating tool designed for individuals from ages 13 and upward and can detect as well as measure level of depression severity. There are three versions of this scale—the original BDI, created in 1961, the revised BDI-1A from 1978, and the BDI-ll (Beck, Steer, & Brown, 1996). The BDI divides the components of depression into two subscales. One of them is the affective component, which affects the emotional state of mind. The second subscale consists of the physical (i.e., somatic) component, which affects the physical body. The reason for these seemingly unwarranted subscales is to aid in the determination of the original cause of the noted depression.

The two articles related to BDI chosen for this project are:

Article 1. Campbell et al. (2012) compared the psychometric strengths of two depression assessment tools—Zung SDS and BDI-II. The sample size was 415 undergraduate student volunteers with a mean average age of 25.2. The researchers' goals include providing clinicians and scholars with information concerning the two depression tools and providing psychometric evidence concerning the tools' reliability when used in patients with depression. The summary of findings for this article indicates the BDI-II demonstrated marginally superior internal consistency, reliability, and psychometric properties to the Zung SDS (Campbell et al., 2012). However, this article did not review sensitivity or specificity of the assessment tools.

Article 2. Jakšić, Ivezić, Jokić-Begić, Surányi, and Stojanović-Špehar (2013) examined the diagnostic validity of BDI-II. The population size was 314 from a medical outpatient setting. Patients were notified of the project by their primary care provider, and the willing ones volunteered. The final finding from this article was that the BDI-II had a high sensitivity and high specificity rating with a satisfactory diagnostic validity in differentiating between healthy and depressed individuals in this setting (Jakšić et al., 2013). The sample population used ranged from ages 25 to 87 with different acute illnesses, and I monitored them over a period of 2 months. One limitation of this article stems from the disclosure of intent of the survey to participants while there are no details about the non-willing population.

Hamilton depression rating scale. In 1960, Max Hamilton developed this scale to be utilized in clinical trials and to monitor the effectiveness of antidepressants versus

placebo (Zimmerman, Martinez, Young, Chelminski, & Dalrymple, 2013). In 1967, it was revised. Hamilton later created the Hamilton depression inventory and the Hamilton anxiety scale. The Hamilton depression scale (HAM-D) is a questionnaire consisting of 17 to 21 items used to assess treatment response for patients previously identified with a depressive disorder. It differs from other scales since its use is the evaluation of a patient's depression before, during, and after his or her personalized treatment. Estimated at 20 minutes, the scale is typically used by a clinician during the patient visit and depends mainly on the skill of the interviewer (Zimmerman et al., 2013).

The two articles related to the Hamilton depression inventory are:

Article 1. Raimo et al. (2015) carried out a study of psychometric properties of HDRS in multiple sclerosis. HDRS is a semi-structured interview consisting of 17 items assessing the whole spectrum of depressive symptoms, including affective, cognitive, and somatic symptoms. In this study were 100 patients among those who attended a multiple sclerosis center in a hospital setting. Other patients were excluded from the survey based on criteria such as general intellectual decline, as identified by age and education-adjusted score lower than 23.8 on mini-mental state examination. In this study, selected ages were between 22-68 years with a mean of 43.3, with multiple sclerosis onset age at between 16-54 years. The educational year was between 5-19 years. The area under the curve was .988, which indicated the good discriminant power of the test. The achieved score of 14.5 provided the best tradeoff between sensitivity (93.33%) and specificity (97%). The advantage of the Hamilton depression rating score is that it has good

psychometric properties in assessing depression and, in this case, reliably tested patients with multiple sclerosis and depression.

Article 2. Schneibel, Brakemeier, Wilbertz, Dykierek, Zobel, and Schramm (2012) carried out research entitled: Sensitivity to detect the change and the correlation of clinical factors with the HDRS and the BDI in depressed inpatients. The authors carried out a study of 105 hospitalized patients with mean age of 41.6 being managed for depressive disorder to determine the discrepancies and the potential prediction between HDRS and BDI to analyze their sensitivity to detect change. HDRS showed a superior sensitivity to change compared to BDI. This study was able to regard HDRS as two complementary rather than redundant or competing instruments, though, small sample sizes in the subgroups provided results that could not be generalized. Also, the reliability of the HDRS rating was obtained using independent interviews of the same subject.

Patient Health Questionnaire. The PHQ-2 and PHQ-9 offer self-assessment tools in diagnosing depression. PHQ-2 is a two-question short assessment that can be used to screen for depression. The first two items of the PHQ-9, which question the degree to which a person has experienced anhedonia and depressed mood within the past 2 weeks, make up the PHQ-2 assessment tool. Alternatively, the PHQ-9 is a nine-item questionnaire that provides the clinician or the client a tool to assess for depression.

Article 1. Manea, Gilbody, and McMillan (2012) conducted a bivariate meta-analysis on the optimal cutoff score for diagnosing depression using the PHQ-9 across a range of 18 studies. The study had a total of 7,180 participants with age ranging from 24.8 to 71.4 years. With a cutoff score between 8-12, the PHQ-9 positively detected

MDD among the population selected. These authors recommended using the assessment tool in conjunction with other diagnostic criteria such as DSM-V or International Classification of Diseases-10 to make a diagnosis of depression, and the authors did not recommend the same cutoff score for all settings.

Article 2. Manea et al. (2016) reviewed the use of the brief PHQ-2 in the assessment of depression in a clinical setting. PHQ-2 is used to assess two critical elements of depression: a loss of interest or pleasure or anhedonia and low mood. PHQ-2 has been found to be very effective in the busy clinical setting due to the limited number of questions. Using bivariate diagnostic meta-analysis to derive sensitivity and specificity, the authors reviewed 21 studies with a total population of 11,175. The authors recommended reading the results of this test with caution because of the high risk of false-positive rate. The sensitivity at the cutoff mark of 2 is 0.91 and specificity 0.70. Composed of the first two questions of the PHQ-9, the PHQ-2 has been found to have a high accuracy level and continues to be a useful tool in screening for depression (Manea et al., 2016).

Major Depression Inventory. MDI is a "self-report measure for depression based on the DSM-system" (Cuijpers, Dekker, Noteboom, Smits, & Peen, 2012, p. 1). This tool can be utilized by clinicians and health professions to diagnose and assess the severity of depression based on DSM-IV and ICD criteria.

Article 1. Bech, Timmerby, Martiny, Lunde, and Soendergaard, (2015) carried out a study to evaluate the standardization of MDI as a tool for assessment of depression. I used data from two formerly published studies in which patients in MINI

neuropsychiatric interviews had verified the diagnosis of DSM-IV major depression. MDI is a standardized tool for measuring depression with a cutoff score of 21, 26, and 31 for mild depression, moderate depression, and severe depression respectively. From this article, the MDI assessment tool yielded a specificity of 82% and sensitivity of 90% for DSM-IV major depression.

***Article 2*.** Fawzi, Fawzi, and Abu-Hindi carried out this research in 2012. The basis was the Arabic version of the MDI as a diagnostic tool to test the reliability and concurrent validity among 50 Egyptians outpatients with MDD compared with 50 healthy Egyptians in the control group. Sensitivity was 88.4% while specificity was 78.9%. The study had excellent reliability and acceptable concurrent and discriminant validity. However, the shortcomings of this study included the small sample size, the individuals involved in the research were aware of the diagnosis of the participants, and the patient group consisted of only patients who met the diagnostic criteria for MDD.

The Zung self-rating scale. Created by a psychiatrist from Duke University, named William W. K. Zung (as cited in Romera, Delgado-Cohen, Perez, Caballero, & Gilaberte, 2008), the Zung self-rating scale provides patients the opportunity to "self-rate" their level of depression. The test is composed of 30 items that judge patients' psychological and somatic depressive indications. The process consists of 10 positive and 10 negative questions with each response noted on a scale of 1-4. At the end of the inquiries, the patient's score is deposited between four ranges: 20-44 (normal), 45-59 (mild depression), 60-69 (moderate depression), 70 and above (severe depression; Romera et al., 2008).

Article 1. Campbell et al. (2012) conducted research with 415 students (75% female, 25% male, mean age 25.2 years) with a goal to compare the psychometric strengths of the Zung SDS and the BDI-II. The research also planned to provide researchers and clinicians interested in measures of depression with psychometric evidence that differentiated the two instruments in an academic setting. Though this study was not designed to evaluate specificity and sensitivity, it was able to make the following resounding conclusion that BDI-II demonstrated marginally superior internal consistency reliability ($\alpha = 0.88$) than the Zung SDS ($\alpha = 0.85$). Zung SDS was less psychometrically sufficient, and the wording was problematic. This study is a valuable tool in the assessment of adolescents and young adults in large treatment settings. The disadvantage of this study is it is limited to the sample size and cannot be generalized beyond the university settings. Also, the whole package of BDI-II is quite expensive for clinicians with limited resources.

Article 2. Ruiz-Grosso et al. (2012) carried out a study of 70 patients in a clinical setting to validate the Spanish version of ZSDS. The accuracy of the tool regarding sensitivity and specificity is as follows: The cutoff scores for the highest proportions of correctly classified individuals among the major depressive episode with evidence of other psychiatric disorders (OPD) was ≥ 47, sensitivity = 85.7%, specificity = 71.4%. However, a major depressive episode with no evidence of psychiatric disorders was ≥ 45, sensitivity = 91.4%, specificity = 91.8%. This study revealed the ZSDZ including the Spanish version to be a valid assessment tool to diagnose depression in medical settings. Also, it is useful for both epidemiological research and primary clinical settings.

Center for Epidemiological Studies Depression Scale (CES-D). This is a tool used for the general population to assess recent onset of depression. According to VanDam and Earleywine (2011), "CES-D is a popular assessment tool that has wide applicability in the general population with the goal to reflect modern diagnostic criteria and improve upon psychometric limitations of its predecessor" (para. 1). The CES-D questionnaire has 20 questions with a possible score of 0-60 points with a score of 16 or over indicative of depressed mood.

Article 1. Chin, Choi, Chan, and Wong (2015) conducted research with 3,686 Chinese adult primary care patients. CES-D was sensitive in detecting the area under the curve > 0.7. The CES-D showed to be a valid, reliable, sensitive, and responsive instrument for screening and monitoring depressive symptoms in adults. Cross-cultural comparisons of depression in multi-center studies also support this instrument.

Article 2. Levine (2013) carried out a study to evaluate a seven-item CES-D short form (SF) using data from the National Longitudinal Survey of Youth. The participants in the 20-item CES-D questionnaire were as follows: (n = 8858) in 1992, to the 7-item CES-D in 1994 (n = 8500) and from 1998 to 2010 if aged 40 (n = 7072) or 50 (n = 1574) or over. Regarding the reliability of the two versions of the CES-D, "a CES-D-SF cutoff score ≥ 8 had acceptable specificity (0.97, 95% CI 0.96, 0.97) and modest sensitivity (0.69, 95% CI 0.67, 0.71) with the standard CES-D cutoff score of 16" (Levine, 2013, p. 1519). With this study, CES-D SF was relevant in assessing for a MDD.

