PATIENT PREFERENCES, REFERRAL PRACTICES, AND SURGEON ENTHUSIASM FOR DEGENERATIVE LUMBAR SPINAL SURGERY

by

S. Samuel Bederman, MD MSc FRCSC

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

The Institute of Medical Science
University of Toronto

© Copyright by S. Samuel Bederman 2010
Patient Preferences, Referral Practices, and Surgeon Enthusiasm for Degenerative Lumbar Spinal Surgery

S. Samuel Bederman, MD MSc FRCSC

Doctor of Philosophy

The Institute of Medical Science
University of Toronto

2010

Abstract

Degenerative disease of the lumbar spine (DDLS) is a common condition for which surgery is beneficial in selected patients. Wide variation in surgical referral and rates of surgery has been observed contributing to unequal access to care.

Our objectives were to examine (1) the variation in preferences for referral and surgery among surgeons, family physicians (FPs) and patients, (2) how FP referral practices compare with preferences and guideline recommendations, and (3) how the ‘enthusiasm’ of patients and physicians influence regional variation in surgical rates.

We used conjoint analysis in a mailed survey to elicit preferences based on clinical vignettes from surgeons, FPs and patients. A Delphi expert panel provided consensus guideline recommendations for surgical referral to compare with actual FP referral practices. Rates of surgery for DDLS, obtained from Ontario hospital discharge data, were used to quantify regional variation and regression models assessed the relationship with patient and physician enthusiasm.

We identified significant differences in preferences for referral and surgery between patients, FPs and surgeons. Surgeons placed high importance on leg-dominant symptoms while patients
had high importance for quality of life symptoms (i.e. severity, duration, walking tolerance).

Surgical referral practices were poorly predicted by individual FP preferences and guideline recommendations based on clinical factors alone. Variation in Ontario surgical rates was higher than that of hip or knee replacements and was highly associated with the enthusiasm of surgeons (p<0.008), rather than FPs or patients.

By appreciating the variation in preferences between patients and physicians, and exploring other non-clinical factors that influence referrals, we may be able to improve the efficiency of referrals and enhance the shared decision making process. With an understanding of the influence that surgeons have in driving variation in surgical rates, further research may allow us to direct strategies to improve access and allow for a more equitable delivery of care for patients suffering from DDLS.
Acknowledgments

I would like to acknowledge the generous salary support from Clinical Research Training Fellowship provided by the Orthopaedic Research and Education Foundation and the American Academy of Orthopaedic Surgeons. Research grant support for this study was generously provided by the Ontario Neurotrauma Foundation along with the Ontario Ministry of Health and Long-Term Care (2007-SCI-SSS-524).

I am grateful to Drs. Steve Gallay of the Ontario Orthopaedic Association and Hamilton Hall along with the Canadian Spine Society for their endorsements, assistance with surgeon participation, and recognition of the importance of this work. I would additionally like to thank the individual physicians who participated on our expert consensus panel.

I gratefully acknowledge the statistical advice from Dr. Derek Stephens and the Biostatistics, Design & Analysis Service of the Child Health Evaluative Sciences Research Program at the Hospital for Sick Children. Statistical and methodological advice for the small-area variation analysis was provided by Ruth Croxford and Dr. Therese Stukel from the Institute for Clinical Evaluative Sciences (ICES). Special recognition also goes to Kinwah Fung from ICES who assisted in preparing the administrative data. Liza Mercier and David Northrup from the Institute for Social Research at York University were instrumental in planning and conducting our province-wide survey.

I would also like to express my sincere gratitude to Drs. Peter Coyte, Hans Kreder, Nizar Mahomed, and Warren McIsaac for participating on this thesis committee. Their input, advice and time commitment to this project have been truly invaluable.

I am indebted to Dr. James Wright for his tireless effort as my thesis supervisor. His thoughtful feedback and comments were always welcomed and sincerely appreciated. His advice both academic and personal was always well received and valued. I am privileged to consider him my mentor, colleague, and friend.

