THE EFFECTS OF PSYCHOLOGICAL HEALTH ON LUMBAR DISCECTOMY OUTCOMES

by

Hannah Van Brenk

Bachelor of Science in Kinesiology (Honours), University of New Brunswick, 2023

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science in Kinesiology

in the Graduate Academic Unit of Kinesiology

Supervisor: Jeffrey Hebert, PhD., Faculty of Kinesiology

Examining Board: Jonathan Edwards, PhD., Faculty of Kinesiology, Chair

Stephan Dombrowski, PhD., Faculty of Kinesiology

Niels Wedderkopp, MD, PhD., Department of Clinical Research, University of Southern Denmark

Gwyneth de Vries, MD., FRCSC, Department of Surgery, Dalhousie University

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

June, 2024

ABSTRACT

Lumbar disc herniations often require surgery, psychosocial factors, specifically depression, predict poor surgical outcomes. The study aimed to investigate the effects of preoperative depression on post-operative pain and disability following discectomy surgery. Psychological health was measured using the PHQ-8 and the Mental Component Score of the SF-12. Disability and pain were collected using the Oswestry Disability Index and Numeric Pain Rating Scales. Data included from 1104 patients (mean[SD]age = 43.90 [13.13]; 52% female sex) in Poisson Regression models showed depression increased the risk of both leg pain and overall poor surgical outcomes (IRR[95% CI]= 1.67[1.14 to 2.50] ; 1.88[1.48 to 2.40]; 1.50[0.99 to 2.29]; 1.67[1.29 to 2.16]; 1.56 1.09 to 2.24] ;1.72[1.40 to 2.10]; 1.61[1.07 to 2.43]; 1.90[1.46 to 2.47]; 1.92[1.20 to 3.09];2.08[1.51 to 2.86]; 1.90 [1.21 to 22.61], 1.93[1.55 to 2.40]) . Psychological factors do have influence on recovery following discectomy surgery and should be acknowledged pre-operatively.

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Jeff Hebert, my supervisory committee, and everyone at the Canada East Spine Centre for their dedication and mentorship.

Thank you to the University of New Brunswick for the funding opportunities provided as well as fostering an amazing campus atmosphere for the duration of my education here.

To my family, husband, and incredible grad lounge friends, thank you for your unwavering support and encouragement. This degree would not have been possible without all of you, you mean the world to me.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

Magnetic resonance Imaging (MRI)

Electromyography (EMG)

Nerve conduction study (NCS)

Somatosensory evoked potentials (SSEP)

Non-steroidal anti-inflammatory drugs (NSAID)

Oswestry Disability Index (ODI)

Numeric Pain Rating Scale (NPRS)

Minimal Clinical Important Change (MCIC)

Minimal Detectable Change (MDC)

Patient Acceptable Symptom State (PASS)

Patient Health Questionnaire (PHQ)

Short Form Health Survey- 36 (Sf-36)

Short Form Health Survey- 12 (SF-12)

Physical Component Score (PCS)

Mental Component Score (MCS)

Canadian Spine Outcomes Research Network (CSORN)

Chapter 1: Introduction

Low back pain is one of the leading causes of disability globally, in 2020 there were 619 million recorded cases of low back pain [1]. Included in these cases are degenerative diseases of the lumbar spine and which create compressive forces on spinal nerves, herniated or bulging discs and are a primary underlying cause of radiculopathy [2–4]. Furthermore, lumbar disc herniations often require surgical intervention.

The vast majority of lumbar radiculopathy cases are caused by lumbar disc herniations [4]. Lumbar radiculopathy results in pain descending into the knee or foot. Common symptoms include; numbness, tingling, prickling, pain to light touch, cold, and heat [5]. About 5-10% of the general population suffering from low back pain will have radiculopathy, with a lifetime prevalence of 49 to 70% [6]. Lumbar radiculopathy caused by disc herniation is a debilitating condition and has a significant impact on an individual's quality of life if not adequately treated.

A variety of treatment options exist for lumbar disc herniations. Conservative treatments are considered first and can include; patient education, active exercise, manual and physical therapy, and nonsteroidal anti-inflammatory drugs (NSAIDs) [2]. While non-invasive treatments may be effective, the majority of patients undergo surgical intervention. Discectomy surgery is considered the gold standard surgical treatment and involves removing the abnormal disc material that is compressing the nerve root or spinal cord [7,8].

Disc disorders, back pain and radiculopathy cost the United States health care system over $1 billion annually, with $300 million being spent on discectomy surgeries [9]. In addition, the number of lumbar surgeries in Ontario, Canada rose from 6.2 to 14.2 procedures per 100 000 between 1993 and 2012. Lumbar disc degeneration and herniation accounted for 4% of these observed cases. Furthermore, direct medical costs rose from $13 million in 2002 to $24 million in 2012 with the overall total of $176 million in combined medical costs [10].

Depression and depression risk are known predictors of lumbar surgical outcomes. Increased pre-operative depressive symptoms can result in increased pain and disability, decreased return to work, and decreased surgical satisfaction following lumbar surgery [11–13]. Baseline depression is associated with increased pain 6-months post-operatively for lumbar spinal stenosis surgery [14]. In addition, patients with depressive symptoms had higher Oswestry Disability Index (ODI) and pain scores 3-months and 10-years post operatively following lumbar stenosis surgery [13]. While depression is correlated with negative surgical outcomes following lumbar surgery, the effects of depression on post-operative pain and disability after discectomy surgery is currently unknown.

Given the prevalence, cost, and invasive nature of surgical intervention for disc related lumbar radiculopathies, research in this area is crucial for informed clinical decision making. Therefore, the aim of the current study are to estimate the effects of

preoperative psychological factors on postoperative pain and disability measures in patients undergoing discectomy surgery for treatment of lumbar radiculopathy.

Chapter 2: Literature Review

2.1 Lumbosacral Radiculopathy

Lumbosacral Radiculopathy is one of the most common disorders seen by neurosurgeons. Degenerative diseases of the lumbar spine are the primary underlying cause of radiculopathy and create compressive forces on the spinal nerves through herniated or bulging discs [2–4]. Sciatica is the most common type of lumbar radiculopathy and is characterized by radiating pain in the leg that descends below the knee and into the foot. Roughly 90% of radiculopathy cases are caused by lumbar disc herniations [4]. Patients experiencing lumbar radiculopathy may describe the pain as sharply dull, aching, burning, throbbing, electric, or sharp [2,3]. L4-L5 radiculopathy is most common and is often caused by disc herniations, although L1-L4 radiculopathy is possible, but rare [3]. The Spine Outcomes Research Trial found that L4-L5 and L5-S1 herniations are the most common [15]. Common symptoms described by individuals with disc related lumbar radiculopathy are; numbness sensation [reported by 94%], sudden pain attacks tingling or prickling [88%], burning, pain to light touch, and pain to cold or heat were also reported [61%, 37 %, 33% respectively] [5].

Diagnosis of lumbar radiculopathy can be a done through a combination of clinical diagnosis and imaging [2]. History and clinical diagnosis are the most common method for diagnosing lumbar radiculopathy. Patients will report pain that radiates through the leg that will follow a predictable dermatomal pattern [6]. If the pain is related

to a herniated disc, it will be exacerbated by leaning forward, sitting, coughing or straining and will be relieved by laying down or walking [3]. There are a variety of physical examinations that can be performed; Lesegue's sign is when leg pain is reproduced in the gluteal region with passive straight leg raising, and can be used to diagnose if the radiculopathy is caused by nerve root compression [sensitivity of 91%, specificity of 26%] [2,6]. In addition, L4 radiculopathy causes weak knee extension and hip abduction, and a foot drop is also seen with L5 radiculopathy. S1 radiculopathies are commonly caused by disc herniations and are associated with weak plantar flexion, knee flexion and hip extension [3]. The Crossed Straight Leg test is able to determine if the radiculopathy is caused by disc herniation. The test involves the unaffected leg being flexed and the patient may encounter pain localized on the affected side. The test has a high specific [88%] for disc related lumbar radiculopathy, but only a sensitivity of 29% [6]. While clinical and historical diagnosis are accurate, Magnetic Resonance Imaging (MRI) is the preferred way to conclude the cause of the reported lumbar radiculopathy [16–18]. Due to the high prevalence of disc herniations and degenerative spinal stenosis in older patients, historical and physical examination can help determine the relevance of MRI findings [17,18]. MRI imaging without contrast is the optimal method as it allows for the compression of the nerve root to be more visible. Electromyography (EMG), nerve conduction velocities (NCV) and somatosensory evoked potentials (SSEP) can help differentiate the radiculopathy from other disorders of the peripheral nervous system but are only necessary in complex cases [2].

The development of lumbar radiculopathy is most commonly caused by nerve root compression, either from disc herniation or spondylosis. Disc herniations can be by caused by an acute injury or chronic degeneration of the spine and intervertebral disc. Disc herniations cause activations of the pain fibers surrounding the tissues such as ligaments, vessels or dura mater. In addition, spondylosis results in the narrowing of the spinal canal, neural foramen or the lateral recess. The most common cause of this canal narrowing is degenerative arthritis of the spine, other possible causes are infection, trauma or vascular disease [19]. Risk factors for the development of lumbar radiculopathy are; 50 years of age or older, history of cancer or unexplained weight loss [3].

2.2 Treatment/Therapy

There are various treatments available for lumbar radiculopathy, and surgical treatment is considered only when conservative treatments have failed. The initial management of acute radiculopathy symptoms is rest and physical therapy. Progressive neurological deficits, such as bowel and bladder disfunction, or numbness are considered an emergency. However, slower symptom development (motor weakness and paresthesia) can be consulted by spinal surgery to discuss conservative and non-conservative treatments [4,9]. Conservative treatments can include; patient education, active exercise, manual and physical therapy, and nonsteroidal anti-inflammatory drugs (NSAIDs). If conservative treatments fail, spinal or pain injections are often the next treatment. Spinal injections can include epidural steroid injections, facet injections or transforaminal injections, and have all been shown to provide long term relief [2].

Normally, if 6 weeks of conservative treatments are unsuccessful, surgery is then considered. The timing of surgery is based on symptom severity and has been shown to be of greater benefit to patients with more severe symptoms [2,9]. An observational cohort study that followed patients with lumbar radiculopathy and herniated discs for 8 years determined that patients who received surgical treatment by 6 weeks had superior surgical outcomes [15].

Discectomy surgery as treatment for lumbar radiculopathy caused by a disc herniation is considered the gold standard by the spinal surgical community [8]. A systematic review has also found that the discectomy procedures produce the most satisfactory results including symptom relief [9].

Surgical Technique

1. A 4-5cm incision is made from one spinous process to another, depending on herniation level. The skin is pulled back using a Beckman retractor;
2. The aponeurosis is cauterized and the multifidus is released from the spinous process on one side using a Cobb bone curette, until the joint is visible. The muscle is retracted using a Taylor retractor on the lateral side of the joint;
3. The ligamentum flavum is exposed and fluoroscopic control is done to confirm the herniation level;
4. The ligamentum flavum is opened and dura mater pushed back to expose the compressed neural structures;

5. Disc forceps are used to remove the herniated disc material, leaving behind the remaining healthy disc;
6. Any potential Cerebrospinal fluid leak is cleared and the patient is closed [7]

2.3 Economic consequence

Discectomy and low back surgery pose a large cost burden, both directly and indirectly. An international comparison study done in 1994 found that the United States had 40% more low back surgeries when standardized for surgeons per capita [20]. Today, lumbar radiculopathy is one of the most common complaints evaluated by a spinal surgeon, it is estimated that 3-5% of the population will see a spinal surgeon for lumbar radiculopathy [2]. In general, 5-10% of patients with low back pain suffer are estimated to suffer from radiculopathy with a lifetime prevalence of low back pain being 49-70%. In addition, the annual prevalence of disc related sciatica is 2.2% [6]. Disc disorders, back pain and radiculopathy cost the US health care system over $1 billion annually with $300 million being spent on discectomy surgeries [9]. The number of instrumented lumbar surgeries in Ontario, Canada rose from 6.2 to 14.2 procedures per 100 000 between 1993 and 2012. Lumbar disc degeneration and herniation accounted for 4% of these observed cases. In addition, the direct medical costs rose from $13 million in 2002 to $24 million in 2012, with an overall total of $176 million in medical costs over this period [10].

2.4 Outcome Measures

There are a variety of available tools that can be used to measure postoperative outcomes following discectomy surgery. The current study uses the Oswestry Disability Index (ODI) and the Numeric Pain Rating Scales (NPRS). However, there are a variety of other measures that have been used in previous studies.

2.4.1 Oswestry Disability Index

The ODI was developed to assess pain-related disability in individuals with acute, subacute or chronic low-back pain. A systematic review of the ODI and its publications has found 4 versions since its original publication in 1980 [21,22]. The questionnaire covers 1 item on pain and 9 items on activities of daily living including; lifting, walking, sitting, standing, sleeping, sex life, social life and travelling. Each item is measured on a 6 point scale which ranges from best case scenario to worst case scenario. The total scores can be calculated and range from 0 (no disability) to 100 (maximal disability). The ODI requires little administration or training and takes less than 5 minutes to complete, which provides an easy and reliable way to measure disability in patients [22].

The ODI has been shown to have a high test-retest reliability [r=0.83-0.99 [23]] [r=0.84 [22]] in patients with various low back pain disorders, including lumbar radiculopathy and disc herniations. The ODI also demonstrates adequate internal reliability [chronbach's alpha= 0.74-0.77 [24], chronbach's alpha= 0.71-0.87[23], chronbach's alpha= 0.71-0.81 [22]]. The minimal detectable change (MDC), or the amount of change

that accounts for measurement error, for the ODI for general low back pain is variable [MDC=13.6; SEM=0.880 [25], MDC=4.0-6.0 [26]]. Specifically for patients with degenerative disc disease the MDC is 11.75 with an SEM=4.24 [27]. In addition, the ODI is considered a valid measure when compared with pain scores on the visual analogue scale (VAS) or the numeric pain rating scales (NPRS) [r=0.75, p<0.001] [28]. The modified ODI, with the question regarding sex life removed, showed high correlations with the original ODI [ICC=0.90, CI=0.78-0.96] [29].

The minimal clinically important change (MCIC) is the smallest amount of change in a measurement where a patient sees beneficial change. When looking at the MCIC in general lumbar spine surgery, a systematic review found that the MCIC can range from 4.0-15.0 on a 100 point scale. It was proposed from this study that an MCIC of at least 10.0 was required on the ODI for lumbar spine surgery [30]. In addition, a prospective review study stated that 12.8 points on the ODI was found to be the MCIC in patients undergoing lumbar spine surgery [31]. An alternative approach to the MCIC is the patient acceptable symptom state (PASS) which establishes an absolute cut-off value representing the point at which patient's consider their health to be acceptable. A cross-sectional study looking at the PASS estimate of the ODI concluded a score of greater than or equal to 22.0 on the 100 point scale for surgery of degenerative diseases of the lumbar spine [32]. Furthermore, a cohort determined the criteria for success for "completely better results" on the ODI and found that a change of 20.0 points on the ODI was required [66% success rate] [33].

When looking at discectomy surgery specifically, the MCIC for achieving benefit for disc herniation was 16.9 with a 32% change [sensitivity=77.2%, specificity= 85%]. Furthermore, change for substantial benefit was 22.0 or 60.6% change [specificity=80.6%, sensitivity=72.5%] [34].

2.4.2 Numeric Pain Rating Scales

The Numeric Pain Rating Scale (NPRS) is an 11, 21 or 101 point scale where a patient ranks their pain. However, the NPRS has been found to be sufficient on a smaller 11 or 21 point scale since patients naturally default to 5 or 10 point increments [35]. The end points of the scales are the extremes of no pain or pain as bad as it could be and can be presented graphically or visually. On the 11-point scale, the individual rates their pain between 0 (no pain) and 10 being the worst pain imaginable [36]. NPRS are easy to administer with roughly 90% of patients able to rate their pain on an NPRS with minimal prompting [37].

The NPRS is a reliable measurement with good test-retest scores in patients with lumbar disc herniations [ICC=0.991] [38]. The MDC of the NPRS for patients with low back pain is 2 points on the 0-10 scale [39]. The NPRS is considered a valid measurement of pain when compared to several other pain rating scales used for various pain sources [37,40]. The Spearman correlations range from 0.64-0.90 when compared to the Verbal Descriptor Scale, Revised Faces Pain Scale and the Iowa Pain Thermometer [37]. A survey in university-aged students found a strong correlation [r=0.90] between the visual analogue scale and the NPRS for general pain rating. Students also indicated that 31%

preferred to rank their pain on the NPRS [41]. In low back pain specifically, there is a high correlation between the VAS and NPRS as well [r=0.97, p<0.001] [28].

The MCIC for the 11-point NPRS that exceeds the expected measurement error and represents meaningful clinical change is 2 points [42]. Previous systematic reviews have found that scientific papers use a 50% reduction in pain as a statistically significant change while clinical important change in pain as reported with "much improved" or "very much improved" that related to a 30% reduction in pain [36]. A systematic review on the NPRS for low back pain found 5 studies that used 1.0-4.50 as the MCIC. Using that data, Ostelo et al., suggested that 2.0 points [or 30% improvement] would be an acceptable MCIC [30]. In addition, 2.5 points on a back pain NPRS and 3.5 on a leg pain scale [success rate of 67%] are needed for a "completely better" result when looking at disc herniation [33]. In order to achieve "benefit" following disc herniation surgery an MCIC for back pain is 2.0 [sensitivity=82.7%, specificity=90.0%] and 3.0 for leg pain scales [sensitivity=72.5, specificity=75.6] [34]. For "completely better" to be achieved back pain NPRS must improve by 2.5 and 3.5 for leg pain [success rate of 67%] [33].