Local Background and Context

APA (2010) recommended the integration of measurement into psychiatric management. There is a need to use appropriate measures such as assessment tools and diagnostic tests when managing patients with depression in an outpatient setting. The use of routine assessments of patients with depressive symptoms in clinical practice can help provider evaluate the progress of depression and can facilitate cooperation and communication between the provider and patients.

Local background of the topic stems from the lack of use of assessment tools in some outpatient behavioral health clinical settings (Okonofua, 2017). In these settings, patients are simply asked to assess the level of their depression using a verbal self-rating technique similar to pain scale assessment. However, this could be a problem. Therefore, health professionals should implement better assessment tools in the outpatient setting. APA (2010) suggested clinicians should integrate rating scales into the initial and ongoing evaluation. The identification and urgent treatment of depression is a local, national, and global issue addressed by national and international. Hence, the need for rapid diagnosis and subsequent treatment necessary to mitigate the negative effects of depression is imperative (NIMH, 2008).

Role of the DNP Student

The doctorate of nursing practice degree is one that prepares advanced practice nurse clinicians with the skill set necessary to apply evidence-based resources to solve practice problems and implement better health care models. The DNP clinician is also able to advocate for better health care policies, serve as an expert clinician utilizing best

practices in assessing, diagnosing, and treating patients with illnesses and most especially, can help shape the best health care practices. The DNP program is such that it develops the necessary competencies in the student. This project helped develop my competencies as an expert nurse practitioner in behavioral health and assisted me in evaluating the quality of health care services. This project also provided the avenue to highlight deficiencies in the use of assessment tools in the management of patients with a diagnosis of depression.

As a doctorate in nursing practice student, one of my professional goals is to contribute to the nursing profession. This doctoral project was a training avenue for me to perfect my scholarship skill set and provide contributions that are useful for managing patients with a diagnosis of depression. My role in this project was that of the author and researcher. This doctoral project was a systematic review of existing literature concerning the management of depression. Hence, the project had no participants or experiments and utilized high-quality resources that were relevant to the topic.

The writer's motivation for this doctoral project stemmed from personal experiences where thorough workups of patients with depressed mood have led to different diagnoses and treatment. Multiple health issues such as hormonal imbalance mimic depression; therefore, using appropriate tools for assessment and relevant guidelines to rule out medical problems was important. As an advanced practice nurse working in psychiatry, it is common to meet patients being treated for depression though their problem is a different health issue. Treating patients with an incorrect diagnosis can lead to poor clinical outcomes, patient suffering, and wasted resources.

The use of assessment tools in managing depression is very important, and for this project, several biases within existing literature were possible. According to APA (2010), one may encounter biases related to culture, ethnicity, gender, society, and age when one uses an assessment tool or interprets the results. This DNP student reviewed the literature for any identifiable biases and issues that could affect the results compiled and identified no potential personal biases.

Role of the Project Team

This project was a systematic review of the literature and consisted of no project team. This DNP student was responsible for the topic, the plan of conducting the literature search, compilation of findings, and final submission of doctoral write-up. The committee chair and members contributed their expert opinions and guidance toward completion of this project.

Summary

The gap-in-practice as identified by this doctoral project was one that deserved adequate attention from health care professionals and clinicians. It is important to note the primary ingredient of diagnosis cannot be assessment tools. According to APA.org (2010), "clinical impressions of treatment response should consider the relative importance of specific symptoms to the patient's function and well-being and the relative effects of specific symptoms on the patient's social environment" (p. 3). Hence, "rating scales should never be used alone to establish a diagnosis or clinical treatment plan; they can augment but not supplant the clinician's evaluation, narrative, and clinical judgment" (APA.org, 2010, p. 1). Section 2 covers the theories applicable to this project, the

relevance of the project to nursing practice, local background and context, and the role of the DNP student and project team. The next section focuses on collection and analysis of evidence as identified in the literature, the sources of evidence and strategies for analysis, and synthesis of evidence collated.

Section 3: Collection and Analysis Evidence

Introduction

The purpose of this study was to examine the literature systematically to determine the accuracy, pros, and cons of assessment tools used for depression. The project also included basic laboratory tests report that could be used to rule out medical illnesses that can present as depressed mood. Hence, this DNP project was a systematic review of the literature of published articles related to depression assessment tools and diagnostic tests for the treatment of depression in a primary care setting. The depression assessment tools reviewed were BDI, HDRS, PHQ- 2 and PHQ-9, MDI, Zung SDS, and CES-D.

The previous sections dealt with the project introduction, problem statement, purpose of the doctoral project, and nature of the project. I also reviewed the significance of the project, the applicable theories, the relevance to nursing practice, local background and context, the role of the DNP student, and the project team. Section 3 deals with collection and analysis of evidence as identified in the literature, the sources of evidence and strategies for analysis, and synthesis of evidence collated.

Practice-Focused Question

I reviewed existing literature for depression assessment tools identified earlier for their accuracy in terms specificity and sensitivity, advantages and disadvantages, demographics of patients, and the health care settings. I also listed diagnostic tests that could be performed to rule out medical illnesses that mimic depression. The final product of the project was a depression treatment guideline for use in primary care settings.

The practice-focused question was: What are the characteristics and indications of the major depression screening tools published in peer-reviewed literature between 2012 and 2017?

There is an existing gap-in-practice whereby management of depression lacks an easily accessible guide that can be used to rate the level of depression and rule out medical illnesses that resemble depression. The findings of this doctoral project could contribute to the gap-in-practice. The compilation of the depression assessment tools in a tabular form make it readily available for health care professionals, clinicians, patients, and family members to read and use for management of depression. Zauszniewski et al. (2012) explained that current psychiatric nursing practice is still grounded in tradition, nonsystematic trial and error, and wisdom that are passed on over time.

In this section of the project, I review the sources of evidence, analysis, and synthesis of the research findings, and I identify the systems used for recording, tracking, organizing, and analyzing the evidence. I also report on the methods used to ensure the integrity of results and a summary emphasizing the key points of this section.

Sources of Evidence

The need to have a more accurate, detailed, and exhaustive search led me to use the following databases in this order until I found relevant articles: Cochrane Database of Systematic Reviews, MEDLINE PLUS CINAHL database, and then Google Scholar database, all retrieved from Walden University Library. I verified relevant articles I found in Google Scholar using Ulrich Periodical Directory to ensure they were peer-reviewed.

Search Strategy

I searched for appropriate articles published from 2012 to 2017 using the following keywords: *BDI*, *HDRS*, *PHQ*, *major depression literature*, *Zung self-rating scale*, *Center for Epidemiologic Studies*, *outpatient*, *clinic*, *specificity*, *sensitivity*, *psychometric*, *accuracy*, and *validity*. I excluded other depression assessment tools such as the geriatric depression scale and the Cornell scale for depression in dementia, and I also excluded populations younger than 18 years. I made exceptions for articles concerned with ages older than 50 years if the age range of participants included ages younger than 50 years.

I appraised the articles using the Melynk Critical Appraisal Guide, which lists seven steps that can be used to critically appraise quantitative studies. The list below shows the seven questions that were answered from the Critical Appraisal Guide for Quantitative Studies for each article reviewed:

1. Why was the study done? • Was there a clear explanation of the purpose of the study and, if so, what was it?
2. What is the sample size? • Were there enough people in the study to establish that the findings did not occur by chance?
3. Are the instruments of the major variables valid and reliable? • How were variables defined? Were the instruments designed to measure a concept valid (did they measure what the researchers said they measured)? Were they reliable (did they measure a concept the same way every time)?

4. How were the data analyzed? • What statistics were used to determine if the purpose of the study was achieved?
5. Were there any untoward events during the study? • Did people leave the study and, if so, was there something special about them?
6. How do the results fit with previous research in the area? • Did the researchers base their work on a thorough literature review?
7. What does this research mean for clinical practice? • Is the study purpose an important clinical issue? (Stillwell et al., 2010)

While reading the articles and answering the questions I have outlined, I sorted the articles into relevant and nonrelevant stacks. From the relevant stack of articles, I used Bandolier's hierarchy levels of evidence to choose two articles that met the criteria for level one, which was the highest level of evidence-based research available: systematic reviews and meta-analysis (Mantzoukas, 2008). This search strategy ensured the highest quality of evidence was selected first from the Cochrane Database for Systematic Review articles followed by articles from peer-reviewed journals in the other databases.

From the compiled resources, the writer hoped to find answers to the questions as identified in the project question which was: What are the characteristics and indications of the major depression screening tools published in peer-reviewed literature between 2012 and 2017?

The evidence collated from the articles concerning depression assessment tools was relevant to the research topic. The goal from the assembled information was to find

information concerning the accuracy, advantages, and disadvantages to have a better understanding of the settings where health care professionals could utilize these tools, and to provide answers to the research question in this project.

Published Outcomes and Research

This project was a systematic review of existing literature concerning the use of appropriate tools for management of depression in an outpatient setting. A systematic review was less biased when compared with narrative reviews. According to Uman (2011), "systematic reviews, as the name implies, typically involve a detailed and comprehensive plan and search strategy derived prior, with the goal of reducing bias by identifying, appraising, and synthesizing all relevant studies on a topic " (p. 57).

By using relevant keywords, developing a search strategy, and using a reference librarian to help in conducting searches, the search was expected to be exhaustive and comprehensive. To ensure authenticity of the project results, the writer reviewed existing studies that showed the use of depression tools in adult patients from ages 18 to 50 years, relevant keywords (e.g., the different depression inventory tools, specificity, sensitivity, validity, reliability), meticulously organized data extracted into an easy to use format, and stored the results of the project within a secured database. To ensure the objectivity of compiled data, the researcher reviewed any biases and reviewed the statistical data (e.g., sensitivity, specificity, reliability, and validity) as presented in the articles.

Evidence Generated for the Doctoral Project

Participants. Three members were involved in reviewing the evidence as presented by the writer. These members, or expert panelists, examined the compiled data

and provided their opinion anonymously. These participants were volunteers and worked as a psychiatrist, nurse practitioner, or psychologist at the contracted clinical site. Based on their professional expertise, these professionals were conversant with the process of diagnosing patients with symptoms of depression and use of the different depression assessment tools.

Procedures. Upon completion of the feedback stage, the researcher presented the findings as a PowerPoint presentation to management at the clinical site. The overall goal was to develop a clinical guideline for diagnosing depression, which could be utilized by new and existing providers. I compiled the data from this project into an Excel spreadsheet showing the results (type of literature—systematic review [high-quality evidence] or peer-reviewed articles [moderate-quality evidence], the number of participants, the age of participants, specificity, sensitivity, advantages, and disadvantages of the depression inventory tools). An additional tool showing laboratory tests to rule out illnesses that mimic depression was presented in Appendix G of this project and referenced within the manuscript). This laboratory test tool was generated using existing textbooks. The different depression instruments were attached as Appendices A-F of the project as well.