Finally, only my wife, Eden, and my children, Maxwell, Sydney, and Bryce could have provided me with the optimal love and support during this time. I spent many of my most cherished moments with my family while working on this thesis. I dedicate this work to the four of you.
# Table of Contents

Acknowledgments.......................................................................................................................... iv  
Table of Contents............................................................................................................................. v  
List of Abbreviations ..................................................................................................................... ix  
List of Tables ......................................................................................................................................... xi  
List of Figures ....................................................................................................................................... xiii  
List of Appendices ........................................................................................................................... xv  

CHAPTER 1: Surgery on the Degenerative Lumbar Spine – A Rationale for Inquiry and Review of the Literature.............................................................................................................. 1  
1 Introduction........................................................................................................................................ 1  
2 The Clinical Picture and Burden of Disease ................................................................................... 1  
3 Cast of Characters ........................................................................................................................ 5  
   3.1 The Family Physician ............................................................................................................. 5  
   3.2 The Surgeon ......................................................................................................................... 8  
   3.3 The Patient ........................................................................................................................... 9  
4 The Shared Decision Making Process ......................................................................................... 10  
5 Current Trends and Regional Variation in Surgical Rates ....................................................... 12  
6 The Influence of Enthusiasm on Practice Patterns .................................................................... 13  
7 Theoretical Perspectives and a Conceptual Framework ........................................................... 14  
8 Research Aims and Hypotheses ................................................................................................. 17  
9 Methodological Considerations ................................................................................................. 18  
   9.1 Selection of Variables ............................................................................................................. 18  
   9.2 Definition of Terms .............................................................................................................. 18  
   9.3 Analytical Methods .............................................................................................................. 19  
10 Potential Significance ............................................................................................................... 19  
11 Dissertation Outline ................................................................................................................... 21
CHAPTER 2: In the Eye of the Beholder: How do the Preferences of Patients, Family Physicians and Surgeons Compare in Decision Making for Lumbar Spinal Surgery?..........................26

1 Introduction...........................................................................................................................................26

1.1 Background.......................................................................................................................................26

1.2 Specific Objective.............................................................................................................................27

2 Methods..................................................................................................................................................27

2.1 Conjoint Analysis...............................................................................................................................27

2.2 Questionnaire Development............................................................................................................28

2.3 Study Population................................................................................................................................31

2.4 Statistical Analyses ..........................................................................................................................33

3 Results...................................................................................................................................................36

3.1 Survey Response and Participant Characteristics........................................................................36

3.2 Conjoint Analysis...............................................................................................................................37

4 Discussion............................................................................................................................................40

5 Conclusions.......................................................................................................................................46

CHAPTER 3: How do Family Physicians’ Referral Practices Compare with their Preferences and Guidelines for Degenerative Lumbar Spinal Surgery?.................................................................59

1 Introduction...........................................................................................................................................59

1.1 The Degenerative Lumbar Spine in Primary Care and the Use of Clinical Practice Guidelines .................................................................59

1.2 Specific Objectives ..........................................................................................................................60

2 Methods...............................................................................................................................................60

2.1 Selection of Guidelines....................................................................................................................60

2.2 Questionnaire Design and Conjoint Analysis.................................................................................60

2.3 Delphi and the Expert-Panel ............................................................................................................61

2.4 Surgical Referral in Primary Care...................................................................................................63

2.5 Statistical Analyses ..........................................................................................................................63
2.3 Explanation of Regional Variation ................................................................. 109

3 Limitations .............................................................................................................. 111

4 Significance and Future Directions ....................................................................... 113
   4.1 Influencing Physician Behaviour ................................................................ 114
   4.2 The Role of the Patient .............................................................................. 116
   4.3 Novel Strategies ......................................................................................... 117

5 Concluding Remarks ............................................................................................... 120

References .................................................................................................................... 124

Appendices ................................................................................................................... 143

Copyright Acknowledgments ....................................................................................... 174
List of Abbreviations