2.5 Outcome Predictors

Previous prospective research studies have identified possible predictors of clinical outcomes following lumbar discectomy surgery for lumbar radiculopathy. It has been seen that better preoperative mental health scores produce better outcomes, especially in terms of patient satisfaction following discectomy surgery [43]. The current

study places particular emphasis on psychological factors, specifically depression, and its effects on discectomy outcomes, including postoperative pain and disability.

A systematic review of psychosocial predictors on outcomes following lumbar spine surgery found 8 studies, found that 60% studies had depression as a statistically significant predictor of poor surgical outcomes [12]. In addition, when primarily looking at back pain levels, individuals that have higher depression levels or risks prior to surgery continue to have higher pain levels after surgery [44]. Similarly, 36 (49.3%) patients with relevant depressive symptoms at baseline had higher Visual Analogue Scale pain levels at baseline, 3 -months and 12-months post-discectomy. Furthermore, the impact of depressive symptoms on long-term surgical outcomes was found to be an independent predictor of poor outcomes at follow-up [$t=7.120$, $p<0.001$] [45]. A prospective cohort study found that preoperative depression and somatic awareness were predictive of clinical meaningful improvement in disability and quality of life following lumbar discectomy. Each point on the Zung Depression Scale is associated with an 11% less likelihood of achieving an MCIC on the ODI [46].

The overall impact of depression on lumbar surgical outcomes has been shown that individuals with depressive symptoms are two times less likely to experience disability improvement post-operatively [11]. A prospective study also found that individuals who reported improved back and leg pain following lumbar surgery also reported less distress and depressive symptoms prior to surgery [47]. Depressive symptoms has been related to poorer surgical results following lumbar spinal stenosis surgery. A

systematic review determined that baseline depression on the Zung Depression Scale was associated with increased pain at the 6-month follow-up for lumbar spinal stenosis surgery [P=0.05] [14]. Furthermore, a longitudinal analysis on lumbar spinal stenosis yielded similar post-operative clinical outcomes. Patients with depression were more likely to experience poor leg outcomes [OR=2.10, 95% CI=1.33-3.31] as well as poor overall outcomes [OR=2.03, 95% CI=1.12-3.69] [48]. Observational prospective studies have also found that depressive symptoms result in poorer short and long term post-operative improvement for lumbar spinal stenosis surgery. Patients with continuous depression had less improvements in symptom severity, disability scores, pain intensity and walking capacity at the 3-month follow up [49]. Similarly, patients with higher depressive symptoms had higher Oswestry Disability Index (ODI) scores 3-months and 10 years post-operatively as well as higher VAS scores and shorter walking distances [13]. In a prospective cohort study, it was determined that patients with preoperative depression had poorer ODI and quality of life scores at baseline, 3-and 12-months post lumbar surgery. The depressive symptoms were also associated with a longer hospital stay following surgery [50]. Moreover, individuals that had lower psychological distress levels had lower pain levels follow elective lumbar surgery [p=0.003]. Depression specifically was related to poorer results on the ODI following surgery [P=0.036] [51]. Depressed patients undergoing surgery for degenerative lumbar spine conditions not only show increased scores on the leg VAS, but also on the ODI [52]. A retrospective review of preoperative depression on quality of life outcomes following lumbar fusion surgery

found that greater depression [OR=0.93, P<0.001] and pain [OR=0.99, P=0,02] were associated and resulted in diminished postoperative improvement [53]. Furthermore, patients with depression suffer from back pain more frequently [81-88%] following non-traumatic lumbar surgery compared to non-depressed patients [41%] [54]. Poorer overall mental health scores are also correlated with lower physical function, elevated pain scores and higher disability following minimally invasive transforaminal lumbar interbody fusion [55]. Depression has also been shown to be a negative predictor of post-operative return to work. Only 10.6% of patients with depression returned to work following lumbar fusion surgery [OR=0.38, P<0.001] [56]. Finally, individuals with preoperative depression not only have longer hospital stays, but are often discharged to an external service or care home [OR=1.9, 95% CI=1.1-3.4] [57].

Other Prognostic Factors

Several non-psychological factors have been associated with pain and disability follow lumbar spine surgery. A systematic review of preoperative predictors of discectomy surgery outcomes found that smoking was generally associated with poorer outcomes in terms of VAS-leg, VAS-back and disability questionnaires. In addition, younger individuals generally have better post-operative outcomes [43]. Shorter duration of symptoms preoperatively has also been associated with better postoperative discectomy outcomes. Patients that have 6-weeks or less of preoperative pain show a larger percentile decrease in pain intensity following surgery [58]. In general lumbar surgery,

fewer preoperative symptoms, higher self-rated health, walking ability and functional disability are all correlated with higher postoperative outcomes [14,44].

2.5.1 Patient Health Questionnaire-9

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item self-reported questionnaire for the diagnosing of depression risk according to the criteria in the 4th Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). One of the 9-items deals with suicidal ideation and has been identified as a potential problem when administering to a larger population. The PHQ-9 has therefore been shortened to the PHQ-8 to remove the ninth item [59]. The PHQ-9 is a 3-page document that can be self-administered by the patient. The scores can range from 0-27 since each item is ranked from 0 (not at all) to 3 (nearly every day) [60].

The internal reliability for the PHQ-9 and the PHQ-8 are high with no statistical significance between the two versions [Chronbach's alpha PHQ-9=0.88; Chronbach's alpha PHQ-8= 0.88] [60,61]. In addition, there is high internal reliability with the PHQ-9 administered in person or over the phone [chronbach's alpha=0.84] which provides the clinician with reliable options for administering the questionnaire [62]. The PHQ-9 is a valid measure of depression in the populations previously diagnosed with depression with the group main effect and all paired comparisons with the PHQ-9 being highly significant [$P<0.001$] [63]. When validated against previous depression diagnosis, a score on the PHQ-9 greater than 10 shows an accurate risk for depression [sensitivity= 70%, specificity = 84%] [60]. There is also a correlation between the PHQ-9 and SF-12 in patients undergoing

minimally invasive lumbar discectomy [r>0.647 for 6-week, 12-week, 6-month and 12-month follow-up] [64].

2.5.2 Short Form Health Survey

The Short Form Health survey-36 (SF-36) is a survey that was designed to address the 8 dimensions of health; physical functioning, bodily pain, role limitations due to physical health problems, social functioning, general mental health, role limitations due to emotional problems, vitality and general health. It addresses 36 items in 11 questions and 2 scores can then be obtained from this, the physical component score (PCS) and mental component score (MCS) [65]. It had been suggested that the scores can be counted separately, for two different components, combining the two measures may cause poorer validity due to systematic errors [66]. In addition, the shorter version, the SF-12 has produced nearly identical scores when compared to the SF-36. The shorter version can be particularly useful when distributing to a larger population size [67,68].

The SF-36 is a reliable measure good test-retest scores [ICC=0.843 [38]; ICC= 0.79-0.83 [67]] in patients with chronic low back pain. There is also adequate internal consistency within the questionnaire [chronbach's alpha= 0.60-0.87 [69]; chronbach's alpha= 0.79-0.83 [70]; Chronbach's alpha= 0.859 [71]; Cjronbach's alpha= 0.802 [38]]. In addition, the SF-12 and SF-36 show strong correlations for patients undergoing lumbar surgery [(0.967-0.934, P <.001] [39].The MCS of the SF-12 provides a summary score, with scores ranging from 0 to 100. Individuals scoring below 46 can be categorized as experiencing depression, while scored less than 37 can be categorized as having severe

psychological symptoms [72]. The MDC on the MCS for patients with lumbar disc herniations is 1.71 [95% CI = 1.52-1.91] and the SEM is 3.93 [38]. The Short-form Health Survey is a valid measure with strong Pearson correlation coefficients between the PHQ-8 and SF-12 at all time points following minimally invasive discectomy surgery [$r>0.647$, $P<0.001$] [64].

2.6 Knowledge Gap

This study aims to investigate the current gap in knowledge of depression and depression risk and how it may impact post-surgical pain and disability ratings following discectomy surgery for lumbar radiculopathy.

References (Literature Review)

1. Ferreira M, de Luca K, Haile L, et al. Global, regional, and national burden of low back pain, 1990–2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. 2023;5:e316-e329.

2. Berry JA, Elia C, Saini HS, Miulli DE. A Review of Lumbar Radiculopathy, Diagnosis, and Treatment. *Cureus*. 2019;11(10):e5934. doi:10.7759/cureus.5934

3. Tarulli AW, Raynor EM. Lumbosacral Radiculopathy. *Neurologic Clinics*. 2007;25(2):387-405. doi:10.1016/j.ncl.2007.01.008

4. Valat JP, Genevay S, Marty M, Rozenberg S, Koes B. Sciatica. *Best Practice & Research Clinical Rheumatology*. 2010;24(2):241-252. doi:10.1016/j.berh.2009.11.005

5. Hasvik E, Haugen AJ, Grøvle L. Symptom descriptors and patterns in lumbar radicular pain caused by disc herniation: a 1-year longitudinal cohort study. *BMJ Open*. 2022;12(12):e065500. doi:10.1136/bmjopen-2022-065500

6. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ*. 2007;334(7607):1313-1317. doi:10.1136/bmj.39223.428495.BE

7. Blamoutier A. Surgical discectomy for lumbar disc herniation: Surgical techniques. *Orthopaedics & Traumatology: Surgery & Research*. 2013;99(1, Supplement):S187-S196. doi:10.1016/j.otsr.2012.11.005

8. Jacobs WCH, van Tulder M, Arts M, et al. Surgery versus conservative management of sciatica due to a lumbar herniated disc: a systematic review. *Eur Spine J*. 2011;20(4):513-522. doi:10.1007/s00586-010-1603-7

9. Schoenfeld AJ, Weiner BK. Treatment of lumbar disc herniation: Evidence-based practice. *Int J Gen Med*. 2010;3:209-214.

10. Xu Y, Yen D, Whitehead M, Xu J, Johnson AP. Use of instrumented lumbar spinal surgery for degenerative conditions: trends and costs over time in Ontario, Canada. *Can J Surg*. 2019;62(6):393-401. doi:10.1503/cjs.017016

11. Adogwa O, Parker SL, Shau DN, et al. Preoperative Zung Depression Scale predicts outcome after revision lumbar surgery for adjacent segment disease, recurrent stenosis, and pseudarthrosis. *The Spine Journal*. 2012;12(3):179-185. doi:10.1016/j.spinee.2011.08.014

12. Wilhelm M, Reiman M, Goode A, et al. Psychological Predictors of Outcomes with Lumbar Spinal Fusion: A Systematic Literature Review. *Physiotherapy Research International*. 2017;22(2):e1648. doi:10.1002/pri.1648

13. Tuomainen I, Pakarinen M, Aalto T, et al. Depression is associated with the long-term outcome of lumbar spinal stenosis surgery: a 10-year follow-up study. *The Spine Journal*. 2018;18(3):458-463. doi:10.1016/j.spinee.2017.08.228

14. Aalto TJ, Malmivaara A, Kovacs F, et al. Preoperative Predictors for Postoperative Clinical Outcome in Lumbar Spinal Stenosis: Systematic Review. *Spine*. 2006;31(18):E648. doi:10.1097/01.brs.0000231727.88477.da

15. Lurie JD, Tosteson TD, Tosteson ANA, et al. Surgical versus Non-Operative Treatment for Lumbar Disc Herniation: Eight-Year Results for the Spine Patient

Outcomes Research Trial (SPORT). *Spine (Phila Pa 1976)*. 2014;39(1):3-16. doi:10.1097/BRS.0000000000000088

16. Andersson GB. Epidemiological features of chronic low-back pain. *The Lancet*. 1999;354(9178):581-585. doi:10.1016/S0140-6736(99)01312-4

17. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *JBJS*. 1990;72(3):403.

18. Jensen RK, Harhangi BS, Huygen F, Koes B. Lumbar spinal stenosis. *BMJ*. 2021;373:n1581. doi:10.1136/bmj.n1581

19. Groen GJ, Baljet B, Drukker J. Nerves and nerve plexuses of the human vertebral column. *Am J Anat*. 1990;188(3):282-296. doi:10.1002/aja.1001880307

20. Cherkin DC, Deyo RA, Loeser JD, Bush T, Waddell G. An international comparison of back surgery rates. *Spine (Phila Pa 1976)*. 1994;19(11):1201-1206. doi:10.1097/00007632-199405310-00001

21. Fairbank JCT, Pynsent PB. The Oswestry Disability Index. *Spine*. 2000;25(22):2940.

22. Smeets R, Köke A, Lin CW, Ferreira M, Demoulin C. Measures of function in low back pain/disorders: Low Back Pain Rating Scale (LBPRS), Oswestry Disability Index (ODI), Progressive Isoinertial Lifting Evaluation (PILE), Quebec Back Pain Disability Scale (QBPDS), and Roland-Morris Disability Questionnaire (RDQ). *Arthritis Care & Research*. 2011;63(S11):S158-S173. doi:10.1002/acr.20542

23. Vianin M. Psychometric properties and clinical usefulness of the Oswestry Disability Index. *J Chiropr Med*. 2008;7(4):161-163. doi:10.1016/j.jcm.2008.07.001

24. Cook CE, Garcia AN, Wright A, Shaffrey C, Gottfried O. Measurement Properties of the Oswestry Disability Index in Recipients of Lumbar Spine Surgery. *Spine*. 2021;46(2):E118. doi:10.1097/BRS.0000000000003732

25. Monticone M, Baiardi P, Vanti C, et al. Responsiveness of the Oswestry Disability Index and the Roland Morris Disability Questionnaire in Italian subjects with sub-acute and chronic low back pain. *Eur Spine J*. 2012;21(1):122-129. doi:10.1007/s00586-011-1959-3

26. Cleland J, Gillani R, Bienen EJ, Sadosky A. Assessing Dimensionality and Responsiveness of Outcomes Measures for Patients with Low Back Pain. *Pain Practice*. 2011;11(1):57-69. doi:10.1111/j.1533-2500.2010.00390.x

27. Johnsen LG, Hellum C, Nygaard ØP, et al. Comparison of the SF6D, the EQ5D, and the oswestry disability index in patients with chronic low back pain and degenerative disc disease. *BMC Musculoskelet Disord*. 2013;14:148. doi:10.1186/1471-2474-14-148

28. Shafshak TS, Elnemr R. The Visual Analogue Scale Versus Numerical Rating Scale in Measuring Pain Severity and Predicting Disability in Low Back Pain. *JCR: Journal of Clinical Rheumatology*. 2021;27(7):282. doi:10.1097/RHU.0000000000001320

29. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther*. 2001;81(2):776-788. doi:10.1093/ptj/81.2.776

30. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting Change Scores for Pain and Functional Status in Low Back Pain: Towards International Consensus Regarding Minimal Important Change. *Spine*. 2008;33(1):90. doi:10.1097/BRS.0b013e31815e3a10

31. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and Pain Scales. *The Spine Journal*. 2008;8(6):968-974. doi:10.1016/j.spinee.2007.11.006

32. van Hooff ML, Mannion AF, Staub LP, Ostelo RWJG, Fairbank JCT. Determination of the Oswestry Disability Index score equivalent to a "satisfactory symptom state" in patients undergoing surgery for degenerative disorders of the lumbar spine—a Spine Tango registry-based study. *The Spine Journal*. 2016;16(10):1221-1230. doi:10.1016/j.spinee.2016.06.010

33. Solberg T, Johnsen LG, Nygaard ØP, Grotle M. Can we define success criteria for lumbar disc surgery? : estimates for a substantial amount of improvement in core outcome measures. *Acta Orthop*. 2013;84(2):196-201. doi:10.3109/17453674.2013.786634

34. Power JD, Perruccio AV, Canizares M, et al. Determining minimal clinically important difference estimates following surgery for degenerative conditions of the lumbar spine: analysis of the Canadian Spine Outcomes and Research Network (CSORN) registry. *The Spine Journal*. 2023;23(9):1323-1333. doi:10.1016/j.spinee.2023.05.001

35. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies Comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for Assessment of Pain Intensity in Adults: A Systematic Literature Review. *Journal of Pain and Symptom Management*. 2011;41(6):1073-1093. doi:10.1016/j.jpainsymman.2010.08.016

36. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs*. 2005;14(7):798-804. doi:10.1111/j.1365-2702.2005.01121.x

37. Ware LJ, Epps CD, Herr K, Packard A. Evaluation of the Revised Faces Pain Scale, Verbal Descriptor Scale, Numeric Rating Scale, and Iowa Pain Thermometer in Older Minority Adults. *Pain Management Nursing*. 2006;7(3):117-125. doi:10.1016/j.pmn.2006.06.005

38. Yao M, Xu B ping, Li Z jun, et al. A comparison between the low back pain scales for patients with lumbar disc herniation: validity, reliability, and responsiveness. *Health & Quality of Life Outcomes*. 2020;18(1):1-12. doi:10.1186/s12955-020-01403-2

39. Boden SH, Farley KX, Campbell C, Boden SD, Gottschalk MB. Rational Selection of Patient-Reported Outcomes Measures in Lumbar Spine Surgery Patients. *International Journal of Spine Surgery*. 2020;14(3):347-354. doi:10.14444/7046

40. Karcioglu O, Topacoglu H, Dikme O, Dikme O. A systematic review of the pain scales in adults: Which to use? *The American Journal of Emergency Medicine*. 2018;36(4):707-714. doi:10.1016/j.ajem.2018.01.008

41. Jensen MP, Castarlenas E, Roy R, et al. The Utility and Construct Validity of Four Measures of Pain Intensity: Results from a University-Based Study in Spain. *Pain Medicine*. 2019;20(12):2411-2420. doi:10.1093/pm/pny319

42. Childs JD, Piva SR, Fritz JM. Responsiveness of the Numeric Pain Rating Scale in Patients with Low Back Pain. *Spine*. 2005;30(11):1331. doi:10.1097/01.brs.0000164099.92112.29

43. Wilson CA, Roffey DM, Chow D, Alkherayf F, Wai EK. A systematic review of preoperative predictors for postoperative clinical outcomes following lumbar discectomy. *The Spine Journal*. 2016;16(11):1413-1422. doi:10.1016/j.spinee.2016.08.003

44. Junge A, Dvorak J, Ahrens S. Predictors of bad and good outcomes of lumbar disc surgery. A prospective clinical study with recommendations for screening to avoid bad outcomes. *Spine (Phila Pa 1976)*. 1995;20(4):460-468. doi:10.1097/00007632-199502001-00009

45. Arpino L, Iavarone A, Parlato C, Moraci A. Prognostic role of depression after lumbar disc surgery. *Neurol Sci*. 2004;25(3):145-147. doi:10.1007/s10072-004-0248-x

46. Chaichana KL, Mukherjee D, Adogwa O, Cheng JS, McGirt MJ. Correlation of preoperative depression and somatic perception scales with postoperative disability and

quality of life after lumbar discectomy: Clinical article. *Journal of Neurosurgery: Spine*. 2011;14(2):261-267. doi:10.3171/2010.10.SPINE10190

47. Trief PM, Grant W, Fredrickson B. A Prospective Study of Psychological Predictors of Lumbar Surgery Outcome. *Spine*. 2000;25(20):2616.