Protections. I presented the expert panelists with a disclosure form concerning their role, questionnaire procedures, voluntary nature of the project, risks and benefits of being in the project, and privacy and their rights concerning the project as well as the topic, practice-focused questions, and Walden University's advocate phone number in case of any questions regarding their rights. The researcher completed Walden Form A,

ensured approval by Walden University institutional review board before engaging with the expert panelists, and provided the expert panelists with an ethics approval number in case they needed to contact the university. The expert panelists signed the form stating an understanding of their role before reviewing the evidence presented in the project and were allowed to withdraw from the project at any point. The expert panelists and participating organization identity were masked and were not shared or printed on the project. The questionnaires completed by the expert panelists will be stored safely for 2 years after the publication of this project in accordance with Walden University guidelines.

Analysis and Synthesis

I entered the information retrieved concerning the different depression assessment tools into a Microsoft Excel spreadsheet. The findings of this report are in tabular form with several headings such as depression assessment tools, the number of articles reviewed, and accuracy of the tools regarding sensitivity, specificity, age groups, and demographics. Another heading was the type of health care setting for the assessment tool and advantages and disadvantages of the assessment tools as identified from existing literature. The findings were a detailed compilation of all articles.

To ensure the integrity of information collated, the writer utilized high-quality, peer-reviewed articles. I reviewed multiple articles and used the most appropriate two articles for this project.

This project was a systematic review of existing literature concerning the use of the indicated depression assessment tool. I presented a summary of all the articles in tabular form in the conclusion. I performed no data analysis on the statistical data used.

Summary

Depressive disorders in medical settings remain under-recognized even though associated with high cost and disability expense (Manea et al., 2016). The CDC (2017) estimated the economic burden of depression, which includes suicide, and workplace related and direct costs, to be $210.5 billion in 2010. CDC (2017) linked depression to increased mortality risks, medical illnesses such as heart disease, and employment-related problems such as lower workplace productivity and absenteeism.

This section is a compilation of the sources of evidence, analysis, and synthesis of the research findings. I reviewed the systems used for recording, tracking, organizing, and analyzing the evidence as well as methods to ensure the integrity of findings. The next section details the findings and recommendations.

Section 4: Findings and Recommendations

Introduction

Local background of the topic stems from the lack of use of assessment tools in some outpatient behavioral health clinical settings (Okonofua, 2017). In such outpatient settings, patients are simply asked to assess the level of their depression using a verbal self-rating technique similar to pain scale assessment. The gap-in-practice is the lack of use of objective depression assessment tools in diagnosis and management of depression in some health care settings. Therefore, I systematically reviewed existing literature and compiled information on accuracy, pros, and cons of different depression inventory tools and diagnostic tests for management of depression with the goal of improving quality of care provided to patients by clinicians. The practice-focused question was: What are the characteristics and indications of the major depression screening tools and diagnostic tests published in peer-reviewed literature between 2012 and 2017 for adult patients?

In this section of the project, I lay out the systematic search of the literature concerning psychometric properties of six depression inventory tools. Within this section, I include the findings, implications for nursing/social change, recommendations made for the clinical site with noted practice deficiencies, strength, and limitations of the project. The six depression tools were as follows: BDI, HDRS, PHQ-2 and PHQ-9, MDI, SDS, and CES-D.

I performed the search for appropriate articles using the following keywords: *BDI*, *HDRS*, *PHQ*, *major depression literature*, *Zung self-rating scale*, *Center for Epidemiologic Studies*, *outpatient*, *clinic*, *specificity*, *sensitivity*, *psychometric*, *accuracy*,

and *validity*. The articles selected were published between 2012 and 2017, with participants between ages 18 and 50 years and the article's goal of reviewing psychometric properties of any of the included depression assessment tools. The need to have a more accurate, detailed, and exhaustive search led to the use of the following databases in this order to find relevant articles: Cochrane Database of Systematic Reviews, MEDLINE PLUS CINAHL database, and then Google Scholar database, all retrieved from the Walden University Library. Relevant articles found in Google Scholar were verified in the Ulrich Periodical Directory to ensure they were peer-reviewed.

Inclusion Criteria

The studies used included literature that met all of the following criteria: (a) reviewed the psychometric properties of any of the six depression assessment tools; (b) published between 2012 and 2017; and (c) age of participants between 18 and 50 years. There were no restrictions to geographical location. The inclusion criteria consisted of whether the article was a systematic review of the literature (or peer-reviewed articles if there was no systematic review of literature article after exhaustive search in the multiple databases or search engines).

Exclusion Criteria

I excluded other depression assessment tools such as the geriatric depression scale and Cornell scale for depression in dementia. This project mainly addressed articles covering the aforementioned six depression inventory tools with population from ages 18 to 50 years. Exceptions included articles with the population older than 50 years when the article age range for participants included less than 50 years of age.

Articles Reviewed

Overall, 10 articles based on the inclusion and exclusion criteria were reviewed for this project, whereas some of the articles covered more than one depression assessment tool. Even though my goal was to use articles of the highest quality (systematic review of literature), of the 10 articles selected, only one article was a systematic review of the literature, whereas the other articles were peer-reviewed articles of high quality.

Analytical Strategies Used

The results of a systematic review of the literature concerning the six depression assessment tools are in tabular form for readers' easy reference. Addressing an aspect of the practice-focused question, Table 1 and Table 2 show the sensitivity and specificity of the depression assessment tools, which help readers to understand the reliability, accuracy, and validity of the tools. These first two tables consist of the sensitivity and specificity of the tools ranked from the least to the highest. Table I1 (Appendix I) is a compilation of the age groups or demographics, the advantages, disadvantages, cutoff scores, cost, and direct references to the depression assessment tools. Table I2 (Appendix I) provides additional information regarding study design, methods, level of evidence, setting, country, number of participants, and the outcomes of the studies.

Findings and Implication

Presently, only a few systematic literature reviews address psychometric properties of depression inventory tools. Of the 10 articles selected for this project, only 1 was a systematic review of the literature. The results of this study revealed characteristics

and indications of the major depression screening tools and diagnostic tests published mostly in peer-reviewed literature between 2012 and 2017 found in the Cochrane Database of Systematic Reviews, MEDLINE PLUS CINAHL database, and Google Scholar database.

Accuracy of the Depression Assessment Tools

I reviewed the accuracy of the depression assessment tools using their sensitivity and specificity ratings from the different articles selected.

Sensitivity of the Six Depression Assessment Tools

Sensitivity refers to the ability of a tool to accurately test positive for an illness. Sensitivity tests are used to rule out other diseases (i.e., the ability of the test to correctly identify those with the disease [true positive rate]). Arranged from the least to the most sensitive, PHQ2 was identified as 83% sensitive, BDI-II at 85%, PHQ-9 at 88%, Zung's SDS from 85.7% to 91.4%, and MDI from 88.4% to 90%. HDRS has the highest ranking with 93%. From this analysis, evidence shows all six depression assessment tools have a high sensitivity rating from 83% to 93%.

Table 1

Sensitivity of the Six Depression Assessment Tools

	PHQ-2	Beck's depression inventory-II	PHQ-9	Zung SDS	MDI	Hamilton Depression Rating Scale	CES-D
Sensitivity	83	85, 85	88	91.4	88.4, 90	93*	*

Note. *Not available. Values for the depression tools were extrapolated from each article reviewed (see Appendix H).

Specificity of the Six Depression Assessment Tools

For specificity, the question was: How often will the test be negative when the patient does not have the disease? Specificity in short rules out those without the disease

(true negative rate). Arranged in order from the least specific to the most specific, Table 2 shows MDI as on average (82%, 78.9%) the least specific rating while HDRS has the highest specific rating at 97%.

Table 2

Specificity of the Six Depression Assessment Tools

	MDI	Beck's depression inventory-II	PHQ-9	PHQ-2	Zung SDS	Hamilton Depression Rating Scale	CES-D
Sensitivity	82, 78.9	78, 88	88	90	91.8	*97	*

Note. *Not available. Values for the depression tools were extrapolated from each article reviewed (see Appendix H).

From the analysis of Tables 1 and 2, evidence shows HDRS to have the highest sensitivity (93%) and specificity (97%) rates. Overall, all six depression rating scales exhibit high ratings on both the specificity (range from 83% to 93%) and sensitivity rate ranging from 78% to 97% (except CES-D that did not include the sensitivity or specificity on either of the articles reviewed). From these data, the writer believed the choice of any of the depression tools was valid and based on clinician preference and factors such as the pros and cons, as well as availability or cost of the tools.

To further answer the practice-focused questions for this project, Table I1 in Appendix I contains essential data such as the pros and cons as well as whether the tools are in public domain or are proprietary. The table answered part of the practice question, which was: What are the characteristics and indications of the major depression screening tools and diagnostic tests published in peer-reviewed literature between 2012 and 2017 for adult patients? Therefore, the table includes findings such as the age groups for the use of the depression assessment tools, type of health care settings, advantages, and disadvantages of the assessment tools from literature. The table also contains information

such as cutoff scores for the depression inventory tools, their costs, and reference to the publisher or reliable sites where these depression assessment tools can be found for purchase or free use.

Other Findings from Analysis and Synthesis of the Collected Evidence

The articles utilized for this project were evaluated using the Critical Appraisal Guide for Quantitative Studies, and all of the articles met all seven steps of this guide. There was a clear explanation of the purpose of the studies; the sample size identified ranged from 70 participants in 1 study to 8500 participants in another study. There was an explanation of the major variables and the instruments measured the relevant concepts showing reliability and validity of the depression inventory tools. The statistical analyses were complex, using software such as SPSS and SAS. The articles identified any untoward events such as participants removed because of not meeting criteria for inclusion or questionnaires not included because of missing values.

Thorough literature research was conducted in all the studies while the results fit previous research in their area. Some of the findings validated earlier findings of the assessment tools such as there was a strong correlation between some of the depression inventory tools. From one of the articles, there was a strong correlation between BDI and Zung SDS, but then BDI was found to have marginal superior internal consistency reliability (Alpha = 0.88) compared to Zung SDS (Alpha = 0.85; Campbell et al., 2012)

Unanticipated Limitations and Their Potential Effects on the Findings

Just as expected in any project, there were unanticipated limitations to this project. These include the following:

- After exhaustive searches, there were very few publications found that were of relevance to the project or that reviewed psychometric properties of depression inventory tools.
- Even though many of the articles found were peer-reviewed, only a few were a systematic review of literature, which is the highest level of evidence this researcher sought.
- Another unanticipated limitation was that some of the selected articles were a comparison of one depression tool with another in other countries such as Croatia, Italy, Germany, Denmark, and Barbados.
- For the information needed to compile the above tables such as the age of patients who could use the assessment tools, there were no results in the articles; the results were on publisher websites.
- There were variations in specificity and sensitivity of the PHQ-9 at different cutoff scores. For instance, the research findings from Manea et al. (2012) indicated a variation in sensitivity and specificity at different cutoff scores. Hence, the question arises: At what cutoff point are some of the tools indicative of depressed mood? For PHQ-9, for instance, Manea et al. (2012) concluded it had acceptable diagnostic properties at a range of cutoff scores (8-11). Therefore, when using PHQ-9, clinicians should exercise caution when choosing a cutoff score by "taking into account the characteristics of the population, the settings and the efficacy of screening on outcomes" (Manea et al., 2012, p. 195). These authors went further to explain that a "cutoff score of 10 could result in false

negatives in hospital settings while more false-positive results may be seen in primary care" (Manea et al., 2012, p. 195).