AAOS  American Academy of Orthopaedic Surgeons
AHCPFR  Agency for Health Care and Policy Research
AUC  Area Under the Curve
CA  Conjoint Analysis
CCHS  Canadian Community Health Survey
CCI  Canadian Classification of Interventions
CI  Confidence Interval
CIHI  Canadian Institutes for Health Information
CPG  Clinical Practice Guidelines
CPSO  College of Physicians and Surgeons of Ontario
CPT  Current Procedural Terminology
CT  Computed Tomography
CV  Coefficient of Variation
DAD  Discharge Abstracts Database
DDLS  Degenerative Disease of the Lumbar Spine
EQ  Extremal Quotient
FP  Family Physician
ICD-9/10  International Classification of Diseases, Ninth/Tenth Revision
ICES  Institute for Clinical Evaluative Sciences
LHIN  Local Health Integration Network
LR  Likelihood Ratio
MRI  Magnetic Resonance Imaging
NA  North American
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHIP</td>
<td>Ontario Health Insurance Plan</td>
</tr>
<tr>
<td>P4P</td>
<td>Pay for Performance</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver Operating Characteristic</td>
</tr>
<tr>
<td>SCV</td>
<td>Systematic Component of Variance</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SE</td>
<td>Standard Error</td>
</tr>
<tr>
<td>SF-36</td>
<td>Short Form-36</td>
</tr>
<tr>
<td>SPORT</td>
<td>Spine Patient Outcomes Research Trials</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>WOMAC</td>
<td>Western Ontario and McMaster Osteoarthritis Scale</td>
</tr>
</tbody>
</table>
List of Tables

Table 2-1. Clinical Factors and Levels for Conjoint Experiment................................................ 47

Table 2-2. Table of Presented Clinical Vignettes. ....................................................................... 48

Table 2-3. Physician Characteristics............................................................................................ 49

Table 2-4. Patient Characteristics. ............................................................................................... 50

Table 2-5. Validation of Conjoint Analysis Responses.............................................................. 51

Table 2-6. Regression Results for Preferences about Surgical Referral..................................... 52

Table 2-7. Regression Results for Preferences about Surgery.................................................... 53

Table 3-1. Panelist-Group Correlations over Two Iterations....................................................... 72

Table 3-2. Patient Characteristics. ............................................................................................... 73

Table 3-3. Regression Results for Consensus Panel (Guideline) Recommendations and Family Physician Preferences for Surgical Referral ................................................................. 74

Table 4-1. Characteristics of DDLS Surgical Discharges in Ontario, 2002-6............................. 92

Table 4-2. Comparison of Small-Area Variation Measures with Other Procedures in Ontario. 93

Table 4-3. Characteristics of Ontario Counties............................................................................ 94

Table 4-4. Regression Results for County Spinal Surgical Rates................................................ 95

Table A 1. Identification of Surgical Cases from CCI Code Discharges ................................. 165

Table A 2. ICD-10 Diagnostic Codes for Exclusions in Surgical Case Finding ......................... 165

Table A 3. OHIP Procedural Codes for Exclusions in Surgical Case Finding.......................... 166

Table A 4. Diagnosis Codes for “Definite” Lumbar Surgery..................................................... 168

Table A 5. Diagnosis codes for “Possible” Lumbar Surgery....................................................... 169
Table A 6. “Fusion” OHIP Procedure Codes, identified by “Definite” Lumbar Surgery Cohort

Table A 7. “Decompression” OHIP Procedure Codes, identified by “Definite” Lumbar Surgery Cohort
List of Figures

Figure 1-1. The Spectrum of Degenerative Disease of the Lumbar Spine ................................. 22
Figure 1-2. Andersen's Behavioural Model of Health Service Use ............................................. 23
Figure 1-3. O'Neill and Kuder's Physician Propensity Framework ........................................... 24
Figure 1-4. Conceptual Model for Preferences of Patients and Physicians in the Utilization of Health Services in Degenerative Disease of the Lumbar Spine ............................................. 25
Figure 2-1. In and Out of County Surgical Discharges by Surgeon Response .................... 54
Figure 2-2. Surgical Referral Preference Quartile Responses .................................................. 55
Figure 2-3. Relative Importance of Clinical Factors in the Decision for Surgical Referral ...... 56
Figure 2-4. Surgical Preference Quartile Responses ................................................................. 57
Figure 2-5. Relative Importance of Clinical Factors in the Decision for Surgery .................... 58
Figure 3-1. Relative Importance of Clinical Factors in the Decision for Surgical Referral ...... 75
Figure 3-2. Receiver Operating Characteristic Curves for Prediction of Referral by Family Physician Preference and Guideline Recommendation ................................................. 76
Figure 4-1. Province-wide Age and Sex Specific Rates of DDLS Surgery, Ontario, 2002-6 ..... 96
Figure 4-2. Total Discharges for DDLS Surgery by Procedure, Ontario, 2002-6 ................... 97
Figure 4-3. Variation in LHIN DDLS Surgical Rates ............................................................... 98
Figure 4-4. Variation in County DDLS Surgical Rates ............................................................ 99
Figure 4-5. Regional variation in DDLS surgical rates across Ontario ................................. 100
Figure 5-1. Correspondence Postcards ................................................................................. 122
Figure 5-2. The Bone and Joint Health Network's Model of Care ......................................... 123
Figure A 1. Map of 49 Ontario Counties ................................................................................... 172