48. Hébert JJ, Abraham E, Wedderkopp N, et al. Preoperative Factors Predict Postoperative Trajectories of Pain and Disability Following Surgery for Degenerative Lumbar Spinal Stenosis. *Spine*. 2020;45(21):E1421. doi:10.1097/BRS.0000000000003587

49. Sinikallio S, Aalto T, Airaksinen O, et al. Depression is associated with poorer outcome of lumbar spinal stenosis surgery. *Eur Spine J*. 2007;16(7):905-912. doi:10.1007/s00586-007-0349-3

50. Chapin L, Ward K, Ryken T. Preoperative Depression, Smoking, and Employment Status are Significant Factors in Patient Satisfaction After Lumbar Spine Surgery. *Clin Spine Surg*. 2017;30(6):E725-E732. doi:10.1097/BSD.0000000000000331

51. Amaral V, Marchi L, Martim H, et al. Influence of psychosocial distress in the results of elective lumbar spine surgery. *J Spine Surg*. 2017;3(3):371-378. doi:10.21037/jss.2017.08.05

52. Vilà-Canet G, Covaro A, Isart A, et al. Elective Lumbar Spine Surgery in Depressed Patients: Is it Worth it? *Int J Spine Surg*. 2021;15(3):418-422. doi:10.14444/8062

53. Miller JA, Derakhshan A, Lubelski D, et al. The impact of preoperative depression on quality of life outcomes after lumbar surgery. *Spine J*. 2015;15(1):58-64. doi:10.1016/j.spinee.2014.06.020

54. Järvimäki V, Kautiainen H, Haanpää M, et al. Depressive symptoms are associated with poor outcome for lumbar spine surgery. *Scandinavian Journal of Pain*. 2016;12(1):13-17. doi:10.1016/j.sjpain.2016.01.008

55. Hartman TJ, Nie JW, Zheng E, Oyetayo OO, MacGregor KR, Singh K. Poor mental health scores correlate with inferior outcomes following minimally invasive transforaminal lumbar interbody fusion. *Acta Neurochir*. 2023;165(7):1931-1942. doi:10.1007/s00701-023-05557-y

56. Anderson JT, Haas AR, Percy R, Woods ST, Ahn UM, Ahn NU. Clinical depression is a strong predictor of poor lumbar fusion outcomes among workers' compensation subjects. *Spine (Phila Pa 1976)*. 2015;40(10):748-756. doi:10.1097/BRS.0000000000000863

57. Mancuso CA, Duculan R, Craig CM, Girardi FP. Psychosocial Variables Contribute to Length of Stay and Discharge Destination After Lumbar Surgery Independent of Demographic and Clinical Variables. *Spine*. 2018;43(4):281. doi:10.1097/BRS.0000000000002312

58. Folman Y, Shabat S, Catz A, Gepstein R. Late results of surgery for herniated lumbar disk as related to duration of preoperative symptoms and type of herniation. *Surgical Neurology*. 2008;70(4):398-401. doi:10.1016/j.surneu.2007.04.022

59. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord*. 2009;114(1-3):163-173. doi:10.1016/j.jad.2008.06.026

60. Kroenke K, Spitzer RL, W WJB. The PHQ-9. *Journal of General Internal Medicine*. 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x

61. Shin C, Lee SH, Han KM, Yoon HK, Han C. Comparison of the Usefulness of the PHQ-8 and PHQ-9 for Screening for Major Depressive Disorder: Analysis of Psychiatric Outpatient Data. *Psychiatry Investig*. 2019;16(4):300-305. doi:10.30773/pi.2019.02.01

62. Costantini L, Pasquarella C, Odone A, et al. Screening for depression in primary care with Patient Health Questionnaire-9 (PHQ-9): A systematic review. *Journal of Affective Disorders*. 2021;279:473-483. doi:10.1016/j.jad.2020.09.131

63. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *General Hospital Psychiatry*. 2006;28(1):71-77. doi:10.1016/j.genhosppsych.2005.07.003

64. Lynch CP, Cha EDK, Jenkins NW, et al. Patient Health Questionnaire-9 Is a Valid Assessment for Depression in Minimally Invasive Lumbar Discectomy. *Neurospine*. 2021;18(2):369-376. doi:10.14245/ns.2142162.031

65. McDowell I. General Health Status and Quality of Life. In: *Measuring Health: A Guide to Rating Scales and Questionnaires*. ; 2006:520-703. https://doi.org/10.1093/acprof:oso/9780195165678.003.0010

66. Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: Scoping review. *SAGE Open Med*. 2016;4:2050312116671725. doi:10.1177/2050312116671725

67. Jenkinson C, Layte R, Jenkinson D, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *Journal of Public Health*. 1997;19(2):179-186. doi:10.1093/oxfordjournals.pubmed.a024606

68. Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care*. 1996;34(3):220.

69. Omari OA, Alkhawaldeh A, ALBashtawy M, Qaddumi J, Holm MB, AlOmari D. A Review of the Short Form Health Survey–Version 2. *Journal of Nursing Measurement*. 2019;27(1):77-86. doi:10.1891/1061-3749.27.1.77

70. LoMartire R, Äng BO, Gerdle B, Vixner L. Psychometric properties of Short Form-36 Health Survey, EuroQol 5-dimensions, and Hospital Anxiety and Depression Scale in patients with chronic pain. *Pain*. 2020;161(1):83-95. doi:10.1097/j.pain.0000000000001700

71. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160-164.

72. Gill SC, Butterworth P, Rodgers B, Mackinnon A. Validity of the mental health component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of

common mental disorders in the general population. *Psychiatry Research*. 2007;152(1):63-71. doi:10.1016/j.psychres.2006.11.005

73. Purvis TE, Neuman BJ, Riley LH, Skolasky RL. Comparison of PROMIS Anxiety and Depression, PHQ-8, and GAD-7 to screen for anxiety and depression among patients presenting for spine surgery. *J Neurosurg Spine*. Published online January 18, 2019:1-8. doi:10.3171/2018.9.SPINE18521

74. Kang JDY, Schafer JL. Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. *Statistical Science*. 2007;22(4):523-539.

75. Imai K, Ratkovic M. Covariate Balancing Propensity Score. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 2014;76(1):243-263. doi:10.1111/rssb.12027

76. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*. 2009;28(25):3083-3107. doi:10.1002/sim.3697

77. Rubin DB. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology*. 2001;2(3-4):169-188.

78. Greenland S. Interpretation and Choice Effect Measures in Epidemiologic Analyses. *American Journal of Epidemiology*. 1987;125(5):761-768. doi:10.1093/oxfordjournals.aje.a114593

79. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Medical Research Methodology*. 2018;18(1):63. doi:10.1186/s12874-018-0519-5

80. Yao C, Zhang Y, Lu P, et al. Exploring the bidirectional relationship between pain and mental disorders: a comprehensive Mendelian randomization study. *The Journal of Headache and Pain*. 2023;24(1):82. doi:10.1186/s10194-023-01612-2

81. Schisterman EF, Cole SR, Platt RW. Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies. *Epidemiology*. 2009;20(4):488-495.

82. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/M16-2607

83. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656-664. doi:10.1093/aje/kwn164

84. Hébert JJ, Abraham E, Wedderkopp N, et al. Preoperative Factors Predict Postoperative Trajectories of Pain and Disability Following Surgery for Degenerative Lumbar Spinal Stenosis. *Spine*. 2020;45(21):E1421. doi:10.1097/BRS.0000000000003587

85. Börsbo B, Peolsson M, Gerdle B. The complex interplay between pain intensity, depression, anxiety and catastrophising with respect to quality of life and disability. *Disability and Rehabilitation*. 2009;31(19):1605-1613. doi:10.1080/09638280903110079

Chapter 3: Scientific Article

3.1 Introduction

Low back pain is one of the leading causes of disability globally, in 2020 there were 619 million recorded cases of low back pain [1]. Included in these cases are degenerative diseases of the lumbar spine which can create compressive forces on spinal nerves, herniated or bulging discs and are a primary underlying cause of radiculopathy [2–4]. Furthermore, lumbar disc herniations often require surgical intervention.

The vast majority of lumbar radiculopathy cases are caused by lumbar disc herniations [4]. Lumbar radiculopathy results in pain descending into the knee or foot. Common symptoms include; numbness, tingling, prickling, pain to light touch, cold, and heat [5]. About 5-10% of the general population suffering from low back pain will have radiculopathy, with a lifetime prevalence of 49 to 70% [6]. Lumbar radiculopathy caused by disc herniation is a debilitating condition and has a significant impact on an individual's quality of life if not adequately treated.

A variety of treatment options exist for lumbar disc herniations. Conservative treatments are considered first and can include; patient education, active exercise, manual and physical therapy, and nonsteroidal anti-inflammatory drugs (NSAIDs) [2]. While non-invasive treatments may be effective, the majority of patients undergo surgical intervention. Discectomy surgery is considered the gold standard surgical treatment and

involves removing the abnormal disc material that is compressing the nerve root or spinal cord [7,8].

Disc disorders, back pain and radiculopathy cost the United States health care system over $1 billion annually, with $300 million being spent on discectomy surgeries [9]. In addition, the number of lumbar surgeries in Ontario, Canada rose from 6.2 to 14.2 procedures per 100 000 between 1993 and 2012. Lumbar disc degeneration and herniation accounted for 4% of these observed cases. Furthermore, direct medical costs rose from $13 million in 2002 to $24 million in 2012 with the overall total of $176 million in combined medical costs [10].

Depression and depression risk are known predictors of lumbar surgical outcomes. Increased pre-operative depressive symptoms can result in increased pain and disability, decreased return to work, and decreased surgical satisfaction following lumbar surgery [11–13]. Baseline depression is associated with increased pain 6-months post-operatively for lumbar spinal stenosis surgery [14]. In addition, patients with depressive symptoms had higher Oswestry Disability Index (ODI) and pain scores 3-months and 10-years post operatively following lumbar stenosis surgery [13]. While depression is correlated with negative surgical outcomes following lumbar surgery, the impacts of depression on post-operative pain and disability after discectomy surgery is currently unknown.

Given the prevalence, cost, and invasive nature of surgical intervention for disc related lumbar radiculopathies, research in this area is crucial for informed clinical

decision making. Therefore, the aim of the current study are to estimate the effects of preoperative psychological factors on postoperative pain and disability measures in patients undergoing discectomy surgery for treatment of lumbar radiculopathy.

3.2 Methods

3.2.1 Study Design and Participants

Data was collected from patients seen at 1 of 22 spine centers participating in the Canadian Spine Outcomes and Research Network (CSORN). The network centers collect patient-level data to investigate various clinical outcomes for patients undergoing surgery for a variety of spinal pathologies. The current study used prospectively collected CSORN data from patients who underwent lumbar discectomy surgery for lumbar radiculopathy. Disability and pain outcomes were collected and analyzed pre-operatively, and 3, 12 and 24 months post-operatively. The study protocol was ethically approved by the University of New Brunswick (REB # 2023-188)

We included data from patients who underwent open or minimally invasive discectomy surgery for treatment of lumbar radiculopathy caused by lumbar disc herniation. Individuals with fusion procedures, deformity, fracture, infection or inflammatory spine disease, spondylolisthesis, degenerative spinal stenosis, tumour or any previous spinal surgery were excluded.

3.2.2 Psychological Health

Depression or Depression Risk

Depression and Depression risk were collected only at the pre-operative time point. The PHQ-8 was used to measure depression risk. The PHQ-8 is a validated measure of depression or depression risk in the general population [59] as well as in patients undergoing spinal surgery [73]. Scores equal to or greater than 10 on the PHQ-9 are associated with a moderate-to-severe risk of depression. This cut-off point has a sensitivity and specificity of 88% [60]. In addition, there is no statistical significance between the PHQ-9 and the PHQ-8 [60,61].

Mental Component score of the SF-12

The Short-Form 12-item health survey is used to collect health-related quality of life data. This survey gives a mental component score (MCS) and a physical component score (PCS); the mental component score is of particular interest to this study. Higher scores on the MCS are related to higher quality of life and higher mental well-being. We used two different scores on the MCS to identify two mental health states. Individuals scoring below 46 on the MCS were labeled as experiencing depression, and those with scores under 37 were categorized as having severe psychological symptoms [72]. These different cut-points were determined through comparisons of the results of structured composite international diagnostic interview, version 2.1 [72]. The SF-12 is a valid and reliable measures in patients with low back pain and/or patients undergoing lumbar spine

surgery [39,67,69]. In addition, the SF-12 is a valid measure of mental well-being and quality of life in patients undergoing minimally invasive discectomy surgery [64].

3.2.2 Clinical Outcomes

Clinical outcome measures were collected at the preoperative baseline, and 3, 12, and 24-months post operatively. Back pain intensity was assessed using an 11 -point Numeric Pain Rating Scale (NPRS). Possible scores range from zero (i.e., 'no pain') to eleven (i.e., 'the worst imaginable pain). The NPRS is reliable in patients with lumbar disc herniations [ICC=0.991] [38]. It is also a valid measurement of pain for diverse pain origins [37,40]. The levels of Minimally Clinically Important Change (MCIC) in patients with lumbar disc herniations range from 2.0 to 4.0 to achieve significant improvement following discectomy surgery [34].

Disability was measured with the Oswestry Disability Index (ODI). The ODI is a widely used disability questionnaire for low back pain. It covers 10 items that assess functional status. The questionnaire is scored on a 100 point scale, with higher scores representing greater pain-related disability [22]. Scores are grouped together into categories: minimal disability (0-20), moderate disability (21-40), severe disability (41-60), crippled (61-80), and bedbound (81-100) [21]. The ODI is a valid and reliable measure of low back pain and disability in a clinical setting [R=0.75; ICC= 0.90] [28,29]. The MCIC is 16.9 points or a 32% change for patients following disc herniation surgery [34].

3.2.3 Potential Confounders and Additional Variables

We collected clinical and health data using standardized forms. All data is kept within the CSORN registry. Additional variables such as age, education, working status, height, sex and weight were collected at the initial assessment.

Potential confounders for the relationship between depression and post-operative pain are: age, sex, smoking status, education level, working status and baseline pain. Possible confounders for the relationship between depression and post-operative disability are: age, sex, smoking status, working status, education level and baseline disability.

3.2.4 Data Analysis

All statistical analyses were performed using STATA/SE versions 18 (StataCorp, College Station, TX). Doubly Robust estimates of the average treatment effect were constructed by combining exposure (propensity score) and outcome (regression) modelling in order to estimate the average treatment effects of depression and depression risk on the outcomes of leg pain, back pain and disability [74].

The study builds on previous work done by part of our research group that identified trajectory models for pain and disability for patients undergoing lumbar discectomy surgery as treatment for lumbar radiculopathy. The researchers created a trajectory groups based on leg pain, a characteristic symptom of lumbar radiculopathy and on overall surgical outcome that included leg pain, back pain, and disability levels.

For the purposes of this research, the 'excellent' and 'good' outcomes were combined and the results will be coded as a binary variable of poor outcome or not.

We first estimated covariate balancing propensity scores (CBPS) to help balance exposure groups and major confounders [75]. We then examined covariate balance by having the standardized differences for confounders be less than 0.1, and between 0.5 and 2.0 for the exposure groups and variance ratios [76,77]. We also visually inspected the distribution of propensity scores with overlap plots (Appendix III)[76]. The propensity scores were then converted to inverse probability of treatment weights to be used in later regression models.