- Limitation of analysis in articles: Bech et al. (2015) used a time frame that covered the past week and not the conventional 2-week time frame for diagnosis of depression.
- Completed data were not available in some of the included information while in some of the articles, there were missing values from the assessment completed by the patient.
- The potential effect of these limitations is difficult to determine based on the strengths of the articles. The articles were thorough in their literature search, data compilation, and data analyses and provided enough information to support the summaries and recommendations made to the clinical site with noted practice deficiency.

Implications of Findings Regarding Individuals, Communities, Institutions, and Systems

The implications resulting from the findings show specific assessment tools as having high reliability, sensitivity, and specificity when compared to others. At the individual level, the findings indicated the depression assessment tools could reliably be used to determine individuals' level of depression based on clinicians' preferences. At the community, institution, and systems levels, it is essential for the clinician to know factors that can affect the results of both self-reported and observer-related depression inventory tools. While some of the assessment tools are self-reported, others are observer-related

and “can be affected by demographics, clinical characteristics of patients such as age, level of education, type or depression and personality type” (Schneibel et al., 2012, p. 63). For the observer-related assessment tools, it is essential to note an observer could be more likely to see improvements in depressive symptomatology than a patient, and the patient may still be affected by cognitive bias (Schneibel et al. 2012).

For PHQ-2, the meta-analysis research conducted by Manea et al. (2016) suggested the possibility of inaccuracy in the cutoff score detecting patients with depression and the authors suggested lowering the cutoff from three. At a cutoff score of two, there is increased sensitivity. This variation in researched sensitivity of the PHQ-2 compared with the original validation study with a cutoff of three shows the need for further research into this issue in the future.

Potential Implications for Positive Social Change

The implication for positive social change includes the fact that the findings of this project could be used by clinicians to determine their choice of depression assessment tools. Researchers, likewise, could use this project for a quick overview of existing publications about the various depression assessment tools addressed within this project. In essence, this project could lead to increased use of depression assessment tools and improved diagnosis and management of depressed mood among patients in any health care settings from outpatient to inpatient. This project also shows the need for objective depression tool usage is a worldwide phenomenon and so the usability of the findings of this project extends beyond the boundaries of the United States and has the potential to affect patient care worldwide.

Recommendations

The gap-in-practice is the lack of use of objective depression assessment tools in diagnosis and management of depression in some health care settings, my clinical site inclusive. Multiple barriers contributing to this problem can be seen at the macro and micro level. At the macro level globally, World Health Organization (2008) identified lack of resources (some of the depression assessment tools are not in public domain and cost money), lack of trained health care providers, and the social stigma associated with mental disorders as barriers to patients being treated adequately using the depression tools. Within the clinical settings, obstacles include time constraints and inherent resistance to change by integrating depression tools into the business process by clinical staff and providers. At the micro level, patients' resistance to seeking treatment, resistance in completing assessment tools, inaccurate report of the severity of depression by patients, erroneous information provided on the assessment tools, and fear of "depression" stigma are some of the barriers to the use of assessment tools (Okonofua, 2017). These obstacles at the micro and micro levels could be challenges to the full implementation of these evidence-based guidelines in the use of depression assessment tools in outpatient settings (Okonofua, 2017). The following were the recommendations made for my clinical site where the use of depression assessment tools was very minimal.

Depression inventory tools are an essential part of objectively assessing a patient's mood in a health care setting. Therefore, the recommendations I made to the clinical practice setting (where I did my clinical rotation and noticed lack of use of depression inventory tools) included the following:

- Choose a depression inventory tool based on its accuracy and reliability.
- Train clinicians and office personnel on the importance of using such objective tools to screen for or to assess a patient's level of depression.
- Develop and implement a quality improvement plan by inculcating the preferred depression assessment tool in electronic form into the electronic health records systems or integrated within a facility website whereby patients could complete the tests from any location and results would be readily available to the health care provider before a patient's visit.
- If patients are unable to complete the online assessment tools, integrate the depression assessment tools into business processes and allow patients to complete the tests while waiting to be seen by the providers either in paper form or on iPad, so that results are easily available to the provider.
- Implement the compulsory use of evidence-based practice resources such as the depression assessment tools into the organization policies and procedures for patient care.
- Encourage patients' involvement and improve patients' understanding of the assessment tools used in psychiatry.
- If clinical practice is not utilizing any depression assessment tools currently, formulate and implement a plan for the change of process within the facility.

Proposed Secondary Products to Guide the use of the Primary Products in Practice

The six depression inventory tools' cutoff scores for detecting depressed symptoms and their scoring instructions are in Appendices A-F of this project. While

these copies are for educational purposes, users of the product of this project should ensure they do not violate copyright of any depression assessment tools by simply copying the tools from this project. A compilation of the publications reviewed for this project is provided in Appendix H.

Contribution of the Doctoral Project Team

While this project utilized no project team members, the products of this project would not have been possible without the valuable inputs of the chair and members of the committee. At this time, there are no plans to extend this project beyond the DNP doctoral project.

Strengths and Limitations of the Project

This project was a comprehensive systematic review of existing literature about psychometric properties of the six depression assessment tools. Ten high-quality peer-reviewed and systematic review of literature articles were reviewed and analyzed. A significant strength of this project was that very marginal variability was noted concerning the reliability and validity of the depression assessment tools from the various articles reviewed. The samples used for the articles, their data gathering, and analysis indicated minimal bias or limitations. The articles used were dispersed geographically, showing the question about psychometric properties of depression assessment tools covers multiple countries and continents.

The limitations of the project included, first, there was very little work on the topic. Second, there were some identified biases in some of the articles. For example, in one of the articles used, study selection was completed by one author. Third, some

patients identified as depressed in some of the articles did not undergo further confirmatory testing, which meant these works had partial verification bias.

Recommendations for Future Projects Addressing Similar Topics with Similar Methods

The findings from this project show very little work completed concerning the psychometric properties of depression assessment tools. Hence, this author encourages future work should be conducted by clinicians, most especially health care professionals in advanced nursing roles. Authors of future studies regarding psychometric properties of the depression tools should consider reporting on the different tools as well as any variations in accuracy based on the different cutoff scores.

Section 5: Dissemination Plan

DNP-prepared nurses encounter stages in their projects when they must disseminate the findings and implications. It is essential to spread evidence-based practice findings to key stakeholders, organizations, and other health care professionals to use the recommendations and innovations for practice in different settings (Forsyth, Wright, Scherb, & Gaspar, 2010, p. 2). The presentation of findings to key stakeholders is essential in any evidence-based project or research. According to Forsyth et al. (2010), to "facilitate timely and quality dissemination of evidence-based practice projects, there is a need for clear criteria identifying the essential information to be shared, how to share it effectively, and how to evaluate the end product" (p. 2).

On completion of this systematic review of the literature concerning the use of depression assessment tools, I provided the findings to the clinical setting as an educational session with the hope of health care professionals using the recommendations in the management of a depression diagnosis. I conducted an oral presentation by using technology such as projectors and PowerPoint slides. The oral presentation afforded me the opportunity to interact with the key stakeholders, to clarify any issues, and to receive feedback. The tactfulness and professionalism of the presenter played a significant role in presentation of the project findings. In short, I presented this DNP project, the findings, and recommendations as an oral presentation to my clinical site.

Based on the nature of this project, other audiences such as advanced nurse practitioners, psychiatrists, and psychologists in private practices, community-based mental health facilities, and primary care settings will benefit from the findings of this

project. A poster presentation is another form by which I could present the findings in these settings or at professional workshops. However, there are key elements in poster development and presentation of which the presenter must be aware. These elements include the following:

- early planning with a clear focus;
- following conference guidelines, such as poster size and type (hanging or freestanding);
- using bullet points or abbreviated wording;
- incorporating pictures or graphics;
- balancing content with white space; and
- using a large font size for viewing at a distance. (Forsyth et al., 2010, p. 2)

Analysis of Self

The DNP education journey has prepared me to be a better nurse leader with the ability to discern issues that can influence patient outcomes. The preparation for my DNP project, the choice of topic, and the process of engaging with my committee chair and committee members have made me a better nurse leader, scholar, writer, and most especially, an advocate for better patient care.

As a nurse practitioner in behavioral health, my passion is in caring for patients with mental health disorders. Hence, for future goals, I intend to seek out opportunities to disseminate the findings of this project through various professional organizations, continue to provide excellent care for my patients, and continue to write and publish my

articles in health-care-related journals. My other goals include seeking out opportunities for psychiatric nursing instruction at graduate and undergraduate levels.

My DNP education at Walden University has prepared me for future professional roles as a nurse leader. As a nurse practitioner, it has sharpened my knowledge of the need to provide evidence-based care to improve health outcomes and to be an advocate for improved patient care within organizations. My DNP education has instilled in me the values of a health care leader, the "eyes of a change agent," and the passion for leaving a mark as a health care leader that can bring about social change at micro and macro levels. My goal is to continue in my professional path by seeking opportunities in academia, within my organization, in my community, and at international levels to be a change agent and to have a positive influence on health care and the lives of people.

Summary

Depression is a mental illness requiring prompt identification and treatment because it has grave consequences if untreated. Depression can affect a person's level of functioning, can lead to worsening health conditions and comorbid substance abuse, and can lead to suicide at its worst. Hence, the need to accurately diagnose and treat patients with depression using assessment tools and diagnostic tests is critical. My chosen topic, use of assessment tools in the management of patients with depression, is an area that I am passionate about and I firmly believe every clinician should use objective assessment tools when caring for patients. By working on my DNP project, I discovered a few published articles of the highest quality of evidence (i.e., a systematic review of literature) concerning psychometric properties of depression assessment tools. In this

systematic literature review, I investigated six depression assessment tools regarding their accuracy (identified by specificity, sensitivity, reliability, and validity), their pros and cons, and the possible demographics and health care settings to use these inventory tools. The six depression inventory tools examined for his project included BDI, HDRS, PHQ-2 and PHQ-9, MDI, Zung SDS, and CES-D. The practice-focused question was: What are the characteristics and indications of the major depression screening tools/diagnostic tests published in peer-reviewed literature between 2012 and 2017 for adult patients? Of the six depression assessment tools reviewed, HDRS was found to have the highest specificity and sensitivity.

A major insight gained from this project and the journey of the DNP degree is that it takes a team to bring about a change. Although there were personal challenges along the way of this project, using the strength of my team, my chair, my colleagues, my preceptors, and the staff at my clinical sites aided me in reaching the end of this project.

References

American Psychiatric Association. (2010). What is depression? Retrieved from https://www.psychiatry.org/patients-families/depression/what-is-depression

Bech, P., Timmerby, N., Martiny, K., Lunde, M., & Soendergaard, S. (2015). Psychometric evaluation of the major depression inventory (MDI) as depression severity scale using the LEAD (longitudinal expert assessment of all data) as index of validity. *BMC Psychiatry*, *15*(1), 190. doi:10.1186/s12888-015-0529-3

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for Beck depression inventory II (BDI-II). *San Antonio, TX, Psychology Corporation*.