Figure A 2. Map of 14 Ontario Local Health Integration Networks (LHINs)............................ 173
List of Appendices

Appendix A. Questionnaires for Surgeons, Family Physicians, and Patients ......................... 143

Appendix B. Diagnostic and Procedural Inclusion and Exclusion Codes for Degenerative Lumbar Spinal Surgery Cohort Identification ................................................................. 165

Appendix C. Maps of Ontario.................................................................................................. 172
CHAPTER 1: Surgery on the Degenerative Lumbar Spine – A Rationale for Inquiry and Review of the Literature

1 Introduction

Degenerative disease of the lumbar spine (DDLS) is a common and often disabling disorder, typically beginning in the fifth or sixth decade of life and usually manifesting as chronic back or leg pain.\textsuperscript{1-4} The annual prevalence of back pain ranges from 25\% to 62\% globally.\textsuperscript{5-7} It comprises one of the most common reasons for visits to the family physician (FP), the costs of this condition are substantial, and surgical rates have been increasing.\textsuperscript{8-15} Furthermore, there exists substantial variation in surgical rates across geographic regions not necessarily attributed to variation in the prevalence of disease.\textsuperscript{16} Variation in referral practices from FPs, recommendations from surgeons, and patient preferences are thought to influence variation in degenerative lumbar spinal surgical rates but this has not been well studied.

2 The Clinical Picture and Burden of Disease

DDLS encompasses a spectrum of disease entities ranging from asymptomatic radiographic disc degeneration to spondylosis with spinal stenosis and degenerative instabilities, such as spondylolisthesis or scoliosis (Figure 1-1).\textsuperscript{17, 18}

Spinal motion occurs via the functional spinal unit which is comprised of the intervertebral disc anteriorly and the facet joint complexes posteriorly as well as the two adjacent vertebrae. The intervertebral disc contains an outer \textit{annulus fibrosis} which resists torsional stresses as well as an inner \textit{nucleus pulposis} which acts in a hydraulic manner to resist compressive loads. The degenerative process usually begins with the disc in early adulthood. Degenerative disc disease results in fissuring of the disc’s outer annulus and dehydration with loss of the normal proteoglycan and water content of the inner nucleus. The disc then loses its ability to withstand compressive loading and often results in disc bulging, fragmentation, herniation or loss of disc height and collapse. With abnormal motion and loss of normal resistance to loading, the facet joints posteriorly are then subjected to abnormal stresses which then accelerate their degeneration. Osteoarthritis of the facet joints, as in other degenerative joint disorders, can then lead to cartilage loss, formation of osteophytes and synovial cysts, and even joint laxity and
instability. Loss of the normal height of the spinal unit (facet cartilage and disc height loss) then leads to redundancy and infolding of the stabilizing interlaminar ligament (*ligamentum flavum*). The combination of the redundant ligament and osteophyte or cyst formation can lead to encroachment on the spinal canal resulting in compression of the neurological structures (i.e. spinal stenosis). Additionally, the loss of disc height along with facet instability can result in typical patterns of degenerative spinal instability, most commonly a forward slippage of one vertebra on another (spondylolisthesis) or even a side-to-side slippage or angulation (scoliosis).19-23

The typical patient presents with longstanding pain either in their back or radiating down their legs. Back pain may arise from the degenerative disc, arthrosis of the facet joints, spinal instability, or from other structures in close proximity (muscles, ligaments, etc.) With compression of spinal nerves from disc fragment herniations or canal encroachment from osteophytes, cysts, or ligament, patients may experience pain radiating down the legs along the dermatome of that nerve, altered sensation and progressive weakness of the limbs or loss of bowel or bladder control. This combination of pain and/or weakness ultimately impedes function and therefore patients’ quality of life.17, 19, 20