We used weighted Poisson Regression models to estimate the average treatment effects. These results are reported as risk ratios and marginal risk estimates to show the probability of poor outcomes in each exposure group. Roust Poisson models are preferred with common binary outcomes as they yield less inflated and more precise estimates compared to logistic regressions [78,79]. To determine if the results were influenced by surgeon clustering, we conducted an analysis incorporating this as a variable and comparing the results with the original analysis. However, there was no evidence that clustering and no significant differences between the different models.

We applied two adjustment sets due to the unclear path between psychological variables and postoperative pain and disability [80]. Initially, we controlled for a certain number of confounding variables to minimize the risk of overadjustment bias [81]; age, sex, education, and smoking status (simple adjustment set). Next, we constructed models that

additionally control for the physical component of the SF-12 and surgical wait time (inclusive adjustment set).

3.2.5 Sensitivity Analyses

We then conducted sensitivity analyses to explore the potential bias due to unmeasured confounders as well as extreme inverse probability of treatment weights. E-value methods were used to estimate the potential for unmeasured confounders in all models [82]. E-values measure how much influence an unmeasured factor would require to completely justify or cancel the observed treatment effect, based on the measured variables. To assess how extreme weights affect variance inflation, we cut off the inverse probability weights at 5% and 95% on all models[83].

3.3 Results

We included data from 1104 patients (52% female sex) from the 22 spine centers. The mean (SD) age of all patients was 43.90 (13.13) years. Baseline demographic and health data are reported in Table 1.

Table 1. Demographic and preoperative data

Variable	Sample Size	Value
Age	1104	43.90 ± 13.13
Sex (female)	1104	582 (52.72%)
Education		
High school Diploma or less	1104	354 (32.07%)
Technical or Associates degree		202 (18.30%)
University/College Degree		548 (49.63%)
Smoking	1104	218 (19.79%)
Moderate to Severe Depression Risk (PHQ-8 ≥10)	1104	631 (57.16%)
SF-12 MCS Depression	1079	688 (63.76%)
SF-12 MCS Severe Psychological Symptoms	1079	358 (33.18%)

NPRS= Numeric Pain Rating Scale; PHQ-8 = Patient Health Questionnaire; SF-12 = Short Form Health Survey; MCS= Mental Component Score

[a]Values are reported as a value (percentage of patients) or mean ± SD

3.3.1 Discectomy Trajectory Results:

The previous trajectory research found that 11.4% of the patient population experienced a poor surgical outcome in regard to leg pain (Figure 1) and 28% of the population experienced a poor overall surgical outcome (Figure 2).

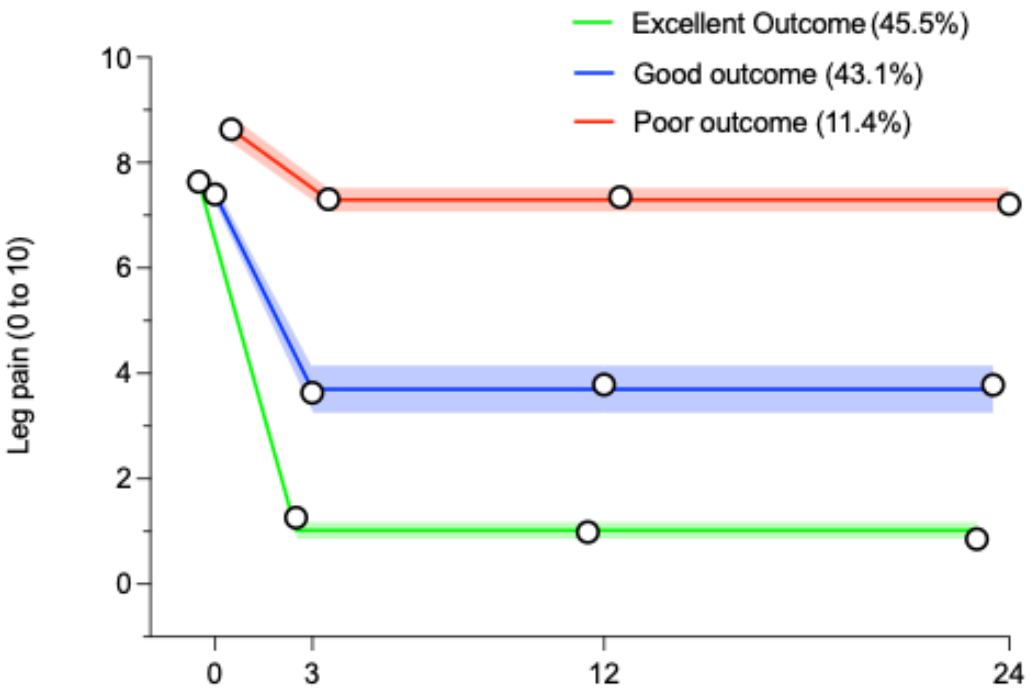

Figure 1. Trajectories of leg pain outcome following discectomy surgery.

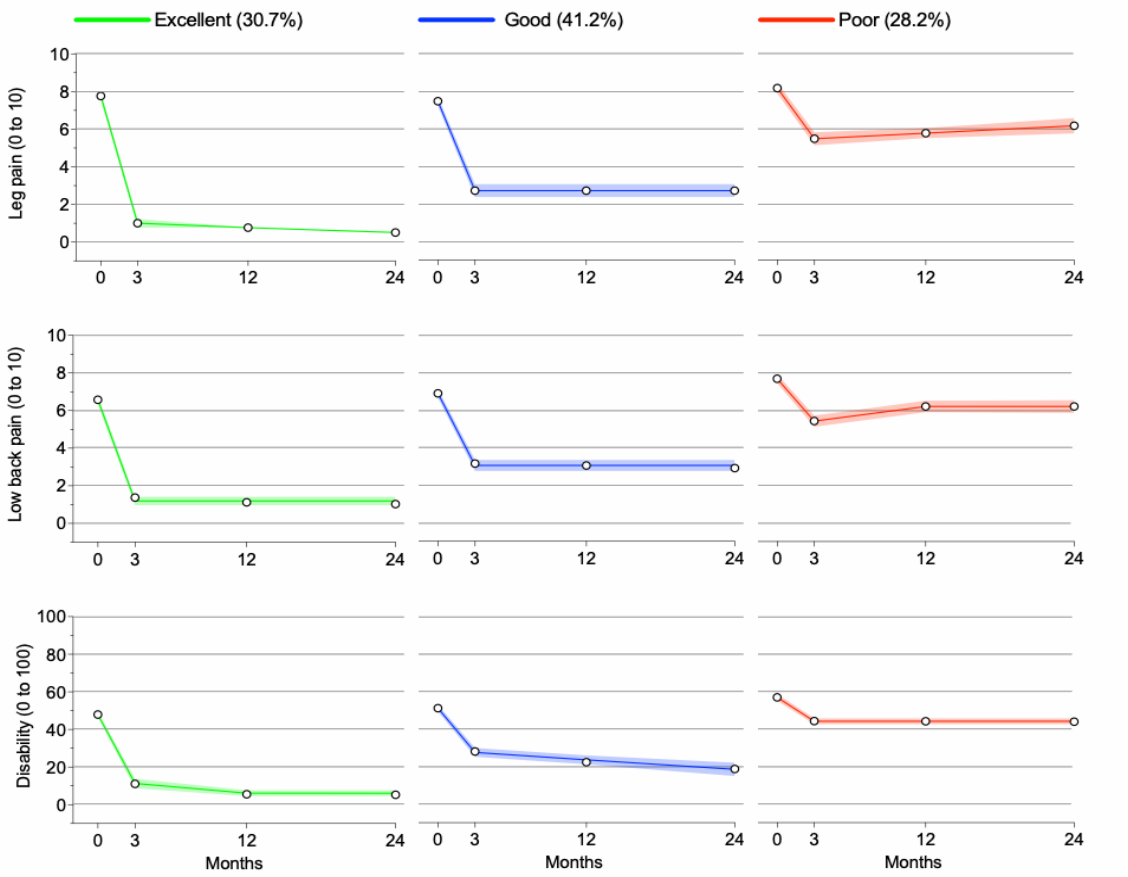

Figure 2. Trajectories of leg pain, back pain and disability outcomes following lumbar discectomy surgery.

3.3.2 Outcome Predictors

Depression Risk

Roughly 60% of patients were identified as being at a moderate-to-severe risk of depression (PHQ-8 ≥ 10). When accounting for common confounders (simple adjustment set), moderate-to-severe depression risk predicted nearly double the risk of poor leg pain (IRR [95% CI] = 1.70 [1.14 to 2.50]) and poor overall (IRR [95% CI] = 1.88 [1.48 to 2.40]) outcomes (Figure 3).

After accounting for further confounding variables (inclusive adjustment set), the effects of depression risk these outcomes were similar. Moderate-to-severe depression risk was associated with roughly a 1.5 times greater risk of poor leg pain (IRR [95% CI]= 1.61 [1.07 to 2.43]) and nearly double the risk of poor overall (IRR [95% CI] = 1.90 [1.46 to 2.4]) outcomes (Figure 3).

Mental Component Scores

Approximately 60% (57.16%) of patients were identified to have depression on the SF-12 MCS. Depression effects on leg pain outcomes were non-significant for the MCS for the simple model. However, it was associated with an increased risk of a poor overall outcome (IRR [95% CI] = 1.67 [1.29 to 2.16]) (Figure 3).

When adjusting for further confounders in the inclusive adjustment set, depression on the MCS was associated with poor leg and overall outcomes. Patients with depression had roughly double the risk of both poor leg and overall outcome respectively (IRR [95% CI] = 1.92 [1.20 to 3.08] ; 2.07 [1.51 to 2.86]) (Figure 3).

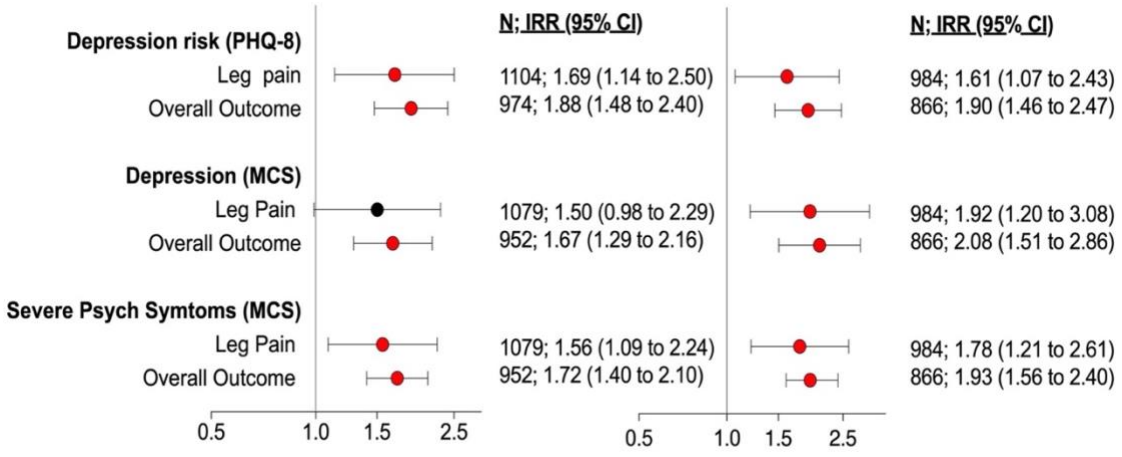

Figure 3. Effects of predictors of poor leg pain and overall surgical outcome for the simple (left) and inclusive (right) adjustment sets, reported as incidence rate ratios and 95% confidence intervals.

PHQ-8 = Patient-Health-Questionnaire-8; MCS = Mental Component Score [a]Red indicates a significant increased risk of poor outcome.

Approximately 30% of patients were identified to have severe psychological symptoms based on the SF-12 MCS. Severe symptomology resulted in roughly a 1.5 times greater risk of having poor leg pain (IRR [95% CI] = 1.55 [1.09 to 2.24]) and poor overall outcomes (IRR [95% CI] = 1.72 [1.40 to 2.10] (Figure 3).

When accounting for further confounders, severe psychological symptoms on the MCS was then associated with increased risk of poor leg and overall outcomes. severe psychological symptomology resulted in roughly 1.5 times greater risk of having a poor leg pain outcome (IRR [95% CI]= 1.78 [1.21 to 2.61]) and nearly double the risk of a poor overall outcome (IRR [95% CI]= 1.93 [1.55 to 2.40]) (Figure 3).

Sensitivity Analysis

Sensitivity analyses indicated that in the case of moderate-to-severe depressive risk, unmeasured confounders associated with both depressive risk and leg pain outcomes would need considerable strength to explain away the effect in both adjustment sets. Unmeasured confounders with an impact of 1.34 or greater could shift the confidence interval to include the null for both adjustment sets. For overall surgical outcome and moderate-to-severe depressive risk, unmeasured confounding would need to have a 3-fold impact on both the exposure and the outcome in both adjustment sets. Unmeasured confounders with an impact of 2.31 to 2.28 could shift the confidence interval to include the null.

For the MCS depression scores, unmeasured confounding would need roughly a 3-fold impact to explain away the results for leg pain outcomes in the inclusive model. An impact of 1.70 would be required to shift the confidence interval to include the null. In terms of overall outcomes, unmeasured confounding would need to effect both the exposure and outcome by 2.72 and 3.57 on the risk ratio scale to explain away the effects.

For severe psychological symptoms on the MCS, unmeasured confounding would need over double the impact for leg pain outcomes in order to explain away the results. Unmeasured confounding would also need a risk ratio over 1.5 to shift the confidence interval to include the null. For overall surgical outcome, unmeasured confounding would need roughly a 3-fold impact on both the exposure and the outcome in order to do away

with the relationship. As well as at least 2.15 on a risk ratio scale to shift the confidence interval to include the null (Table 2 & 3).

Table 2. E-values with their corresponding IRRs and 95% confidence intervals for the simple adjustment set.

Model	Value	
	IRR (95% CI)	E-value Point Estimate (Lower-Bound Estimate)
Moderate-to-Severe Depression Risk – Leg Pain	1.69 (1.39 to 2.50)	2.77 (1.34)
Moderate-to-Severe Depression Risk – Overall Outcome	1.88 (1.48 to 2.40)	3.17 (2.31)
Mental Component Score Depression – Leg pain	1.50 (0.99 to 2.29)	-
Mental Component Score Depression – Overall Outcome	1.67 (1.29 to 2.16)	2.72 (1.89)
Mental Component Score Severe Psychological Symptoms – Leg Pain	1.56 (1.09 to 2.24)	2.49 (1.39)
Mental Component Score Severe Psychological Symptoms – Overall Outcome	1.72 (1.40 to 2.10)	2.83 (2.15)

Table 3. E-values with their corresponding IRRs and 95% confidence intervals for the inclusive adjustment set.

Model	**Value**	
	IRR (95% CI)	E-value Point Estimate (Lower-Bound Estimate)
Moderate-to-Severe Depression Risk – Leg Pain	1.61 (1.07 to 2.43)	2.60 (1.34)
Moderate-to-Severe Depression Risk – Overall Outcome	1.90 (1.46 to 2.47)	3.20 (2.28)
Mental Component Score Depression – Leg pain	1.92 (1.20 to 3.07)	3.26 (1.70)
Mental Component Score Depression – Overall Outcome	2.08 (1.51 to 2.86)	3.57 (2.38)
Mental Component Score Severe Psychological Symptoms – Leg Pain	1.78 (1.20 to 2.61)	2.95 (1.71)
Mental Component Score Severe Psychological Symptoms – Overall Outcome	1.93 (1.55 to 2.40)	3.27 (2.47)

When conducting an analysis truncating the weights at 5% and 95% , the results remained essentially unchanged. Only minor changes were seen between the two different analyses, with either no effect or very small increases in the risk ratios (Table 4 & 5). The consistency between the main analysis and sensitivity analysis indicates that our models were not heavily influenced by the extreme probabilities.

Table 4. IRRs and confidence intervals for the simple adjustment set and truncated weight analysis.

Model	Truncated Weight Analysis IRR (95% CI)
Moderate-to-Severe Depression Risk – Leg Pain	1.70 (1.15, 2.52)
Moderate-to-Severe Depression Risk – Overall Outcome	1.90 (1.48, 2.40)
Mental Component Score Depression – Leg pain	1.56 (1.02, 2.37)
Mental Component Score Depression – Overall Outcome	1.70 (1.30, 2.20)
Mental Component Score Severe Psychological Symptoms – Leg Pain	1.56 (1.08, 2.23)
Mental Component Score Severe Psychological Symptoms – Overall Outcome	1.72 (1.40, 2.10)

Table 5. IRRs and confidence intervals for the inclusive adjustment set and truncated weight analysis.

Model	Truncated Weight Analysis IRR (95% CI)
Moderate-to-Severe Depression Risk – Leg Pain	1.58 (1.02, 2.45)
Moderate-to-Severe Depression Risk – Overall Outcome	1.93 (1.48, 2.05)
Mental Component Score Depression – Leg pain	1.98 (1.25, 3.12)
Mental Component Score Depression – Overall Outcome	2.18 (1.62, 2.93)
Mental Component Score Severe Psychological Symptoms – Leg Pain	1.76 (1.20, 2.58)
Mental Component Score Severe Psychological Symptoms – Overall Outcome	1.95 (1.57, 2.43)

3.4 Discussion

This study aimed to estimate the effects of preoperative psychological factors, specifically depression, on postoperative pain and disability measures in patients undergoing lumbar discectomy surgery as treatment for lumbar radiculopathy. We found that depression did increase the risk of poor outcomes following discectomy surgery. When adjusting for a more inclusive set of confounders, moderate-to-severe risk of depression significantly increased the risk of poor leg pain and poor overall surgical outcomes. In addition, both depression and severe psychological symptoms on the MCS were associated with an increased risk of poor leg pain and poor overall surgical outcomes.