Bieling, P. J., McCabe, R. E., & Antony, M. M. (2004). Measurement issues in preventing anxiety and depression: Concepts and instruments. *The prevention of anxiety and depression: Theory, research, and practice*, 43-71. Washington, DC, US: American Psychological Association.

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., De Girolamo, G., . . . & Karam, A. N. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*, *9*(1), 90. doi:10.1186/1741-7015-9-90

Campbell, M., Maynard, D., Roberti, J. and Emmanuel, M. (2012). A comparison of the psychometric strengths of the public-domain Zung self-rating depression scale with the proprietary beck depression inventory-II in Barbados. *West Indian Medical Journal*, 61(5). doi:10.7727/wimj.2010.145

Centers for Disease Control. (2017). Depression. Retrieved from https://www.cdc.gov/mentalhealth/basics/mental-illness/depression.htm

Chin, W. Y., Choi, E. P., Chan, K. T., & Wong, C. K. (2015). The psychometric properties of the Center for Epidemiologic Studies Depression Scale in Chinese primary care patients: Factor structure, construct validity, reliability, sensitivity, and responsiveness. *PloS One*, *10*(8), e0135131. doi:10.1371/journal.pone.0135131

Cuijpers, P., Dekker, J., Noteboom, A., Smits, N., & Peen, J. (2012). Sensitivity and specificity of the Major Depression Inventory in outpatients. *BMC Psychiatry*, *7*(1), 39. doi:10.1186/1471-244X-7-39

Davis, L., Uezato, A., Newell, J. M., & Frazier, E. (2008). Major depression and comorbid substance use disorders. *Current Opinion in Psychiatry*, *21*(1), 14-18. doi:10.1097/YCO.0b013e3282f32408

Doenges, M. E., & Moorhouse, M. F. (2012). *Application of nursing process and nursing diagnosis: An interactive text for diagnostic reasoning*. Philadelphia, PA: FA Davis. doi:10.1037/10722-000

Faust, C. (2002). Orlando's deliberative nursing process theory: A practice application in an extended care facility [Clinical outlook]. *Journal of Gerontology Nursing, 28*(7), 14-18. doi:10.3928/0098-9134-20020701-05

Fawzi, M., Fawzi, M., & Abu-Hindi, W. (2012). Arabic version of the major depression inventory as a diagnostic tool: reliability and concurrent and discriminant validity. *Eastern Mediterranean Health Journal*, *18*(4), 304-310. Retrieved from http://www.who.int/iris/handle/10665/118314

Forsyth, D. M., Wright, T. L., Scherb, C. A., & Gaspar, P. M. (2010). Disseminating evidence-based practice projects: Poster design and evaluation. *Clinical Scholars Review*, *3*(1), 14-21. doi:10.1891/1939-2095.3.1.14

Institute of Clinical Systems Improvement. (2017). Depression. Retrieved from https://www.icsi.org/guideline_sub-pages/depression/depression_suspected/

Jakšić, N., Ivezić, E., Jokić-Begić, N., Surányi, Z., & Stojanović-Špehar, S. (2013). Factorial and diagnostic validity of the Beck Depression Inventory-II (BDI-II) in Croatian primary health care. *Journal of Clinical Psychology in Medical Settings*, *20*(3), 311-322. doi:10.1007/s10880-013-9363-2

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The patient health questionnaire-2: Validity of a two-item depression screener. *Medical Care*, *41*(11), 1284-1292. doi:10.1097/01.MLR.0000093487.78664.3C

Levine, S. Z. (2013). Evaluating the seven-item Center for Epidemiologic Studies depression scale short form: A longitudinal U.S. community study. *Social Psychiatry and Psychiatric Epidemiology*, *48*(9), 1519-1526. doi:10.1007/s00127-012-0650-2

Lichtblau, L. (2011). *Psychopharmacology demystified*. Clifton Park, NY: Delmar Cengage Learning.

Louch, P. (2008). Depression management in primary care. *Primary Health Care*, 15(10), 20-22. doi:10.7748/phc2005.12.15.10.20.c559

Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cutoff score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis.

Canadian Medical Association Journal, *184*(3), E191-E196. doi:10.1503/cmaj.110829

Manea, L., Gilbody, S., Hewitt, C., North, A., Plummer, F., Richardson, R., & . . . McMillan, D. (2016). Identifying depression with the PHQ-2: A diagnostic meta-analysis. *Journal of Affective Disorders*, *20*(3), 382-395. doi:10.1016/j.jad.2016.06.003

Mantzoukas, S. (2008). A review of evidence-based practice, nursing research, and reflection: Leveling the hierarchy. *Journal of Clinical Nursing*, *17*(2), 214-223. doi:10.1111/j.1365-2702.2006.01912.x

Martiny, K., Lunde, M., Unden, M., Dam, H., & Bech, P. (2005). Adjunctive bright light in non-seasonal major depression: Results from clinician-rated depression scales. *Acta Psychiatry Scandinavia, 112*, 117-125. doi:10.1111/j.1600-0447.2005.00574.x

Maurer, D. (2012). Screening for depression. *American Family Physician, 85*(2), 139-144. Retrieved from http://www.usafp.org/wp-content/uploads/2013/12/Screening-for-Depression.pdf

National Institute of Mental Health [NIMH]. (2008). Brain basics. Retrieved from https://www.nimh.nih.gov/health/educational-resources/brain-basics/brain-basics.shtml

NIMH. (2016a). 12-month prevalence of depression with severe impairment. Retrieved from https://www.nimh.nih.gov/health/statistics/prevalence/major-depression-with-severe-impairment-among-adults.shtml

NIMH. (2016b). 12-month prevalence of major depressive episodes among U.S. adults. Retrieved from https://www.nimh.nih.gov/health/statistics/prevalence/major-depression-among-adults.shtml

Okonofua, M. (2017). *Use of assessment tools in management of depression*.

Orlando, I. (1987). Nursing in the 21st century-alternate paths. *Journal of Advanced Nursing, 12*(4), 405-412. doi:10.1111/j.1365-2648.1987.tb01349.x

Perese, E. F. (2012). *Psychiatric advanced practice nursing: A biopsychosocial foundation for practice*. Philadelphia, PA: F. A. Davis.

Pratt, L. A., & Brody, D. J. (2014). Depression in the U.S. household population, 2009-2012. Retrieved from https://www.cdc.gov/nchs/data/databriefs/db172.pdf

Psychcentral. (2017). Depression symptoms. Retrieved from https://psychcentral.com/disorders/depression/depression-symptoms-major-depressive-disorder/

Raimo, S., Trojano, L., Spitaleri, D., Petretta, V., Grossi, D., & Santangelo, G. (2015). Psychometric properties of the Hamilton depression rating scale in multiple sclerosis. *Quality of Life Research*, *24*(8), 1973-1980. doi:10.1007/s11136-015-0940-8

Romera, I., Delgado-Cohen, H., Perez, T., Caballero, L., & Gilaberte, I. (2008). Factor analysis of the Zung self-rating depression scale in a large sample of patients with major depressive disorder in primary care. *BMC Psychiatry*, *8*(1), 4. doi:10.1186/1471-244X-8-4

Ruiz-Grosso, P., de Mola, C. L., Vega-Dienstmaier, J. M., Arevalo, J. M., Chavez, K., Vilela, A., & Huapaya, J. (2012). Validation of the Spanish center for epidemiological studies depression and Zung self-rating depression scales: A comparative validation study. *PloS one*, *7*(10), e45413. doi:10.1371/journal.pone.0045413

Schneibel, R., Brakemeier, E. L., Wilbertz, G., Dykierek, P., Zobel, I., & Schramm, E. (2012). Sensitivity to detect change and the correlation of clinical factors with the Hamilton depression rating scale and the Beck depression inventory in depressed inpatients. *Psychiatry research*, *198*(1), 62-67. doi:10.1016/j.psychres.2011.11.014

Straaso, B., Lauritzen, L., Lunde, M., Vinberg, M., Lindberg, L., Larsen, E. R., . . . Bech, P. (2014). Dose-remission of pulsating electromagnetic fields as augmentation in therapy-resistant depression: a randomized, double-blind controlled study. *Acta Neuropsychiatry, 26*, 272-279. doi:10.1017/neu.2014.5

Stillwell, S., Fineout-Overholt, E., Melnyk, B., & Williamson, K. (2010). Evidence-based practice, step by step: Asking the clinical question. *American Journal of Nursing*, *110*(3), 58-61. doi:10.1097/01.NAJ.0000368959.11129.79

U.S. Preventive Services Task Force (USPSTF). (2016). Screening for depression in adults. U.S. Preventive Services Task Force Recommendation Statement. *Journal of the American Medical Association, 315*(4), 380-387. doi:10.1001/jama.2015.18392

Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 20*(1), 57-59. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3024725/

VanDam, N. T., & Earleywine, M. (2011). Validation of the Center for Epidemiologic Studies depression scale-revised (CESD-R): Pragmatic depression assessment in the general population. *Psychiatry Research*, *186*(1), 128-132. doi:10.1016/j.psychres.2010.08.018

World Health Organization. (2008). *The global burden of disease: 2004 update*. Retrieved from http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf

Zauszniewski, A. J., Bekhet, A., & Haberlein, S. (2012). A decade of published evidence for psychiatric and mental health nursing interventions. *OJIN: The Online Journal of Issues in Nursing*, *17*(3). doi:10.3912/OJIN.Vol17No03HirshPsy01

Zimmerman, M., Martinez, J. H., Young, D., Chelminski, I., & Dalrymple, K. (2013). Severity classification on the Hamilton depression rating scale. *Journal of affective disorders*, *150*(2), 384-388. doi:10.1007/s11136-015-0940-8

Zung, W. W. K. (1965). A self-rating scale for depression. *Archives of General Psychiatry, 12*, 63-70. doi:10.1001/archpsyc.1965.01720310065008

Appendix A: Beck Depression Inventory

Beck Depression Inventory is a copyright protected assessment tool that cannot be reproduced in this work. It is available for purchase at www.Psychcorp.com.

Appendix B: Hamilton Depression Rating Scale

Name_____Date_____

1. Depressed Mood (sadness, hopelessness, helplessness, worthlessness)
- ☐ 0 = Absent
- ☐ 1 = These feeling states indicated only on questioning
- ☐ 2 = These feeling states spontaneously reported verbally
- ☐ 3 = Communicates feeling states nonverbally (i.e., facial expression, posture, voice, tendency to weep)
- ☐ 4 = Reports virtually only these feeling states in spontaneous verbal and nonverbal communication

2. Feelings of Guilt
- ☐ 0 = Absent
- ☐ 1 = Self-reproach, feels he/she has let people down
- ☐ 2 = Ideas of guilt or rumination over past errors or "sinful" deeds
- ☐ 3 = Present illness is a punishment; delusions of guilt
- ☐ 4 = Hears accusatory or denunciatory voices and/or experiences threatening visual hallucinations

3. Suicide
- ☐ 0 = Absent
- ☐ 1 = Feels life is not worth living
- ☐ 2 = Wishes he/she were dead or has any thoughts of possible death to self
- ☐ 3 = Suicidal ideas or gestures
- ☐ 4 = Attempts at suicide (any serious attempt rates "4")

4. Insomnia—Early
- ☐ 0 = No difficulty falling asleep
- ☐ 1 = Complains of occasional difficulty falling asleep (i.e., >1/2 hour)
- ☐ 2 = Complains of nightly difficulty falling asleep

5. Insomnia—Middle
- ☐ 0 = No difficulty
- ☐ 1 = Complains of being restless and disturbed during the night
- ☐ 2 = Wakes during the night—getting out of bed rates "2" (except for purposes of voiding)

6. Insomnia—Late
- ☐ 0 = No difficulty
- ☐ 1 = Wakes in early hours of the morning but falls back to sleep
- ☐ 2 = Unable to fall asleep again if he/she gets out of bed

7. Work and Activities
- ☐ 0 = No difficulty
- ☐ 1 = Thoughts and feelings of incapacity; fatigue or weakness related to activities, work or hobbies
- ☐ 2 = Loss of interest in activity, hobbies or work—either directly reported by patient or indirectly in listlessness, indecision and vacillation (feels he/she has to push self to work or for activities)
- ☐ 3 = Decrease in actual time spent in activities or decrease in productivity
- ☐ 4 = Stopped working because of present illness

8. Retardation (slowness of thought and speech; impaired ability to concentrate; decreased motor activity)
- ☐ 0 = Normal speech and thought
- ☐ 1 = Slight retardation at interview
- ☐ 2 = Obvious retardation at interview
- ☐ 3 = Interview difficult
- ☐ 4 = Complete stupor

9. Agitation
- ☐ 0 = None
- ☐ 1 = Fidgetiness
- ☐ 2 = "Playing with" hands, hair, etc.
- ☐ 3 = Moving about, can't sit still
- ☐ 4 = Hand wringing, nail biting, hair pulling, lip biting

10. Anxiety—Psychic
- ☐ 0 = No difficulty
- ☐ 1 = Subjective tension and irritability
- ☐ 2 = Worries about minor matters
- ☐ 3 = Apprehensive attitude apparent in face or speech
- ☐ 4 = Fears expressed without questioning

11. Anxiety—Somatic (physiological concomitants of anxiety such as gastrointestinal: dry mouth, flatulence, indigestion, diarrhea, cramps, belching; cardiovascular: palpitations, headaches; respiratory: hyperventilation, sighing; urinary frequency; sweating)
- ☐ 0 = Absent
- ☐ 1 = Mild
- ☐ 2 = Moderate
- ☐ 3 = Severe
- ☐ 4 = Incapacitating

12. Somatic Symptoms—Gastrointestinal
- ☐ 0 = None
- ☐ 1 = Loss of appetite, but eating; heavy feelings in abdomen
- ☐ 2 = Difficulty eating without urging; requests or requires laxatives or medication for bowels or medication for GI symptoms

13. Somatic Symptoms—General
- ☐ 0 = None
- ☐ 1 = Heaviness in limbs, back of head; backache, headache, muscle ache; loss of energy and fatigability
- ☐ 2 = Any clear-cut symptoms rate "2"

14. Genital Symptoms (i.e., loss of libido, menstrual disturbances)
- ☐ 0 = Absent
- ☐ 1 = Mild
- ☐ 2 = Severe

15. Hypochondriasis
- ☐ 0 = Not present
- ☐ 1 = Self-absorption (bodily)
- ☐ 2 = Preoccupation with health
- ☐ 3 = Frequent complaints, requests for help, etc.
- ☐ 4 = Hypochondriacal delusions

16. Weight Loss
- ☐ 0 = No weight loss
- ☐ 1 = Slight or doubtful weight loss
- ☐ 2 = Obvious or severe weight loss

17. Insight
- ☐ 0 = Acknowledges being depressed and ill
- ☐ 1 = Acknowledges illness but attributes cause to bad food, climate, overwork, virus, need for rest, etc.
- ☐ 2 = Denies being ill at all

18. Diurnal Variation
- ☐ 0 = No variation
- ☐ 1 = Mild: doubtful or slight variation
- ☐ 2 = Severe: clear or marked variation; if applicable, note whether symptoms are worse in AM ☐ or PM ☐

19. Depersonalization and Derealization (feelings of unreality, nihilistic ideas)
- ☐ 0 = Absent
- ☐ 1 = Mild
- ☐ 2 = Moderate
- ☐ 3 = Severe
- ☐ 4 = Incapacitating

20. Paranoid Symptoms
- ☐ 0 = None
- ☐ 1 = Suspicious
- ☐ 2 = Ideas of reference
- ☐ 3 = Delusions of reference and persecution
- ☐ 4 = Paranoid hallucinations

21. Obsessive/Compulsive Symptoms
- ☐ 0 = Absent
- ☐ 1 = Mild
- ☐ 2 = Severe

Total HAM-D Score: _____

The Hamilton Depression Rating Scale is the most widely used interview scale, developed in 1960 to measure the severity of depression in an inpatient population. Since then, many versions have been adapted, including structured interview guides, self-report forms, and computerized versions.

In the original clinician-administered scale, the first 17 items are tallied for the total score, while items 18-21 are used to further qualify the depression. The scale takes 20-30 minutes to administer. Scores of 0-7 are considered normal, and scores greater than or equal to 20 indicate moderately severe depression. Each item either is scored on a 5-point scale, representing absent, mild, moderate, or severe symptoms, or on a 3-point scale, representing absent, slight or doubtful, and clearly present symptoms. The HDRS contains a relatively large number of somatic symptoms and relatively few cognitive or affective symptoms. The 21 items it assesses are as follows:

- Depressed mood
- Feelings of guilt
- Thoughts of suicide
- Insomnia
- Work and activities
- Psychomotor retardation
- Psychomotor agitation
- Psychic anxiety
- Somatic anxiety
- Gastrointestinal symptoms
- General somatic symptoms
- Genital symptoms
- Hypochondriasis
- Loss of insight
- Loss of weight
- Diurnal variation
- Depersonalization and derealization
- Paranoid symptoms
- Obsessional and compulsive symptoms

Source. Bienenfeld, D., & Stinson, K. N. (2014). Screening tests for depression.

Appendix C: PHQ-2/PHQ9

PHQ-2

Over the past two weeks, how often have you been bothered by any of the following problems?
Little interest or pleasure in doing things.
0 = Not at all
1 = Several days
2 = More than half the days
3 = Nearly every day
Feeling down, depressed, or hopeless.
0 = Not at all
1 = Several days
2 = More than half the days
3 = Nearly every day
Total point score: _____

Score interpretation:

PHQ-2 score	*Probability of major depressive disorder(%)*	*Probability of any depressive disorder(%)*
1	15.4	36.9
2	21.1	48.3
3	38.4	75.0
4	45.5	81.2
5	56.4	84.6
6	78.6	92.9

Source: Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire-2: validity of a two-item depression screener. *MedCare, 41*, 1284-92.

Appendix D: Major Depression Inventory

Major Depression Inventory

	All of the time 5	Most of the time 4	Slightly more than half of the time 3	Slightly less than half of the time 2	Some of the time 1	At no time 0
1 Have you felt low in spirits or sad?						
2 Have you lost interest in your daily activities?						
3 Have you felt lacking in energy and strength?						
4 Have you felt less self-confident?						
5 Have you had a bad conscience of feelings of guilt?						
6 Have you felt that life wasn't worth living?						
7 Have you had difficulty in concentrating, e.g., when reading the newspaper or watching television?						
8a Have you felt very restless?						
8b Have you felt subdued?						
9 Have you had trouble sleeping at night?						
10a Have you suffered from reduced appetite?						
10 b Have you suffered from increased appetite?						

Source: Bienenfeld, D., & Stinson, K. N. (2014). Screening tests for depression.

Appendix E: Zung Self-Rating Depression Scale

Instructions: For each item below, please place a check mark (√) in the column which best describes how often you felt or behaved this way during the past several days.

Place check mark (√) in correct column	**A little of the time**	**Some of the time**	**Good part of the time**	**Most of the time**
1. I feel down-hearted and blue.				
2. Morning is when I feel the best.				
3. I have crying spells or feel like it.				
4. I have trouble sleeping at night.				
5. I eat as much as I used to.				
6. I still enjoy sex.				
7. I notice that I am losing weight.				
8. I have trouble with constipation.				
9. My heart beats faster than usual.				
10. I get tired for no reason.				
11. My mind is as clear as it used to be.				
12. I find it easy to do the things I used to.				
13. I am restless and can't keep still.				
14. I feel hopeful about the future.				
15. I am more irritable than usual.				
16. I find it easy to make decisions.				
17. I feel that I am useful and needed.				
18. My life is pretty full.				
19. I feel that others would be better off if I were dead.				
20. I still enjoy the things I used to do.				

Source: Zung, W. W. K. (1965). A self-rating scale for depression. Archives of General Psychiatry 12, 63-70.

Scoring the Zung Self-Rating Depression Scale

In scoring the SDS, a value of 1, 2, 3 and 4 is assigned to a response depending upon whether the item is worded positively or negatively.

For items 1, 3, 4, 7, 8, 9, 10, 13, 15, 19 the scoring is:
- A little of the time = 1
- Some of the time = 2
- Good part of the time = 3
- Most of the time = 4

Items 2, 5, 6, 11, 12, 14, 16, 17, 18, 20 are reverse scored as follows:
- Most of the time = 1
- Good part of the time = 2
- Some of the time = 3
- A little of the time = 4

The SDS index is derived by dividing the sum of the values (raw scores) obtained on the 20 items by the maximum possible score of 80 and expressed as a decimal point. The table below converts raw scores into SDS index scores.

A Table for the Conversion of Self-Rated Raw Scores to the SDS Index

Raw Score	SDS	Raw Score	SDS	Raw Score	SDS
2	0	4	0.5	6	0
2	0	4	0.5	6	0
2	0	4	0.5	6	0
2	0	4	0.5	6	0
2	0	4	0.5	6	0
2	0	4	0.5	6	0
2	0	4	0.5	6	0
2	0	4	0.5	6	0
2	0	4	0.6	6	0
2	0	4	0.6	6	0
3	0	5	0.6	7	0
3	0	5	0.6	7	0
3	0	5	0.6	7	0
3	0	5	0.6	7	0
3	0	5	0.6	7	0
3	0	5	0.6	7	0
3	0	5	0.7	7	0
3	0	5	0.7	7	0
3	0	5	0.7	7	0
3	0	5	0.7	7	0
				8	1

Source: Zung, W. W. K. (1965). A self-rating scale for depression. *Archives of General Psychiatry, 12*, 63-70.

Appendix F: Center for Epidemiologic Studies Depression Scale (CES-D)

Center for Epidemiologic Studies Depression Scale (CES-D), NIMH

Below is a list of the ways you might have felt or behaved. Please tell me how often you have felt this way during the past week.