Sensitivity analyses allowed us to estimate the potential impact of unmeasured confounding [82], and showed that substantial unmeasured confounding would be needed to completely explain away our results. This helps to provide us with increased confidence in our estimates of the effects of depression on postoperative pain and disability.

Previous studies have reported similar results regarding mental health status and postoperative pain and disability in lumbar spine surgery. To our knowledge, this is the first study that looked specifically at how depression impacts both pain and depression following lumbar discectomy surgery. While the relationship between depression, pain and disability is complex with no clear causal path, our findings are consistent with the results of previous studies. Previous research has indicated that depression is associated with increased postoperative pain following lumbar surgeries [11,13,45,52]. In addition, lumbar surgery and depression has been correlated with poorer surgical outcomes, including disability outcomes [13,14,46,84]. These previous findings are consistent with our results, and indicates that individuals with a risk of depression and lower mental well-being had less improvement in disability and pain scales following surgery. A plausible explanation of these results could be that the catastrophizing thoughts associated with depression may put these individuals at risk for lower improvements. In addition, patients who have lower mental health also tend to have lower quality of life and more pain-related disability scores [85].

The study's results are relevant and applicable to both clinical and research settings. Clinically, the findings could help surgeons to provide better care for surgical

patients with depression. Understanding the relationship between depression and postoperative outcomes allows for better risk assessment in surgical patients. Recognizing the impact of depression on the postoperative outcomes can allow surgeons and health care professional to enhance patient outcomes. Furthermore, We found the prevalence of depression in our surgical population to be 57% and 64% on the PHQ-8 and the MCS respectively. However, in the general population the prevalence of depression on these screening measures in 6% and 5% [86,87]. A nearly a ten-fold difference between surgical and non-surgical patients underscores the necessity for preoperative psychological screening and interventions. Addressing depression in lumbar discectomy patients could potentially improve overall surgical outcomes, enhance recovery and overall patient well-being. For example, managing depression alongside regular postoperative care could lead to improved patient outcomes, including faster recovery, reduced disability and pain, and better quality of life postoperatively. Research in this area could be done to create targeted interventions. Future intervention research could focus on specific interventions targeted to mitigate the impact of depression on postoperative pain and disability for surgical patients. Moreover, conducting research on the effects of preoperative antidepressant usage on postoperative outcomes could offer substantial contributions to this field.

Strength and Limitations

Strengths of this study primarily lie in its sample size. The ability to access the CSORN database allowed us to have a large sample size for statistical analysis. The large

sample size ensured that our analysis had increased precision and statistical power so that we could detect true associations in our results.

In terms of limitations, we were unable to use a clinical diagnosis of depression or psychological health. The psychological health measures used do not provide a clinical diagnosis of the mental condition and only provide an estimate of the risk or presence of the symptoms. However, the psychological surveys used are highly valid when compared to the diagnostic interviews associated with the DSM-IV. Since clinical diagnosis is a lengthy and time consuming process, the psychological measures provide a more time effective and valid measure of mental health.

In addition, the ideal study design for our study would have been a randomized control trial. However, our patient population is real-world patients and surgeons and we are unable to manipulate any care decisions for the sake of randomization. Using propensity weighting in our analysis did provide the most accurate representation of a trial and allowed for causal analysis between depression and postoperative outcomes.

3.5 Conclusion

The results in the current study indicate that poor mental status, including depressive risk and severe psychological symptoms, are predictive of poor pain and disability outcomes. After all confounding adjustments, psychological factors had significant effects on both pain and disability outcomes postoperatively. These finding may be applicable in clinical settings by indicating the importance of considering

depression pre-operatively and potentially screening for depressive risk and evaluating the need for psychological treatment prior to surgery in order to maximize the likelihood improvement.

Future research should focus on treatments and interventions to mitigate the effects of depression prior to surgery and its subsequent effects on surgical outcomes. Investigation on the use of anti-depressant medicate may also be of interest. Finally, it may be beneficial to conduct research that uses randomization tools to modify depression levels to determine potential causal links.

References

1. Ferreira M, de Luca K, Haile L, et al. Global, regional, and national burden of low back pain, 1990–2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. 2023;5:e316-e329.

2. Berry JA, Elia C, Saini HS, Miulli DE. A Review of Lumbar Radiculopathy, Diagnosis, and Treatment. *Cureus*. 2019;11(10):e5934. doi:10.7759/cureus.5934

3. Tarulli AW, Raynor EM. Lumbosacral Radiculopathy. *Neurologic Clinics*. 2007;25(2):387-405. doi:10.1016/j.ncl.2007.01.008

4. Valat JP, Genevay S, Marty M, Rozenberg S, Koes B. Sciatica. *Best Practice & Research Clinical Rheumatology*. 2010;24(2):241-252. doi:10.1016/j.berh.2009.11.005

5. Hasvik E, Haugen AJ, Grøvle L. Symptom descriptors and patterns in lumbar radicular pain caused by disc herniation: a 1-year longitudinal cohort study. *BMJ Open*. 2022;12(12):e065500. doi:10.1136/bmjopen-2022-065500

6. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ*. 2007;334(7607):1313-1317. doi:10.1136/bmj.39223.428495.BE

7. Blamoutier A. Surgical discectomy for lumbar disc herniation: Surgical techniques. *Orthopaedics & Traumatology: Surgery & Research*. 2013;99(1, Supplement):S187-S196. doi:10.1016/j.otsr.2012.11.005

8. Jacobs WCH, van Tulder M, Arts M, et al. Surgery versus conservative management of sciatica due to a lumbar herniated disc: a systematic review. *Eur Spine J*. 2011;20(4):513-522. doi:10.1007/s00586-010-1603-7

9. Schoenfeld AJ, Weiner BK. Treatment of lumbar disc herniation: Evidence-based practice. *Int J Gen Med*. 2010;3:209-214.

10. Xu Y, Yen D, Whitehead M, Xu J, Johnson AP. Use of instrumented lumbar spinal surgery for degenerative conditions: trends and costs over time in Ontario, Canada. *Can J Surg*. 2019;62(6):393-401. doi:10.1503/cjs.017016

11. Adogwa O, Parker SL, Shau DN, et al. Preoperative Zung Depression Scale predicts outcome after revision lumbar surgery for adjacent segment disease, recurrent stenosis, and pseudarthrosis. *The Spine Journal*. 2012;12(3):179-185. doi:10.1016/j.spinee.2011.08.014

12. Wilhelm M, Reiman M, Goode A, et al. Psychological Predictors of Outcomes with Lumbar Spinal Fusion: A Systematic Literature Review. *Physiotherapy Research International*. 2017;22(2):e1648. doi:10.1002/pri.1648

13. Tuomainen I, Pakarinen M, Aalto T, et al. Depression is associated with the long-term outcome of lumbar spinal stenosis surgery: a 10-year follow-up study. *The Spine Journal*. 2018;18(3):458-463. doi:10.1016/j.spinee.2017.08.228

14. Aalto TJ, Malmivaara A, Kovacs F, et al. Preoperative Predictors for Postoperative Clinical Outcome in Lumbar Spinal Stenosis: Systematic Review. *Spine*. 2006;31(18):E648. doi:10.1097/01.brs.0000231727.88477.da

15. Lurie JD, Tosteson TD, Tosteson ANA, et al. Surgical versus Non-Operative Treatment for Lumbar Disc Herniation: Eight-Year Results for the Spine Patient

Outcomes Research Trial (SPORT). *Spine (Phila Pa 1976)*. 2014;39(1):3-16. doi:10.1097/BRS.0000000000000088

16. Andersson GB. Epidemiological features of chronic low-back pain. *The Lancet*. 1999;354(9178):581-585. doi:10.1016/S0140-6736(99)01312-4

17. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *JBJS*. 1990;72(3):403.

18. Jensen RK, Harhangi BS, Huygen F, Koes B. Lumbar spinal stenosis. *BMJ*. 2021;373:n1581. doi:10.1136/bmj.n1581

19. Groen GJ, Baljet B, Drukker J. Nerves and nerve plexuses of the human vertebral column. *Am J Anat*. 1990;188(3):282-296. doi:10.1002/aja.1001880307

20. Cherkin DC, Deyo RA, Loeser JD, Bush T, Waddell G. An international comparison of back surgery rates. *Spine (Phila Pa 1976)*. 1994;19(11):1201-1206. doi:10.1097/00007632-199405310-00001

21. Fairbank JCT, Pynsent PB. The Oswestry Disability Index. *Spine*. 2000;25(22):2940.

22. Smeets R, Köke A, Lin CW, Ferreira M, Demoulin C. Measures of function in low back pain/disorders: Low Back Pain Rating Scale (LBPRS), Oswestry Disability Index (ODI), Progressive Isoinertial Lifting Evaluation (PILE), Quebec Back Pain Disability Scale (QBPDS), and Roland-Morris Disability Questionnaire (RDQ). *Arthritis Care & Research*. 2011;63(S11):S158-S173. doi:10.1002/acr.20542

23. Vianin M. Psychometric properties and clinical usefulness of the Oswestry Disability Index. *J Chiropr Med*. 2008;7(4):161-163. doi:10.1016/j.jcm.2008.07.001

24. Cook CE, Garcia AN, Wright A, Shaffrey C, Gottfried O. Measurement Properties of the Oswestry Disability Index in Recipients of Lumbar Spine Surgery. *Spine*. 2021;46(2):E118. doi:10.1097/BRS.0000000000003732

25. Monticone M, Baiardi P, Vanti C, et al. Responsiveness of the Oswestry Disability Index and the Roland Morris Disability Questionnaire in Italian subjects with sub-acute and chronic low back pain. *Eur Spine J*. 2012;21(1):122-129. doi:10.1007/s00586-011-1959-3

26. Cleland J, Gillani R, Bienen EJ, Sadosky A. Assessing Dimensionality and Responsiveness of Outcomes Measures for Patients with Low Back Pain. *Pain Practice*. 2011;11(1):57-69. doi:10.1111/j.1533-2500.2010.00390.x

27. Johnsen LG, Hellum C, Nygaard ØP, et al. Comparison of the SF6D, the EQ5D, and the oswestry disability index in patients with chronic low back pain and degenerative disc disease. *BMC Musculoskelet Disord*. 2013;14:148. doi:10.1186/1471-2474-14-148

28. Shafshak TS, Elnemr R. The Visual Analogue Scale Versus Numerical Rating Scale in Measuring Pain Severity and Predicting Disability in Low Back Pain. *JCR: Journal of Clinical Rheumatology*. 2021;27(7):282. doi:10.1097/RHU.0000000000001320

29. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther*. 2001;81(2):776-788. doi:10.1093/ptj/81.2.776

30. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting Change Scores for Pain and Functional Status in Low Back Pain: Towards International Consensus Regarding Minimal Important Change. *Spine*. 2008;33(1):90. doi:10.1097/BRS.0b013e31815e3a10

31. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and Pain Scales. *The Spine Journal*. 2008;8(6):968-974. doi:10.1016/j.spinee.2007.11.006

32. van Hooff ML, Mannion AF, Staub LP, Ostelo RWJG, Fairbank JCT. Determination of the Oswestry Disability Index score equivalent to a "satisfactory symptom state" in patients undergoing surgery for degenerative disorders of the lumbar spine—a Spine Tango registry-based study. *The Spine Journal*. 2016;16(10):1221-1230. doi:10.1016/j.spinee.2016.06.010

33. Solberg T, Johnsen LG, Nygaard ØP, Grotle M. Can we define success criteria for lumbar disc surgery? : estimates for a substantial amount of improvement in core outcome measures. *Acta Orthop*. 2013;84(2):196-201. doi:10.3109/17453674.2013.786634

34. Power JD, Perruccio AV, Canizares M, et al. Determining minimal clinically important difference estimates following surgery for degenerative conditions of the lumbar spine: analysis of the Canadian Spine Outcomes and Research Network (CSORN) registry. *The Spine Journal*. 2023;23(9):1323-1333. doi:10.1016/j.spinee.2023.05.001

35. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies Comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for Assessment of Pain Intensity in Adults: A Systematic Literature Review. *Journal of Pain and Symptom Management*. 2011;41(6):1073-1093. doi:10.1016/j.jpainsymman.2010.08.016

36. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs*. 2005;14(7):798-804. doi:10.1111/j.1365-2702.2005.01121.x

37. Ware LJ, Epps CD, Herr K, Packard A. Evaluation of the Revised Faces Pain Scale, Verbal Descriptor Scale, Numeric Rating Scale, and Iowa Pain Thermometer in Older Minority Adults. *Pain Management Nursing*. 2006;7(3):117-125. doi:10.1016/j.pmn.2006.06.005

38. Yao M, Xu B ping, Li Z jun, et al. A comparison between the low back pain scales for patients with lumbar disc herniation: validity, reliability, and responsiveness. *Health & Quality of Life Outcomes*. 2020;18(1):1-12. doi:10.1186/s12955-020-01403-2

39. Boden SH, Farley KX, Campbell C, Boden SD, Gottschalk MB. Rational Selection of Patient-Reported Outcomes Measures in Lumbar Spine Surgery Patients. *International Journal of Spine Surgery*. 2020;14(3):347-354. doi:10.14444/7046

40. Karcioglu O, Topacoglu H, Dikme O, Dikme O. A systematic review of the pain scales in adults: Which to use? *The American Journal of Emergency Medicine*. 2018;36(4):707-714. doi:10.1016/j.ajem.2018.01.008

41. Jensen MP, Castarlenas E, Roy R, et al. The Utility and Construct Validity of Four Measures of Pain Intensity: Results from a University-Based Study in Spain. *Pain Medicine*. 2019;20(12):2411-2420. doi:10.1093/pm/pny319

42. Childs JD, Piva SR, Fritz JM. Responsiveness of the Numeric Pain Rating Scale in Patients with Low Back Pain. *Spine*. 2005;30(11):1331. doi:10.1097/01.brs.0000164099.92112.29

43. Wilson CA, Roffey DM, Chow D, Alkherayf F, Wai EK. A systematic review of preoperative predictors for postoperative clinical outcomes following lumbar discectomy. *The Spine Journal*. 2016;16(11):1413-1422. doi:10.1016/j.spinee.2016.08.003

44. Junge A, Dvorak J, Ahrens S. Predictors of bad and good outcomes of lumbar disc surgery. A prospective clinical study with recommendations for screening to avoid bad outcomes. *Spine (Phila Pa 1976)*. 1995;20(4):460-468. doi:10.1097/00007632-199502001-00009

45. Arpino L, Iavarone A, Parlato C, Moraci A. Prognostic role of depression after lumbar disc surgery. *Neurol Sci*. 2004;25(3):145-147. doi:10.1007/s10072-004-0248-x

46. Chaichana KL, Mukherjee D, Adogwa O, Cheng JS, McGirt MJ. Correlation of preoperative depression and somatic perception scales with postoperative disability and

quality of life after lumbar discectomy: Clinical article. *Journal of Neurosurgery: Spine*. 2011;14(2):261-267. doi:10.3171/2010.10.SPINE10190

47. Trief PM, Grant W, Fredrickson B. A Prospective Study of Psychological Predictors of Lumbar Surgery Outcome. *Spine*. 2000;25(20):2616.