	During the Past Week			
	Rarely or none of the time (less than 1 day)	Some or a little of the time (1-2 days)	Occasionally or a moderate amount of time (3-4 days)	Most or all of the time (5-7 days)
1. I was bothered by things that usually don't bother me.	☐	☐	☐	☐
2. I did not feel like eating; my appetite was poor.	☐	☐	☐	☐
3. I felt that I could not shake off the blues even with help from my family or friends.	☐	☐	☐	☐
4. I felt I was just as good as other people.	☐	☐	☐	☐
5. I had trouble keeping my mind on what I was doing.	☐	☐	☐	☐
6. I felt depressed.	☐	☐	☐	☐
7. I felt that everything I did was an effort.	☐	☐	☐	☐
8. I felt hopeful about the future.	☐	☐	☐	☐
9. I thought my life had been a failure.	☐	☐	☐	☐
10. I felt fearful.	☐	☐	☐	☐
11. My sleep was restless.	☐	☐	☐	☐
12. I was happy.	☐	☐	☐	☐
13. I talked less than usual.	☐	☐	☐	☐
14. I felt lonely.	☐	☐	☐	☐
15. People were unfriendly.	☐	☐	☐	☐
16. I enjoyed life.	☐	☐	☐	☐
17. I had crying spells.	☐	☐	☐	☐
18. I felt sad.	☐	☐	☐	☐
19. I felt that people dislike me.	☐	☐	☐	☐
20. I could not get "going."	☐	☐	☐	☐

SCORING: zero for answers in the first column, 1 for answers in the second column, 2 for answers in the third column, 3 for answers in the fourth column. The scoring of positive items is reversed. Possible range of scores is zero to 60, with the higher scores indicating the presence of more symptomatology.

Source: Bienenfeld, D., & Stinson, K. N. (2014). Screening tests for depression.

Appendix G: Physical Sources, Medical Conditions, Medications, and Laboratory Screening Associated with Major Depressive Disorder

Physical Sources	Medical Conditions	Laboratory Screening	Medications
Food allergies	Brain cancers	Recommended:	Amphetamines (withdrawal from)
Hypoglycemia	Cancer (Pancreatic cancer and lung cancer)	Complete blood count with differential	Anabolic steroids
Lack of exercise	Cardiac conditions (e.g., congestive heart failure, ischemic heart disease, myocardial infarction)	TSH	Antihistamines
Nutritional deficiencies		Serum chemistries (Electrolytes, including calcium, phosphate, and magnesium levels) / vitamins (vitamin B12, Vit D)	Anti-inflammatory agents
Poisoning with heavy metals (mercury, lead, aluminum, cadmium, and thallium)	Chronic inflammation		Antipsychotic drugs
	Chronic pain		Benzodiazepines
	Diabetes		Beta-blockers
	Epilepsy	Liver and renal function tests (BUN, LFT)	Cimetidine (Tagamet)
	Fibromyalgia		Corticosteroids (adrenal hormone agents)
Premenstrual syndrome	Head injury	Rapid plasma regin -RPR (blood screening test for syphilis)	
Selenium toxicity	Infections: AIDS, Influenza, infectious mononucleosis, Lyme disease, syphilis (late stage), tuberculosis, viral hepatitis, viral pneumonia		Cyclosporine (an antibiotic)
Sleep disturbances		Urinalysis	
		Urine toxicology screen for drugs of abuse	High blood pressure medications
		Serum alcohol level	Indomethacin (Indocin)
		Urine pregnancy test	Marijuana
		HIV serology	Opioids
	Liver disease	Arterial blood gas (ABG)	Oral contraceptives
	Lung disease		Phencyclidine
	Multiple sclerosis	Not recommended due to little justification for their use in routine screening for psychiatric illness and limited in identifying medical causes of psychiatric disorders	Ranitidine (Zantac)
	Parkinson's disease		Reserpine (Serpasil)
	Rheumatoid Arthritis		Thiazide diuretics
	Stroke		Tranquilizers and sedative-hypnotics
	Thyroid diseases		Vinblastine (anticancer agents)
	Systemic lupus Erythematosus		
		Computed tomography	
		Magnetic resonance imaging	

Sources: Perese, E. F. (2012). Psychiatric Advanced Practice Nursing: A Biopsychosocial Foundation for Practice. Philadelphia, PA: F. A. Davis. / Bienenfeld, D., & Stinson, K. N. (2016). *Screening tests for depression*. Medscape.

Appendix H: Compilation of the Eight Publications Reviewed for This Project.

Depression Diagnostic Tool	Article #1 or Article #2	References for Articles
Beck depression inventory	Article #1	*Campbell, M. H., Maynard, D., Roberti, J. W., & Emmanuel, M. K. (2012). A comparison of the psychometric strengths of the public-domain Zung self-rating depression scale with the proprietary Beck Depression Inventory-II in Barbados. *West Indian Medical Journal*, *61*(5), 483-488.
	Article #2	Jakšić, N., Ivezić, E., Jokić-Begić, N., Surányi, Z., & Stojanović-Špehar, S. (2013). Factorial and diagnostic validity of the Beck Depression Inventory-II (BDI-II) in Croatian primary health care. *Journal of clinical psychology in medical settings*, *20*(3), 311-322.
Hamilton depression rating scale	Article #1	Raimo, S., Trojano, L., Spitaleri, D., Petretta, V., Grossi, D., & Santangelo, G. (2015). Psychometric properties of the Hamilton Depression Rating Scale in multiple sclerosis. *Quality of Life Research*, *24*(8), 1973-1980.
	Article #2	Schneibel, R., Brakemeier, E. L., Wilbertz, G., Dykierek, P., Zobel, I., & Schramm, E. (2012). Sensitivity to detect change and the correlation of clinical factors with the Hamilton Depression Rating Scale and the Beck Depression Inventory in depressed inpatients. *Psychiatry research*, *198*(1), 62-67.
PHQ-9	Article #1	Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cutoff score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *Canadian Medical Association Journal*, *184*(3), E191-E196.
PHQ-2	Article #2	Manea, L., Gilbody, S., Hewitt, C., North, A., Plummer, F., Richardson, R., & McMillan, D. (2016). Identifying depression with the PHQ-2: A diagnostic meta-analysis. *Journal of affective disorders*, *203*, 382-395.
Major depression inventory	Article #1	Bech, P., Timmerby, N., Martiny, K., Lunde, M., & Soendergaard, S. (2015). Psychometric evaluation of the Major Depression Inventory (MDI) as depression severity scale using the LEAD (Longitudinal Expert Assessment of All Data) as index of validity. *BMC Psychiatry*, *15*(1), 190.
	Article #2	Fawzi, M. H., Fawzi, M. M., & Abu Hindi, W. (2012). Arabic version of the Major Depression Inventory as a diagnostic tool: reliability and concurrent and discriminant validity.
Zung self-rating depression scales	Article #1	*Campbell, M. H., Maynard, D., Roberti, J. W., & Emmanuel, M. K. (2012). A comparison of the psychometric strengths of the public-domain Zung self-rating depression scale with the proprietary Beck depression inventory-II in Barbados. *West Indian Medical Journal*, *61*(5), 483-488.
	Article #2	Ruiz-Grosso, P., de Mola, C. L., Vega-Dienstmaier, J. M., Arevalo, J. M., Chavez, K., Vilela, A., . . . & Huapaya, J. (2012). Validation of the spanish center for epidemiological studies depression and Zung self-rating depression scales: a comparative validation study. *PloS one*, *7*(10), e45413.
Center for Epidemiologic Studies depression scale	Article #1	Levine, S. Z. (2013). Evaluating the seven-item Center for Epidemiologic Studies Depression Scale short-form: a longitudinal US community study. *Social psychiatry and psychiatric epidemiology*, *48*(9), 1519-1526.

	Article #2	Chin, W. Y., Choi, E. P., Chan, K. T., & Wong, C. K. (2015). The psychometric properties of the Center for Epidemiologic Studies Depression Scale in Chinese primary care patients: factor structure, construct validity, reliability, sensitivity, and responsiveness. *PloS one*, *10*(8), e0135131.

Appendix I: Tables

Table 13

Demographics, Health Care Settings, and Advantages and Disadvantages of the Six Depression Assessment Tools

Depression Assessment Tools	Reliability of the Tools	Age Groups or Demographics for Use with the Tool	Type of Health Care Settings	Advantages of Assessment Tools	Disadvantages of Assessment Tools	Cutoff Scores	Cost and Reference
Beck's depression inventory-II	First article (by Campbell et al., 2012) $\propto = 0.88$ Second article (by Jakšić et al., 2013). At a cutoff score of 12 indicating MDD, $\propto = 0.89$	Paper version appropriate for age 13 to age 80	Brevity and simplicity of BDI makes it a good fit or any setting such as inpatient, outpatient, or primary care	-Widely and thoroughly researched depression tool -Simple to complete, taking approximately 5 minutes -Available in Spanish - Use of cutoff scores to designate the level of depression -BDI more psychometrically adequate compared with the Zung SDS as seen in 53% of respondents completing the Zung SDS compared with 81% completing the entire assessment questions for BDI -Wording of Beck's less problematic for patients to complete -BDI-II was translated into many languages - Has shown good qualities as a tool for screening and measuring the severity of depression -BDI-II has shown strong and comparable psychometric properties regarding internal consistency, factor structure, and convergent validity	-Copyrighted, -Costs $ (not in public domain) -One BDI item ("loss of interest in sex") has low response rate from survey participants, which was below 95% -Assessment tool is not free, copyrighted and priced per assessment -Across multiple studies, BDI-II has inconsistent factor structure and is controversial -Cutoff scores are not culturally independent even though the expression of depression may differ cross-culturally	-Optimal cutoff for MDD $\geq$ 12 -Range of score for BDI = 0-63 with 0-13 (minimal depression), 14-19 (mild), 20-28 = moderate, and 29-63 = severe	Cost for assessment paper-based or online and scoring tools. Sold by Pearsons at https://www.pearsonclinical.com/psychology/products/100000159/beck-depression-inventory-ii-bdi-ii.html

(table continues)

Depression Assessment Tools	Reliability of the Tools	Age Groups or Demographics for Use with the Tool	Type of Health Care Settings	Advantages of Assessment Tools	Disadvantages of Assessment Tools	Cutoff Scores	Cost and Reference
Hamilton depression rating scale	Article 1 (Raimo, Trojano, Spitaleri, Petretta, Grossi, & Santangelo, 2015) $\propto = 0.8$ ---------- Article 2 (Schneibel et al., 2012) (No sensitivity or specificity rating on the article)	Observer-rated	Hospital setting and outpatient	- Easy to administer and acceptable -Has fair internal consistency with Cronbach's alpha at 0.8 Available in multiple languages such as French, German, Italian, Thai, and Turkish. Has interactive voice response version (IVR) -Available in public domain	-Assessment based on observer-report may be susceptible to incorrect evaluation based on personal traits -Assessment procedures can be time-consuming	A cutoff of 14.5 can identify depressive symptoms	http://dcf.psychiatry.ufl.edu/files/2011/05/HAMILTON-DEPRESSION.pdf
PHQ-9/ PHQ-2	Article 1 (Manea et al., 2012) ---------- Article 2 (Manea et al., 2016) PHQ-2:	PHQ-9 & PHQ-2 are self-rating tools	PHQ-9: Used to screen for and measure symptoms of depression in primary care setting PHQ-2: Comprised of the first two questions of PHQ-9 (low mood and loss of interest/pleasure)	PHQ-9: Especially designed to capture depression diagnosis and not for measuring depression severity. - -Free assessment tool according to Pfizer Inc, the legal copyright holder -Takes 5-10 minutes to complete the questionnaire and 1 minute to score - Available in multiple languages Self-rating scale susceptible to an erroneous report by patients	PHQ-9 defined by the DSM-IV symptoms of depression, thus, not designed for ICD-10 depression -Variation of sensitivity and specificity at different cutoff points, leading to the possibility of false-positive results	For PHQ9: Cutoff score of 10 recommended for diagnosing depression 1-4: minimal 5-9: mild 10-14: moderate 15-19: moderately severe 20-27 severe For PHQ-2: cutoff ≥ 3 indicative of depression	Free Available on http://www.phqscreeners.com

(table continues)

Depression Assessment Tools	Reliability of the Tools	Age Groups or Demographics for Use with the Tool	Type of Health Care Settings	Advantages of Assessment Tools	Disadvantages of Assessment Tools	Cutoff Scores	Cost and Reference
Major depression inventory	Article 1 (Bech et al., 2015) ---------- Article 2 (Fawzi et al., 2012) $\propto = 0.91$	Self-reporting mood questionnaire, usable with ICD-10 and DSM-V criteria for initial assessment and follow-up management of depression. -Developed by World Health Organization	Can be used in any health care setting	-Can be used for depression severity measurement from 0 to 50 scale and also as a diagnostic scale. Translated into multiple languages. Free depression tool According to Bech et al. (2015), "MDI is superior to the Zung SDS and the BDI" -Available in other languages (e.g., German, Turkish, Spanish, Danish, French)	Multiple options for questions 8 and 10, novice clinician or patient can double count	Cutoff score of 21 = normal 21-25 = mild 26-30= moderate 31-50 = severe	Free Available online for download Reference: https://psychology-tools.com/major-depression-inventory/
Zung's SDS	Article 1 (Campbell et al., 2012) For Zung SDS: $\propto = 0.85$ ---------- Article 2: Ruiz-Grosso et al. (2012) For Zung SDS: $\propto = 0.89$	Adult	Can be administered in any health setting, self-report tool but administered by an interview in the literature for patients with low literacy level	No cost to assessor (available in public domain), Multiple versions with modified recent -Short, simple and includes all the symptoms of depression -Easy to complete in as little as 5 minutes, scoring instructions simple, no training requirements -Validating instrument for depressive in primary care setting and diverse patient populations -Can be used for monitoring changes in depression level	-The response rate to the 20 Zung SDS was lower than 95% from one survey indicating a problem with adequacy or wording of Zung SDS questions	Total score ranges from 0-100 25-49 = normal 50-59 = mild 60-69 = moderate > 69 = severe	Free tool found in public domain: https://psychology-tools.com/zung-depression-scale/
6 Center for Epidemiologic Studies depression scale (CES-D)	Article 1 (Chin, Choi, Chan, & Wong, 2015) -McDonald's Omega hierarchical value = 0.855 ---------- Article 2 (Levine, 2013)	Telephone and self-administered versions	Can be used by clinicians for diagnosis and to manage depression over time -Used in research work for depression screening	Reliability, sensitivity, and responsiveness levels high CES-D is a good screening tool for depression, monitor disease progression and valid for use in cross-cultural comparative studies. -One of the most widely used instruments in psychiatric epidemiology -Free assessment tool	Self-rating tool- susceptibly patient not completing all questions on questionnaire 10% of respondents cease halfway on the 20-item CES-D (Levine, 2013)	On CES-D: cutoff of 16 or greater indicative of depression CES-D- SF- cutoff greater than or equal to 8	Free: Available in public domain http://cesd-r.com/

(table continues)

While Tables 1, 2 and I3 answered the practice questions, Table I4 is a summary of the articles reviewed for this project. The table contains details of the articles, the depression tools evaluated by those articles, the study design, methods of data analysis, level of evidence, the setting, country, demographics of participants, and the outcomes of the study.

Table I4

Articles Reviewed for This Project

Depression Inventory Tool	Article	Author, Year	Study Design/Methods	Level of Evidence	Setting/ Country	Participants	Outcome of Study
Beck depression inventory	1	Campbell et al. (2012)	-Study compared psychometric properties of BDI-II and Zung SDS using correlational analyses -Missing value analyses and corrected item-total correlations were reported as well for each assessment tool	Peer-reviewed	Barbados (Caribbean)	*N* = 415 undergraduate students age *M* = 25.2 75% females 25% males	BDI-II demonstrated marginal superior internal consistency, reliability, and psychometric properties than the Zung SDS -Strong correlation between BDI-II and Zung's SDS
	2	Jakšić, Ivezić, Jokie-Begic	Purpose of the study was to examine the diagnostic validity of BDI-II Receiver operating characteristics analysis	Peer-reviewed	Croatia	*N* = 314 participants from a medical outpatient setting Ages 25-87 years	Reliability proved to be high using internal consistency and Velicer Minimum Average Partial test
Hamilton depression rating scale	1	Raimo et al. (2015)	-Goal to assess psychometric properties of HDRS in screening for depression in patients with Multiple sclerosis -Internal consistency using Cronbach's alpha with ≥ 70 acceptable -Less than 5% missing values or invalid items on the HRDS questionnaire -Data analysis using SPSS version 20 -Assessment of construct convergent validity, and divergent validity -Clinical diagnosis of depression using DSM-IV	Peer-reviewed	Italy	70 selected patients (of 100 screened) with multiple sclerosis in a hospital setting from ages 22-68 with mean age of 43.3 11 males 59 females	HDRS is an easy tool to administer, with a fair consistency. -Good convergent and divergent validity concerning Neuropsychiatric inventory (NPI) subdomains of depression
	2	Schneibel et al. (2012)	Research conducted to determine discrepancies and predictive abilities of HDRS compared with BDI-II in sensing depression severity change -Use of analysis of variance	Peer-reviewed	Germany	*N* = 105 hospitalized patients with mean age of 41.6	HDRS showed a superior sensitivity to change compared with BDI-II

(table continues)

Depression Inventory Tool	Article	Author, Year	Study Design/Methods	Level of Evidence	Setting/ Country	Participants	Outcome of Study
PHQ-2/PHQ9	1	Manea et al. (2012)	Goal of project was to summarize psychometric properties of PHQ-9 across a range of studies Random-effects bivariate meta-analysis on the optimal cutoff score for diagnosing depression using the PHQ-9 across 18 studies	Meta-analysis	United Kingdom	N=7,180 participants with age range of 24.8 to 71.4	The cutoff between 8 and 12 positively detected depression -Recommendation that assessment tools should be used in conjunction with DSM-V or ICD-10 for diagnosing depression -Same cutoff score not recommended for all settings
	2	Manea et al. (2016)	-Research was done to review studies that had reviewed psychometric properties of PHQ-2 -Quality assessment was done at study level using the revised tool for the quality assessment of diagnostic accuracy studies (QUADAS-2) -Databases searched are MEDLINE, PsycINFO, gray literature databases from inception to Aug 2014 using search terms based on the inclusion and exclusion criteria	Data from studies combined using bivariate diagnostic meta-analysis	United Kingdom	19 of 21 studies met inclusion criteria totaling $N = 11,175$ What is age of participants?	Data cutoff point of ≥ 3 indicative of depression while the authors recommended reading results of the test with caution due to high false rated. Authors believed the sensitivity of the PHQ-2 was lower than that reported in the original validation study, hence suggesting lowering the cutoff point
Major depression inventory	1	Bech et al., (2015)	-Objective was to evaluate MDI as a depression severity scale using longitudinal expert assessment of all data as an index for validity. -The article used data from two previously published articles: Study 1 (Martiny et al., 2005) used a randomized, double-blind trial, sample of 102 patients Study 2 (Straaso et al., 2014) used a randomized, double-blind controlled dose-remission-remission study The data were analyzed using SAS statistical package version 9.0.0	Peer-reviewed	Denmark	Study 1: 102 patients with mean age of 44.7 Study 2: 65 patients with mean age of 48.1	MDI accepted for measuring depression at a cutoff score of 21, 26, 31 for mild, moderate, and severe depression respectively
	2	Fawzi et al. (2012)	Article sought to translate MDI into the Arabic language, test its reliability, concurrent and discriminant validity of Arabic version of MDI -Used descriptive statistics, chi-square test for categorical and *t* test -Cronbach for internal consistency reliability of the MDI-A	Peer-Reviewed	Egypt	$N = 100$ (50 Egyptian outpatients with a diagnosis of MDD, and 50 healthy controls) The age range of 18-60 years	MDI-A (Arabic) has excellent reliability, an acceptable concurrent and discriminant validity MDI-A has a strong positive correlation of scores with BDI

(table continues)

Depression Inventory Tool	Article	Author, Year	Study Design/Methods	Level of Evidence	Setting/ Country	Participants	Outcome of Study
Zung self-rating scale	1	Campbell et al. (2012)	Study compared psychometric properties of Zung SDS and BDI-II using correlational analyses -Missing value analyses and corrected item-total correlations were reported as well for each assessment tool -Research conducted in academic setting	Peer-reviewed	Barbados (Caribbean)	*N* = 415 undergraduate students with a mean age of 25.2 75% females 25% males	Zung SDS demonstrated lower inferior internal consistency, reliability, and psychometric properties compared with BDI-II Zung SDS is a valuable tool in assessment of adolescents and young adults in large treatment settings
	2	Ruiz-Grosso et al. (2012)	Study was to validate and compare psychometric properties of the Spanish version of ZSDS with CES-D -Cross-sectional study -Cronbach's Alpha and Hierarchical McDonald Omega for polychromic variables were used to establish validity	Peer-reviewed	Peru	*N* = 70 patients	Zung SDS is a valid instrument for detecting depression in a clinical setting and can be useful in epidemiological research
Center for Epidemiologic Studies depression scale (CES-D)	1	Chin et al. (2015)	Study was to validate the use of CES-D in Chinese primary care patients regarding its psychometric properties -Assessed convergence validity by reviewing correlation between CES-D, PHQ-9, and Short Form-12 Health Survey (version 2) Mental Component Summary, Internal consistency assessed with McDonald's Omega hierarchical (wH)	Peer-reviewed	Hong Kong	*N* = 3,686 Chinese adult primary care patients in Hong Kong Age *M* = 49.4 58.1% female 49.9% male	CES-D has a strong correlation with PHQ-9 Two-week test-retest reliability was high
	2	Levine S. (2013)	Study goal was to evaluate potential use of the CES-D- SF (with seven items) -Reviewed bivariate correlation between CES-D and CES-D-SF	Peer-reviewed	U.S.A	*N* = 8500 Ages 28-51 and over	Compared with CES-D, the CES-D-SF also has high internal specificity