48. Hébert JJ, Abraham E, Wedderkopp N, et al. Preoperative Factors Predict Postoperative Trajectories of Pain and Disability Following Surgery for Degenerative Lumbar Spinal Stenosis. *Spine*. 2020;45(21):E1421. doi:10.1097/BRS.0000000000003587

49. Sinikallio S, Aalto T, Airaksinen O, et al. Depression is associated with poorer outcome of lumbar spinal stenosis surgery. *Eur Spine J*. 2007;16(7):905-912. doi:10.1007/s00586-007-0349-3

50. Chapin L, Ward K, Ryken T. Preoperative Depression, Smoking, and Employment Status are Significant Factors in Patient Satisfaction After Lumbar Spine Surgery. *Clin Spine Surg*. 2017;30(6):E725-E732. doi:10.1097/BSD.0000000000000331

51. Amaral V, Marchi L, Martim H, et al. Influence of psychosocial distress in the results of elective lumbar spine surgery. *J Spine Surg*. 2017;3(3):371-378. doi:10.21037/jss.2017.08.05

52. Vilà-Canet G, Covaro A, Isart A, et al. Elective Lumbar Spine Surgery in Depressed Patients: Is it Worth it? *Int J Spine Surg*. 2021;15(3):418-422. doi:10.14444/8062

53. Miller JA, Derakhshan A, Lubelski D, et al. The impact of preoperative depression on quality of life outcomes after lumbar surgery. *Spine J*. 2015;15(1):58-64. doi:10.1016/j.spinee.2014.06.020

54. Järvimäki V, Kautiainen H, Haanpää M, et al. Depressive symptoms are associated with poor outcome for lumbar spine surgery. *Scandinavian Journal of Pain*. 2016;12(1):13-17. doi:10.1016/j.sjpain.2016.01.008

55. Hartman TJ, Nie JW, Zheng E, Oyetayo OO, MacGregor KR, Singh K. Poor mental health scores correlate with inferior outcomes following minimally invasive transforaminal lumbar interbody fusion. *Acta Neurochir*. 2023;165(7):1931-1942. doi:10.1007/s00701-023-05557-y

56. Anderson JT, Haas AR, Percy R, Woods ST, Ahn UM, Ahn NU. Clinical depression is a strong predictor of poor lumbar fusion outcomes among workers' compensation subjects. *Spine (Phila Pa 1976)*. 2015;40(10):748-756. doi:10.1097/BRS.0000000000000863

57. Mancuso CA, Duculan R, Craig CM, Girardi FP. Psychosocial Variables Contribute to Length of Stay and Discharge Destination After Lumbar Surgery Independent of Demographic and Clinical Variables. *Spine*. 2018;43(4):281. doi:10.1097/BRS.0000000000002312

58. Folman Y, Shabat S, Catz A, Gepstein R. Late results of surgery for herniated lumbar disk as related to duration of preoperative symptoms and type of herniation. *Surgical Neurology*. 2008;70(4):398-401. doi:10.1016/j.surneu.2007.04.022

59. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord*. 2009;114(1-3):163-173. doi:10.1016/j.jad.2008.06.026

60. Kroenke K, Spitzer RL, W WJB. The PHQ-9. *Journal of General Internal Medicine*. 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x

61. Shin C, Lee SH, Han KM, Yoon HK, Han C. Comparison of the Usefulness of the PHQ-8 and PHQ-9 for Screening for Major Depressive Disorder: Analysis of Psychiatric Outpatient Data. *Psychiatry Investig*. 2019;16(4):300-305. doi:10.30773/pi.2019.02.01

62. Costantini L, Pasquarella C, Odone A, et al. Screening for depression in primary care with Patient Health Questionnaire-9 (PHQ-9): A systematic review. *Journal of Affective Disorders*. 2021;279:473-483. doi:10.1016/j.jad.2020.09.131

63. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *General Hospital Psychiatry*. 2006;28(1):71-77. doi:10.1016/j.genhosppsych.2005.07.003

64. Lynch CP, Cha EDK, Jenkins NW, et al. Patient Health Questionnaire-9 Is a Valid Assessment for Depression in Minimally Invasive Lumbar Discectomy. *Neurospine*. 2021;18(2):369-376. doi:10.14245/ns.2142162.031

65. McDowell I. General Health Status and Quality of Life. In: *Measuring Health: A Guide to Rating Scales and Questionnaires*. ; 2006:520-703. https://doi.org/10.1093/acprof:oso/9780195165678.003.0010

66. Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: Scoping review. *SAGE Open Med*. 2016;4:2050312116671725. doi:10.1177/2050312116671725

67. Jenkinson C, Layte R, Jenkinson D, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *Journal of Public Health*. 1997;19(2):179-186. doi:10.1093/oxfordjournals.pubmed.a024606

68. Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care*. 1996;34(3):220.

69. Omari OA, Alkhawaldeh A, ALBashtawy M, Qaddumi J, Holm MB, AlOmari D. A Review of the Short Form Health Survey–Version 2. *Journal of Nursing Measurement*. 2019;27(1):77-86. doi:10.1891/1061-3749.27.1.77

70. LoMartire R, Äng BO, Gerdle B, Vixner L. Psychometric properties of Short Form-36 Health Survey, EuroQol 5-dimensions, and Hospital Anxiety and Depression Scale in patients with chronic pain. *Pain*. 2020;161(1):83-95. doi:10.1097/j.pain.0000000000001700

71. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160-164.

72. Gill SC, Butterworth P, Rodgers B, Mackinnon A. Validity of the mental health component scale of the 12-item Short-Form Health Survey (MCS-12) as measure of

common mental disorders in the general population. *Psychiatry Research*. 2007;152(1):63-71. doi:10.1016/j.psychres.2006.11.005

73. Purvis TE, Neuman BJ, Riley LH, Skolasky RL. Comparison of PROMIS Anxiety and Depression, PHQ-8, and GAD-7 to screen for anxiety and depression among patients presenting for spine surgery. *J Neurosurg Spine*. Published online January 18, 2019:1-8. doi:10.3171/2018.9.SPINE18521

74. Kang JDY, Schafer JL. Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. *Statistical Science*. 2007;22(4):523-539.

75. Imai K, Ratkovic M. Covariate Balancing Propensity Score. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 2014;76(1):243-263. doi:10.1111/rssb.12027

76. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*. 2009;28(25):3083-3107. doi:10.1002/sim.3697

77. Rubin DB. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology*. 2001;2(3-4):169-188.

78. Greenland S. Interpretation and Choice Effect Measures in Epidemiologic Analyses. *American Journal of Epidemiology*. 1987;125(5):761-768. doi:10.1093/oxfordjournals.aje.a114593

79. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Medical Research Methodology*. 2018;18(1):63. doi:10.1186/s12874-018-0519-5

80. Yao C, Zhang Y, Lu P, et al. Exploring the bidirectional relationship between pain and mental disorders: a comprehensive Mendelian randomization study. *The Journal of Headache and Pain*. 2023;24(1):82. doi:10.1186/s10194-023-01612-2

81. Schisterman EF, Cole SR, Platt RW. Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies. *Epidemiology*. 2009;20(4):488-495.

82. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/M16-2607

83. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656-664. doi:10.1093/aje/kwn164

84. Hébert JJ, Abraham E, Wedderkopp N, et al. Preoperative Factors Predict Postoperative Trajectories of Pain and Disability Following Surgery for Degenerative Lumbar Spinal Stenosis. *Spine*. 2020;45(21):E1421. doi:10.1097/BRS.0000000000003587

85. Börsbo B, Peolsson M, Gerdle B. The complex interplay between pain intensity, depression, anxiety and catastrophising with respect to quality of life and disability. *Disability and Rehabilitation*. 2009;31(19):1605-1613. doi:10.1080/09638280903110079

86. Vilagut G, Forero CG, Pinto-Meza A, et al. The Mental Component of the Short-Form 12 Health Survey (SF-12) as a Measure of Depressive Disorders in the General Population: Results with Three Alternative Scoring Methods. *Value in Health*. 2013;16(4):564-573. doi:10.1016/j.jval.2013.01.006

87. Arias-de La Torre J, Vilagut G, Ronaldson A, et al. Prevalence and variability of current depressive disorder in 27 European countries: a population-based study. *The Lancet Public Health*. 2021;6(10):e729-e738. doi:10.1016/S2468-2667(21)00047-5

Appendix I – Diagnostic Results

Simple Adjustment Set:

Table 6: Standard Difference Diagnostic Results for the PHQ-8 and Leg Outcomes Simple Model

	Raw			Matched (ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4659271	0.6088795	-0.2894331	0.5251318	0.5251561	-0.0000493
Age	43.17433	44.87949	-0.1298619	43.97587	43.97617	-0.0000227
Education (1)	0.2012678	0.1585624	0.1112489	0.1823014	0.1822939	0.0000195
Education (2)	0.3835182	0.3805497	0.0061039	0.3830549	0.3830546	6.51e-07
Education (3)	0.0744849	0.167019	-0.286616	0.1118861	0.1119017	-0.0000483
Smoking Status	0.2123613	0.1775899	0.0877706	0.1986733	0.1986676	0.0000144

Table 7: Variance Ratio Diagnostic Results for the PHQ-8 and Leg Outcomes Simple Model

	Raw			Matched (ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.249234	0.2386498	1.04435	0.2497642	0.2498955	0.9994747
Age	167.8299	176.996	0.9482126	176.1283	174.6812	1.008284
Education (1)	0.1610143	0.133703	1.204268	0.1493042	0.1493786	0.9995017
Education (2)	0.2368073	0.2362311	1.002439	0.236699	0.2368245	0.9994701
Education (3)	0.0690464	0.1394184	0.4952456	0.0995253	0.0995903	0.999348
Smoking Status	0.1675295	0.1463611	1.144631	0.1594549	0.1595361	0.9994913

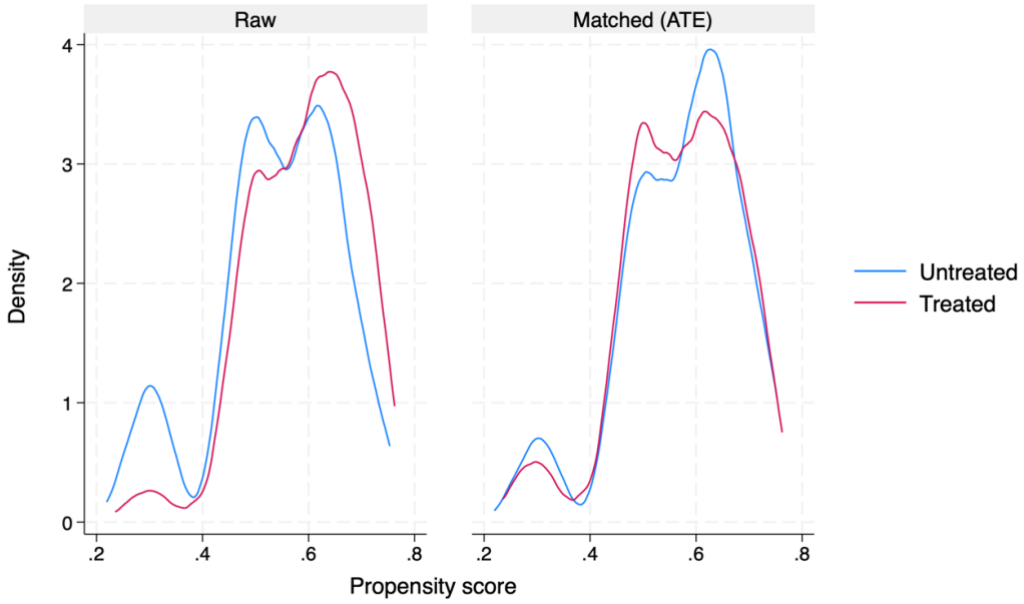

Figure 4: Overlap Plot Diagnostic Graphs for the PHQ-8 and Leg Outcomes Simple Model.

Table 8: Standard Difference Diagnostic Results for the PHQ-8 and Overall Outcomes Simple Model.

	Raw			Matched (ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4591837	0.6062176	-0.297633	0.5172378	0.5172595	-0.0000438
Age	43.20748	45.29275	-0.160805	44.1653	44.16563	-0.000025
Education (1)	0.2006803	0.1735751	0.0694661	0.1891333	0.1891293	0.0000104
Education (2)	0.372449	0.3756477	-0.006604	0.3745357	0.3745362	-1.11E-06
Education (3)	0.0731293	0.1554404	-0.260606	0.1042116	0.1042235	-0.0000377
Smoking Status	0.2210884	0.1917098	0.0725606	0.2123344	0.2123305	9.66E-06

Table 9: Variance Ratio Diagnostic Results for the PHQ-8 and Overall Outcomes Simple Model.

	Raw			Matched (ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.249234	0.2386498	1.04435	0.2497642	0.2498955	0.9994747
Age	167.8299	176.996	0.9482126	176.1283	174.6812	1.008284
Education (1)	0.1610143	0.133703	1.204268	0.1493042	0.1493786	0.9995017
Education (2)	0.2368073	0.2362311	1.002439	0.236699	0.2368245	0.9994701
Education (3)	0.0690464	0.1394184	0.4952456	0.0995253	0.0995903	0.999348
Smoking Status	0.1675295	0.1463611	1.144631	0.1594549	0.1595361	0.9994913

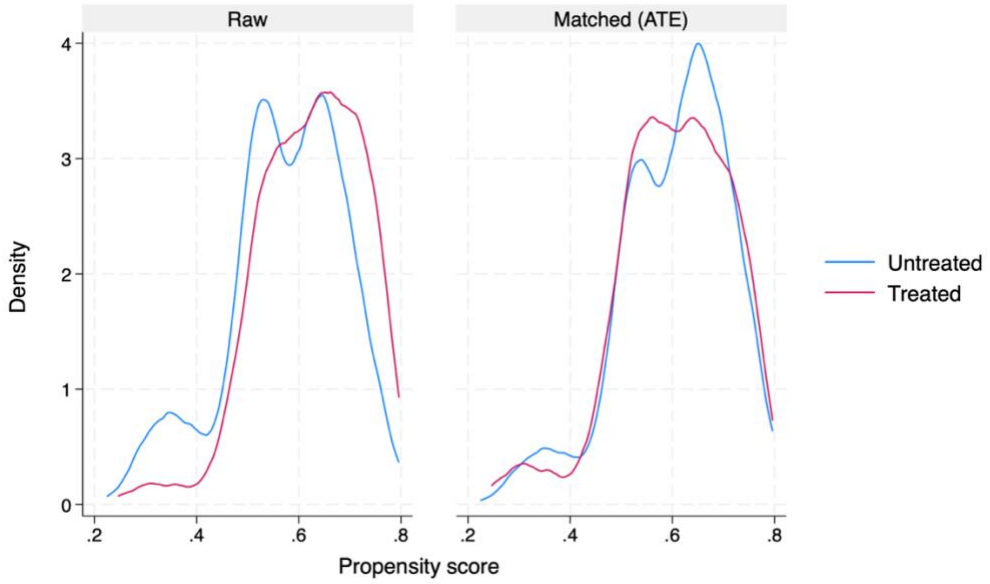

Figure 5: Overlap Plot Diagnostic Graphs for the PHQ-8 and Overall Outcomes Model.

Table 10: Standard Difference Diagnostic Results for the MCS Depression and Leg Outcomes Simple Model.

	Raw			Matched (ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4622093	0.6445013	-0.37263	0.5284398	0.5284653	-0.0000521
Age	43.86483	43.78261	0.0062933	43.86602	43.86601	3.00E-07
Education (1)	0.181686	0.1918159	-0.025969	0.1840534	0.1840545	-2.88E-06
Education (2)	0.3851744	0.3810742	0.0084258	0.3883786	0.3883787	-3.19E-07
Education (3)	0.0944767	0.1457801	-0.158129	0.1112195	0.1112264	-0.0000214
Smoking Status	0.2005814	0.1841432	0.0416723	0.1914675	0.1914647	7.16E-06

Table 11: Variance Ratio Diagnostic Results for the MCS Depression and Leg Outcomes Simple Model.

	Raw			Matched (ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2487571	0.2393379	1.039355	0.2501282	0.2503507	0.9991115
Age	167.1323	169.1894	0.9878417	177.2067	163.4305	1.084294
Education (1)	0.160681	0.1438194	1.117241	0.1536232	0.1537577	0.9991249
Education (2)	0.2341289	0.2351457	0.995676	0.2346578	0.2348673	0.9991079
Education (3)	0.0678968	0.1316197	0.5158563	0.0935106	0.0936035	0.9990076
Smoking Status	0.1725017	0.1553597	1.110338	0.1675334	0.1676807	0.9991219

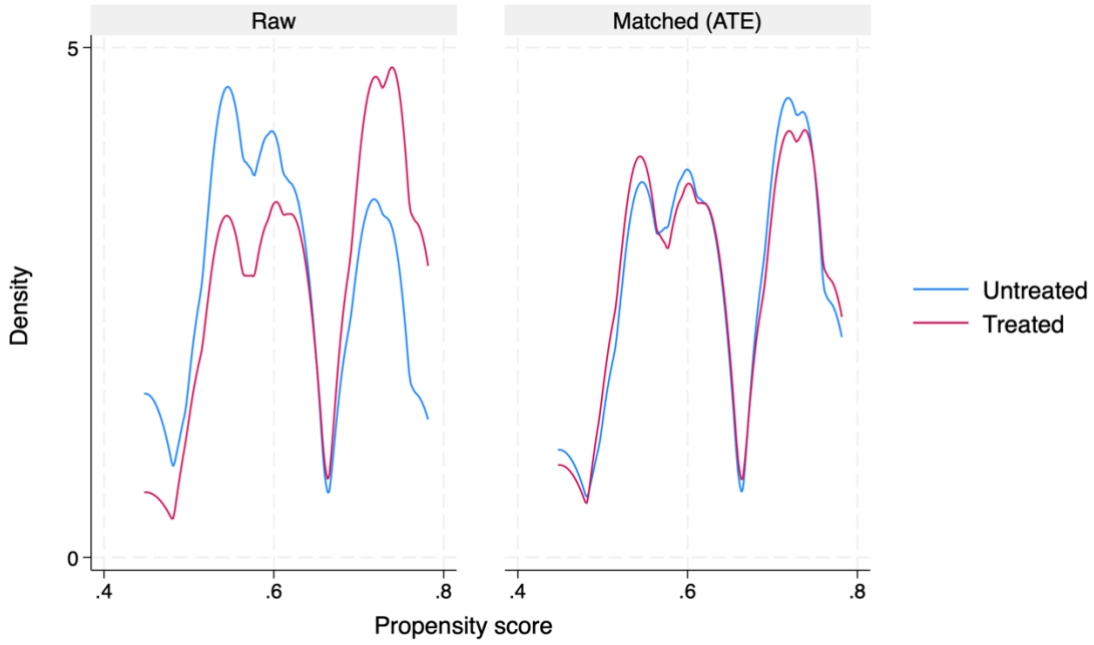

Figure 6: Overlap Plot Diagnostic Graphs for the MCS Depression and Leg Outcomes Simple Model.

Table 12: Standard Difference Diagnostic Results for the MCS Depression and Overall Outcomes Simple Model.

	Raw			Matched (ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4532488	0.6479751	-0.3987213	0.519559	0.5196104	-0.0001053
Age	43.96197	43.90654	0.0043131	43.97004	43.97003	3.32E-07
Education (1)	0.1838352	0.2087227	-0.0626162	0.1912257	0.1912319	-0.0000155
Education (2)	0.3740095	0.3738318	0.0003669	0.3787146	0.3787164	-3.67E-06
Education (3)	0.0935024	0.1308411	-0.1183803	0.1043357	0.1043448	-0.0000289
Smoking Status	0.2123613	0.1931464	0.0477503	0.2036188	0.2036129	0.0000147

Table 13: Variance Ratio Diagnostic Results for the MCS Depression and Overall Outcomes Simple Model.

	Raw			Matched (ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2482077	0.2288162	1.084747	0.2500137	0.2503955	0.9984751
Age	171.4144	158.835	1.079198	173.1821	162.2879	1.067129
Education (1)	0.150278	0.1656737	0.9070721	0.154904	0.1551456	0.9984425
Education (2)	0.234498	0.2348131	0.9986583	0.2356633	0.2360256	0.9984653
Education (3)	0.0848942	0.1140771	0.7441828	0.0935981	0.093749	0.99839
Smoking Status	0.1675295	0.1563279	1.071655	0.1624156	0.1626614	0.9984886

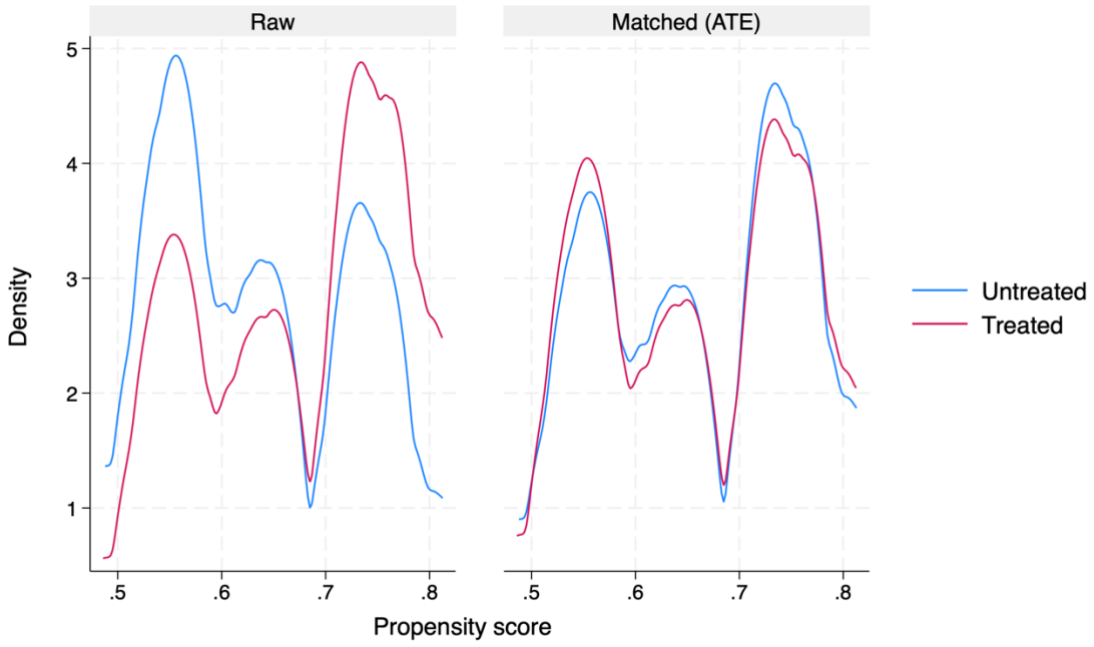

Figure 7: Overlap Plot Diagnostic Graphs for the MCS Depression and Overall Outcomes Simple Model.

Table 14: Standard Difference Diagnostic Results for the MCS Severe Psychological Symptoms and Leg Outcomes Simple Model.

	Raw			Matched (ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4162011	0.5839112	-0.339883	0.528346	0.5283602	-0.0000288
Age	43.32682	44.08738	-0.058402	43.81276	43.81283	-5.61E-06
Education (1)	0.1899441	0.1830791	0.0176067	0.186223	0.1862221	2.38E-06
Education (2)	0.3631285	0.3938974	-0.063405	0.3836006	0.3836034	-5.69E-06
Education (3)	0.0977654	0.1206657	-0.073395	0.1101076	0.11011	-7.71E-06
Smoking Status	0.2374302	0.1733703	0.1588943	0.192627	0.1926216	0.0000134

Table 15: Variance Ratio Diagnostic Results for the MCS Severe Psychological Symptoms and Leg Outcomes Simple Model.

	Raw			Matched(ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2436584	0.2432963	1.001488	0.2498945	0.2495418	1.001414
Age	165.1842	174.0021	0.9493232	171.0901	172.7008	0.9906735
Education (1)	0.1542964	0.1497688	1.03023	0.1519685	0.1517539	1.001414
Education (2)	0.231914	0.2390738	0.9700518	0.2371135	0.2367802	1.001408
Education (3)	0.0884544	0.1062529	0.8324891	0.0982583	0.0981218	1.001391
Smoking Status	0.1815642	0.1435121	1.265149	0.1559575	0.1557345	1.001432

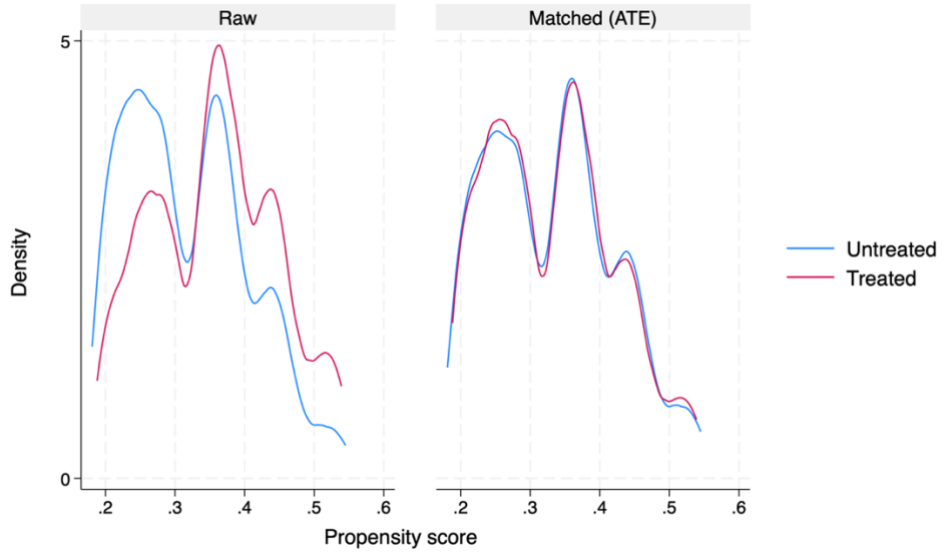

Figure 8: Overlap Plot Diagnostic Graphs for the MCS Severe Psychological Symptoms and Leg Outcomes Simple Model.

Table 16: Standard Difference Diagnostic Results for the MCS Severe Psychological Symptoms and Overall Outcomes Simple Model

	Raw			Matched(ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4066265	0.5790323	-0.3496858	0.5191504	0.5191428	0.0000155
Age	43.56627	44.14516	-0.0448459	43.9608	43.96078	1.53E-06
Education (1)	0.189759	0.1935484	-0.0096164	0.1933195	0.1933193	4.45E-07
Education (2)	0.3493976	0.3870968	-0.0781306	0.3735224	0.3735207	3.62E-06
Education (3)	0.0993976	0.1096774	-0.0335654	0.1032786	0.1032781	1.79E-06
Smoking Status	0.246988	0.183871	0.1537949	0.204584	0.2045866	-6.43E-06

Table 17: Variance Ratio Diagnostic Results for the MCS Severe Psychological Symptoms and Overall Outcomes Simple Model.

	Raw			Matched(ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2420103	0.2441477	0.9912457	0.2503874	0.2500368	1.001402
Age	165.0802	168.1824	0.9815542	171.0036	166.4587	1.027303
Education (1)	0.154215	0.1563396	0.9864108	0.1564182	0.1561989	1.001404
Education (2)	0.2280057	0.2376361	0.9594739	0.2347104	0.234381	1.001405
Education (3)	0.0897882	0.097806	0.9180227	0.092892	0.0927613	1.001408
Smoking Status	0.1865468	0.1503049	1.241123	0.163221	0.1629938	1.001394

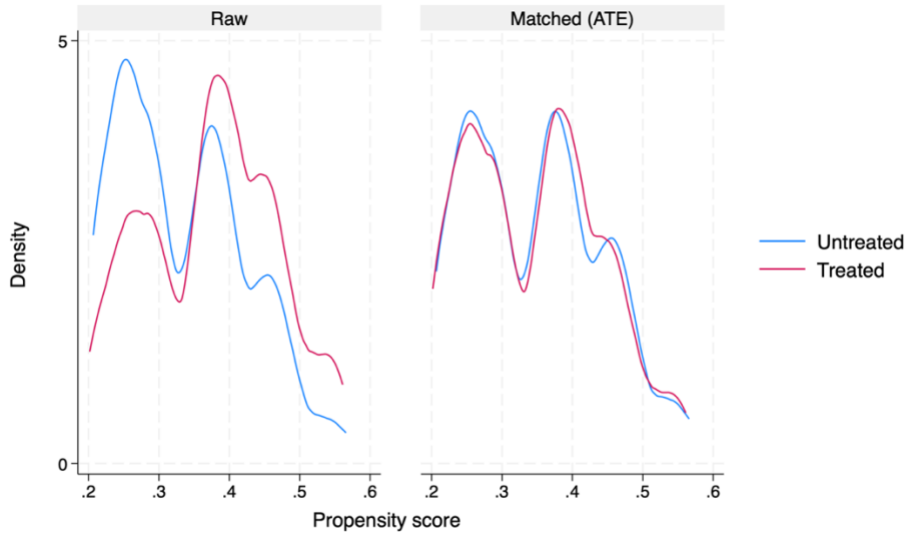

Figure 9: Overlap Plot Diagnostic Graphs for the MCS Severe Psychological Symptoms and Overall Outcomes Simple Model.

Inclusive Adjustment Set:

Table 18: Standard Difference Diagnostic Results for the PHQ-8 and Leg Outcomes Inclusive Model.

	Raw			Matched(ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4707965	0.6109785	-0.2838364	0.5298425	0.529892	-0.0001004
Age	43.15221	44.62291	-0.1117497	43.97458	43.97515	-0.0000428
Education (1)	0.2088496	0.1551313	0.1394214	0.1866395	0.1866206	0.0000491
Education (2)	0.3769912	0.3818616	-0.0100267	0.3798036	0.3798055	-3.87E-06
Education (3)	0.079646	0.1670644	-0.2679233	0.1158419	0.1158727	-0.0000943
Smoking Status	0.2070796	0.1813842	0.0649188	0.1962137	0.1962046	0.000023
Pre-Op PCS	28.80345	30.80574	-0.2605289	29.57736	29.57805	-0.0000902
Surgical Wait Time	105.492	111.2554	-0.0390233	108.397	108.3991	-0.0000145

Table 19: Variance Ratio Diagnostic Results for the PHQ-8 and Leg outcomes Inclusive Model.

	Raw			Matched(ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2495889	0.2382524	1.047582	0.2495511	0.2497024	0.9993941
Age	168.491	177.9149	0.9470314	176.5061	175.4755	1.005873
Education (1)	0.1655244	0.1313791	1.259899	0.1520743	0.1521565	0.9994602
Education (2)	0.2352853	0.236608	0.9944095	0.2359705	0.2361168	0.9993803
Education (3)	0.0734325	0.1394868	0.5264476	0.1026041	0.1026913	0.9991515
Smoking Status	0.1644888	0.1488392	1.105144	0.1579935	0.1580856	0.9994172
Pre-Op PCS	50.59492	67.53913	0.7491201	53.18622	64.92346	0.8192141
Surgical Wait Time	18915.65	24708.83	0.7655422	19830.31	21224.01	0.9343337

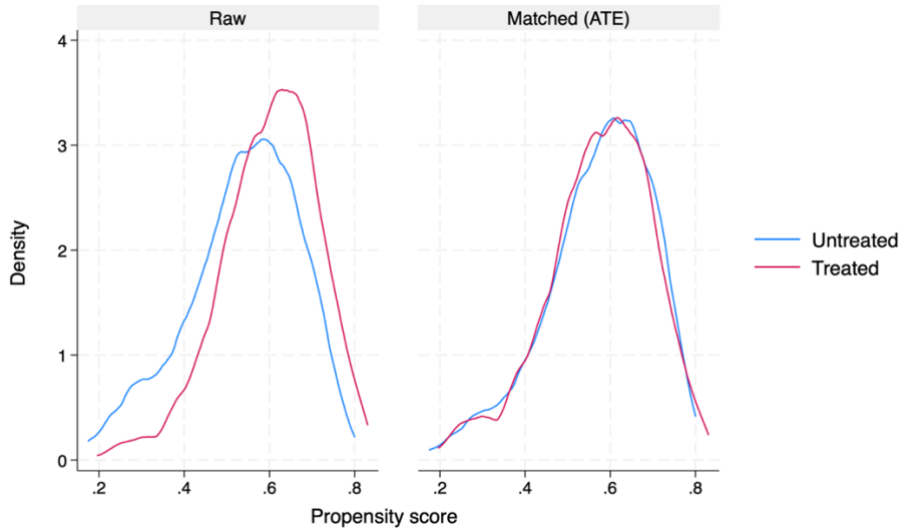

Figure 10: Overlap Plot Diagnostic Results for the PHQ-8 and Leg Outcomes Inclusive Model.

Table 20: Standard Difference Diagnostic Results for the PHQ-8nand Overall Outcomes Inclusive Model.

	Raw			Matched(ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4628571	0.6129032	-0.3040569	0.5228083	0.5228911	-0.0001678
Age	43.1981	44.7478	-0.1191266	43.98181	43.98275	-0.0000719
Education (1)	0.207619	0.1671554	0.1037101	0.1916388	0.1916166	0.000057
Education (2)	0.3657143	0.3782991	-0.0260079	0.3725047	0.3725125	-0.0000161
Education (3)	0.08	0.1524927	-0.2273407	0.1074003	0.1074395	-0.000123
Smoking Status	0.2152381	0.1935484	0.0537411	.2071208	0.2071091	0.000029
Pre-Op PCS	28.61666	29.68273	-0.1456866	29.03484	29.03542	-0.0000799
Surgical Wait Time	108.5105	107.9765	0.0038149	109.7683	109.7688	-3.03E-06

Table 21: Variance Ratio Diagnostic Results for the PHQ-8 and Overall outcomes Inclusive Model.

	Raw			Matched(ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2490949	0.2379507	1.046834	0.2499559	0.2502098	0.9989854
Age	167.0027	171.4597	0.9740052	175.5355	167.8135	1.046015
Education (1)	0.1648273	0.1396239	1.180509	0.155209	0.1553553	0.9990587
Education (2)	0.23241	0.2358806	0.9852866	0.234191	0.2344344	0.9989618
Education (3)	0.0737405	0.1296188	0.5689026	0.0960484	0.0961783	0.9986494
Smoking Status	0.169233	0.1565465	1.08104	0.1645352	0.1646979	0.9990119
Pre-Op PCS	49.9345	57.15788	0.8736242	51.04215	55.59027	0.9181851
Surgical Wait Time	19932.69	19246	1.03568	20267.49	18764.68	1.080087

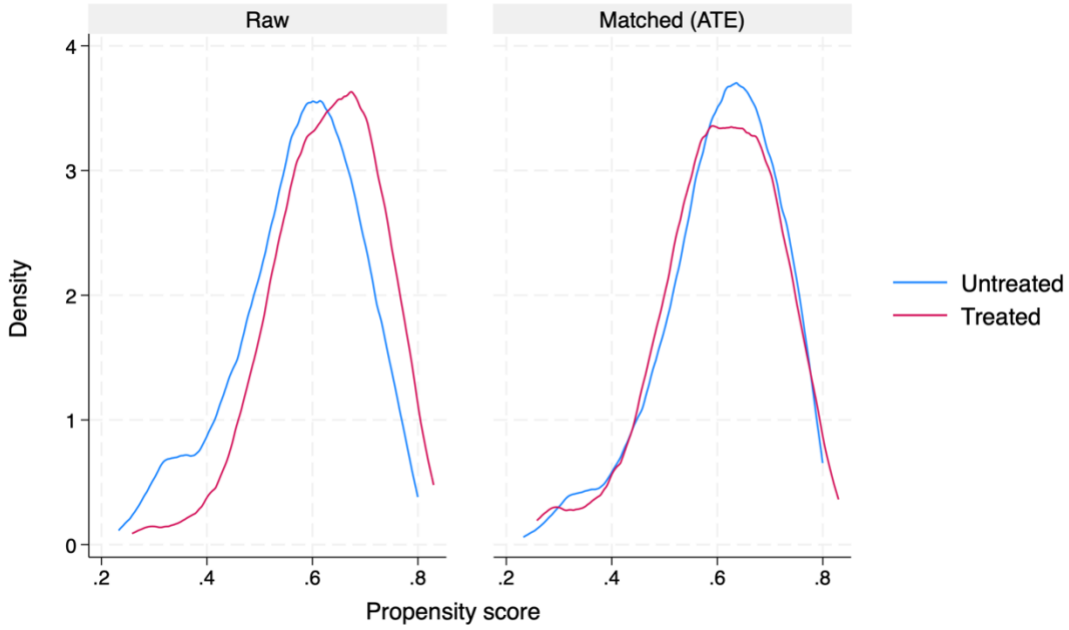

Figure 11: Overlap Plot Diagnostic Results for the PHQ-8 and Overall Outcomes Inclusive Model.

Table 22: Standard Difference Diagnostic Results for the MCS Depression and Leg Outcomes Inclusive Model.

		Raw			Matched(ATE)	
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.468254	0.6412429	-0.3530791	0.5284184	0.5280609	0.0007296
Age	43.67143	43.96893	-0.0226436	43.72271	43.72223	0.0000368
Education (1)	0.1873016	0.1836158	0.0094727	0.183257	0.1832716	-0.0000377
Education (2)	0.3793651	0.3785311	0.0017173	0.3837804	0.38377	0.0000213
Education (3)	0.1	0.1468927	-0.1427515	0.1158362	0.1157406	0.0002912
Smoking Status	0.2015873	0.1864407	0.0382683	0.192784	0.1928242	-0.0001016
Pre-Op PCS	30.73954	27.7278	0.3929443	29.92541	29.93104	-0.0007343
Surgical Wait Time	100.7841	120.6921	-0.131832	108.9212	108.877	0.000293

Table 23: Variance Ratio Diagnostic Results for the MCS Depression and Leg Outcomes Inclusive Model.

		Raw			Matched(ATE)	
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.249388	0.2307021	1.080996	0.2495886	0.2499186	0.9986796
Age	174.0302	171.1973	1.016547	176.0767	176.4132	0.9980928
Education (1)	0.1524617	0.1503257	1.014209	0.1499118	0.1501072	0.9986985
Education (2)	0.2358215	0.2359117	0.9996177	0.236869	0.2371605	0.9987706
Education (3)	0.0901431	0.1256702	0.7172988	0.102581	0.1026346	0.9994777
Smoking Status	0.1612057	0.1521102	1.059795	0.1558657	0.1560839	0.9986019
Pre-Op PCS	51.45475	66.03567	0.7791964	50.22478	73.06219	0.6874251
Surgical Wait Time	17446.51	28161.56	0.619515	20915.87	18351.05	1.139764

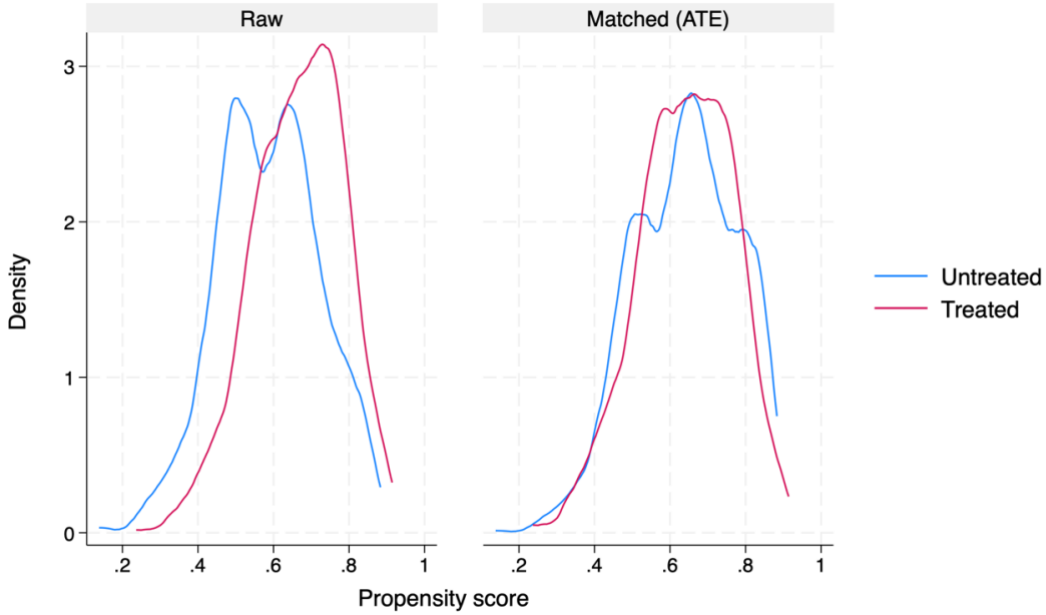

Figure 12: Overlap Plot Diagnostic Graphs for the MCS Depression and Leg Outcomes Inclusive Model

Table 24: Standard Difference Diagnostic Results for the MCS Depression and Overall Outcomes Inclusive Model

	Raw			Matched(ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4618056	0.6413793	-0.366637	0.5206738	0.5200342	0.0013059
Age	43.72222	43.97931	-0.0198707	43.72102	43.72056	0.0000355
Education (1)	0.1892361	0.1965517	-0.0185173	0.1883006	0.1882923	0.0000209
Education (2)	0.3680556	0.3758621	-0.0161311	0.3746888	0.3746393	0.0001024
Education (3)	0.0989583	0.1275862	-0.0903007	0.1081106	0.1080098	0.0003177
Smoking Status	0.2135417	0.1931034	0.0507332	0.2056592	0.2057383	-0.0001964
Pre-Op PCS	30.24561	26.63478	0.501728	29.31479	29.3263	-0.0015995
Surgical Wait Time	103.2865	118.2586	-0.1050265	109.8554	109.7932	0.0004363

Table 25: Variance Ratio Diagnostic Results for the MCS Depression and Overall Outcomes Inclusive Model

	Raw			Matched(ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2489734	0.2308078	1.078705	0.2500066	0.2504623	0.9981807
Age	173.2027	161.5843	1.071903	175.3513	165.5975	1.0589
Education (1)	0.1536926	0.1584656	0.9698802	0.1531093	0.1533672	0.9983185
Education (2)	0.2329952	0.2354015	0.9897777	0.2347046	0.2350954	0.9983378
Education (3)	0.0893207	0.1116931	0.799697	0.0965904	0.0966771	0.999103
Smoking Status	0.1682337	0.1563537	1.075982	0.1636476	0.1639755	0.9980004
Pre-Op PCS	46.79211	56.79569	0.8238674	46.53472	65.50281	0.7104232
Surgical Wait Time	18178.1	22466.38	0.8091245	20964.75	15905.99	1.318041

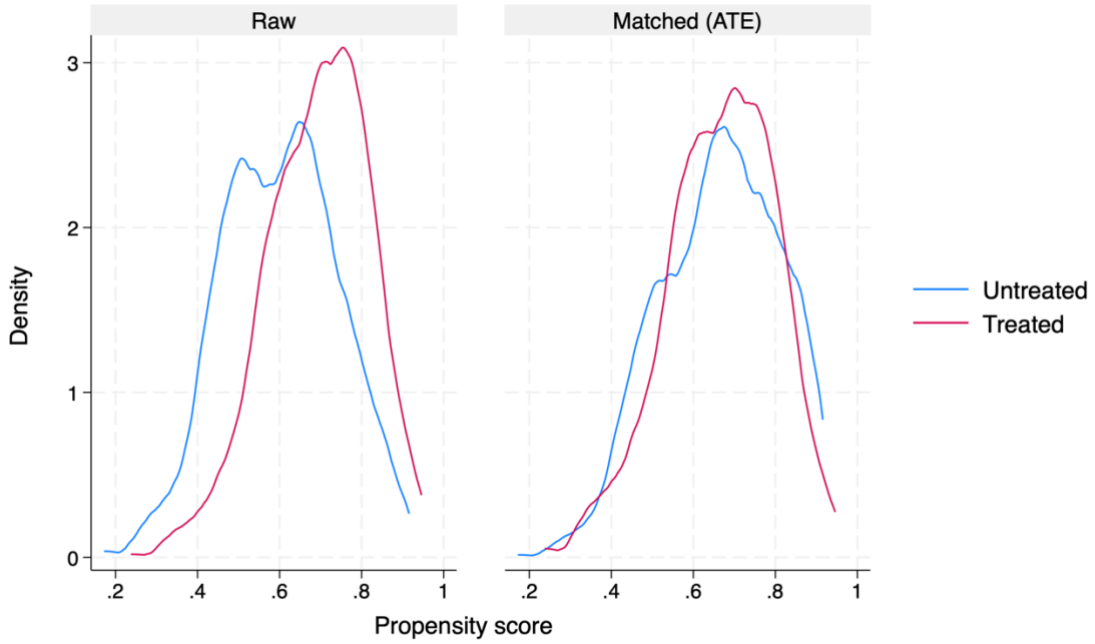

Figure 13: Overlap Plot Diagnostic Graphs for the MCS Depression and Overall Outcomes Inclusive Model

Table 26: Standard Difference Diagnostic Results for the MCS Severe Psychological Symptoms and Leg Outcomes Inclusive Model.

	Raw			Matched(ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4215385	0.5842185	-0.3293791	0.5285145	0.5290408	-0.0010656
Age	43.01538	44.15478	-0.0871017	43.73942	43.74323	-0.0002913
Education (1)	0.1907692	0.1836115	0.0183296	0.1856239	0.185603	0.0000535
Education (2)	0.36	0.3884674	-0.0587842	0.3757146	0.3758215	-0.0002207
Education (3)	0.1046154	0.1229135	-0.0575866	0.1149058	0.1149735	-0.0002131
Smoking Status	0.2430769	0.1729894	0.1731111	0.193985	0.1937728	0.0005241
Pre-Op PCS	31.15808	28.91529	0.3052434	30.05037	30.04132	0.0012311
Surgical Wait Time	95.98462	113.8452	-0.1253116	109.5902	109.6389	-0.0003418

Table 27: Variance Ratio Diagnostic Results for the MCS Severe Psychological Symptoms and Leg Outcomes Inclusive Model.

	Raw			Matched(ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2445964	0.2432764	1.005426	0.249956	0.2495353	1.001686
Age	166.3362	175.9	0.9456291	174.8166	174.6471	1.00097
Education (1)	0.1548528	0.1501261	1.031485	0.1516342	0.1513842	1.001651
Education (2)	0.2311111	0.2379215	0.9713754	0.2352771	0.2349362	1.001451
Education (3)	0.0939601	0.1079696	0.8702459	0.1020164	0.1019093	1.001051
Smoking Status	0.1845584	0.1432815	1.288083	0.1568374	0.1564623	1.002397
Pre-Op PCS	43.14134	64.83102	0.6654429	39.63534	71.1572	0.5570109
Surgical Wait Time	17361.73	23267.6	0.7461764	21716.88	19979.36	1.086966

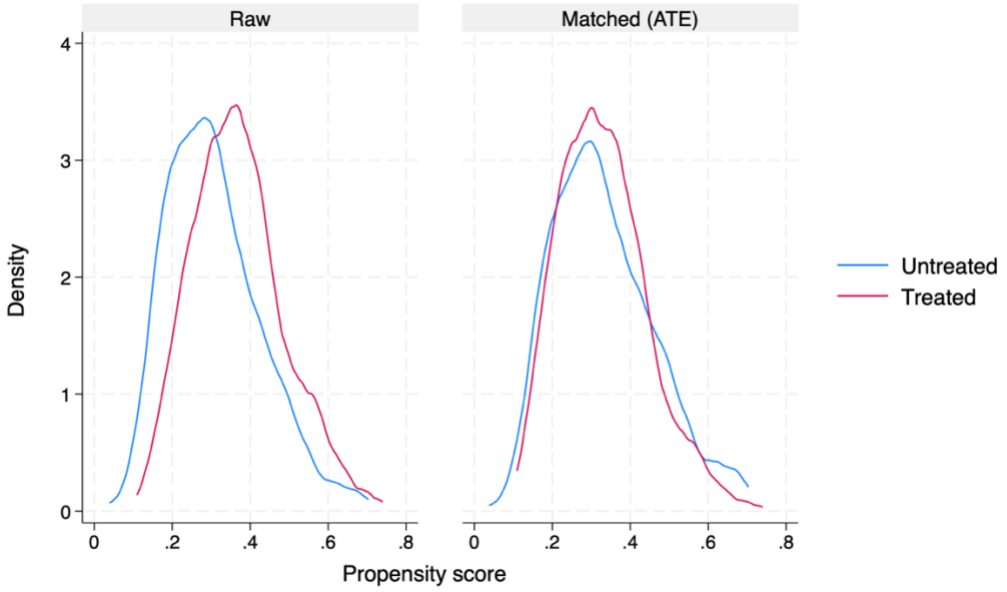

Figure 14: Overlap Plot Diagnostic Graphs for the MCS Severe Psychological Symptoms and Leg Outcomes Inclusive Model.

Table 28: Standard Difference Diagnostic Results for the MCS Severe Psychological Symptoms and Overall Outcomes Model

	Raw			Matched (ATE)		
Means	Treated	Untreated	StdDif	Treated	Untreated	StdDif
Sex	0.4152824	0.5787611	-0.3309999	0.522562	0.5231185	-0.0011269
Age	43.2093	44.12743	-0.0706769	43.80751	43.81066	-0.0002421
Education (1)	0.1893688	0.1929204	-0.0090211	0.1920048	0.1920155	-0.0000271
Education (2)	0.345515	0.3840708	-0.0800586	0.3662339	0.3663862	-0.0003163
Education (3)	0.1063123	0.1097345	-0.011011	0.1065041	0.1065252	-0.0000679
Smoking Status	0.2524917	0.1823009	0.170564	0.2059198	0.2056832	0.000575
Pre-Op PCS	30.83818	28.07658	0.3946465	29.4521	29.44073	0.0016238
Surgical Wait Time	99.13289	113.1841	-0.1011141	109.1383	109.1826	-0.0003183

Table 29: Variance Ratio Diagnostic Results for the MCS Severe Psychological Symptoms and Overall Outcomes Inclusive Model.

	Raw			Matched(ATE)		
Variances	Treated	Untreated	Ratio	Treated	Untreated	Ratio
Sex	0.2436323	0.244229	0.9975571	0.2503226	0.2499078	1.00166
Age	167.4927	170.0156	0.9851606	175.1414	166.26	1.053419
Education (1)	0.1540199	0.1559782	0.9874455	0.1556561	0.1554206	1.001515
Education (2)	0.2268882	0.2369799	0.9574154	0.2328803	0.232559	1.001382
Education (3)	0.0953267	0.0978661	0.9740525	0.0954782	0.0953463	1.001383
Smoking Status	0.1893688	0.1493316	1.268109	0.1640619	0.1636673	1.002411
Pre-Op PCS	41.28158	56.65275	0.7286775	37.14367	64.02484	0.5801446
Surgical Wait Time	18306.95	20314.84	0.9011613	21037.6	17565.98	1.197633

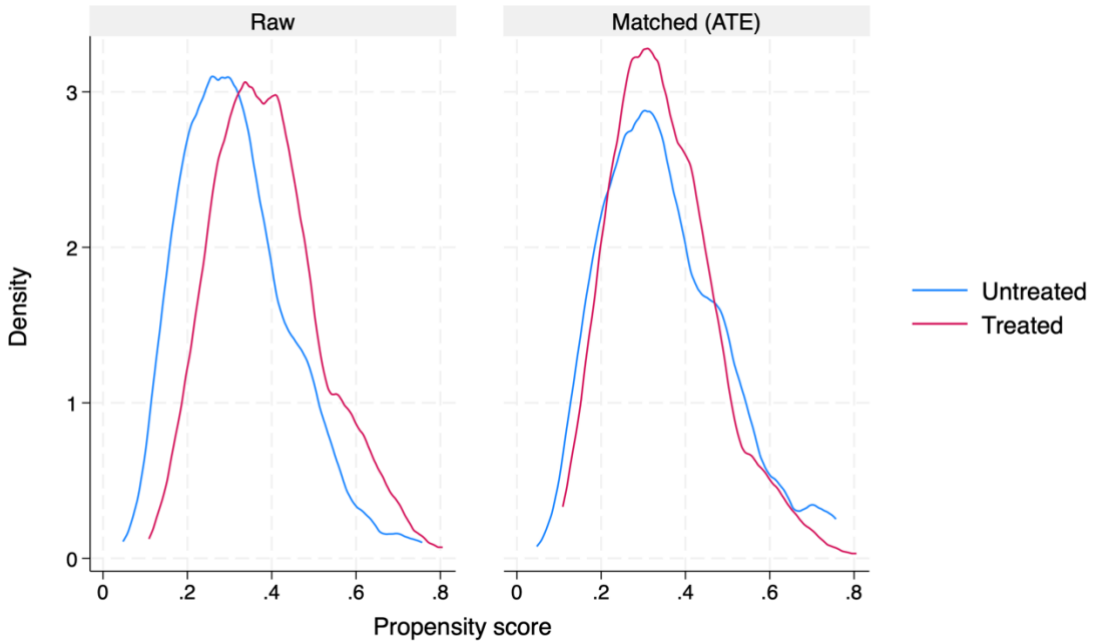

Figure 15: Overlap Plot Diagnostic Graphs for the MCS Severe Psychological Symptoms and Overall Outcomes Inclusive Model.

Curriculum Vitae

Hannah Van Brenk

University of New Brunswick, Bachelor of Science in Kinesiology (Honours), 2023

Publications: N/A

Conference Presentations: N/A