Complaints of low back pain account for the one of the most common reasons to seek care from a primary care physician second only to chest pain and abdominal pain.8 Carey et al.(1995) noted in their patient survey, that of those suffering from chronic low back pain, 73% had sought care from a health care provider, 91% of which was a medical practitioner.9 Although degenerative changes in the lumbar spine are a common finding on diagnostic imaging modalities, the majority of people remain asymptomatic.24 Few studies have addressed the natural history of spinal stenosis. Reviews have suggested that for most patients, symptoms remain stable or gradually worsen. It is believed that approximately 15% will improve over a 4-year period, 70% will remain the same, and another 15% will deteriorate.25-27

Because DDLS is common, it exists on a spectrum of disease severity, and it may remain stable for long periods of time, the choice of provider and treatment varies widely. Care-seeking from both medical as well as non-medical providers is commonplace and includes such practitioners as primary care providers, orthopaedic or neurosurgeons, neurologists, rheumatologists, physiatrists, pain specialists, as well as chiropractors, and physical therapists.28 Treatment
options for back pain are even more plentiful than types of providers and include such modalities as medications (antidepressants, non-steroidal anti-inflammatories [NSAIDs], narcotics, muscle relaxants, and herbals), acupuncture, back schools, cognitive behavioural therapy, exercise, injections, orthotics, laser therapy, braces, massage, bedrest, manual therapy, traction, radiofrequency denervation, and many others, however, only a few are supported by evidence (Cochrane Back Review Group [www.cochrane.iwh.on.ca]). Similarly, for patients with spinal stenosis or sciatica, many of these modalities have also been evaluated using a systematic evidence-based approach by the Cochrane Review Group and the Agency for Healthcare Research and Quality (formerly the Agency for Health Care Policy and Research). For patients with sciatica, for example, traction, bed rest, or epidural steroid injections have not been shown to have any significant long-term benefit despite their widespread use.

In practice, however, for symptomatic patients with degenerative spinal conditions, the mainstay of treatment consists of non-operative modalities, such as NSAID medications and physical therapy. For a minority of patients with moderate to severe symptoms who have failed to improve following non-operative treatment, surgery can yield significant benefit. Mirza and Deyo (2007), in their recent systematic review of lumbar fusion for chronic low back pain, found that fusion surgery poses some benefit over unstructured non-operative treatments, however, it may not be more beneficial than a structured course of cognitive behavioural therapy. More recent case series, however, suggest that the benefit of fusion for back pain depends on the diagnostic indication and that not all patients with back pain are equivalent. For example, patients with spondylolisthesis improve more following surgery than those with degenerative disc disease who, in turn, improve more than those with chronic low back pain.

For patients with spinal stenosis, Amundsen et al. (2000) assessed 31 patients who were randomly allocated to surgical and non-surgical treatment. The authors found that at four year follow-up, half of those who were randomized to non-surgical treatment had excellent or fair results, while four fifths of those who underwent surgery improved. The recent Cochrane systematic review found that decompression with lumbar fusion is beneficial in patients with stenosis associated with spondylolisthesis and that instrumentation increases the rate of fusion without significantly improving clinical results. Because prior systematic reviews failed to show a significant benefit of surgery over non-operative treatment for spinal stenosis, the Spine Patient Outcomes Research Trials (SPORT) were designed, as randomized trials with
observational preference arms, to determine the benefit of surgery for spinal stenosis, degenerative spondylolisthesis, and disc herniation.\textsuperscript{71-74}

In the trial for degenerative spondylolisthesis, the authors found that due to strong patient preferences there was tremendous cross-over (approximately 40% in each direction) within the randomized cohort questioning the value of a typical intention-to-treat analysis.\textsuperscript{71} As expected, the intention-to-treat analysis failed to demonstrate any difference, however, the observational as-treated comparison, controlling for measured confounders, found a substantially greater improvement for surgical treatment (most commonly decompression and instrumented fusion) over nonsurgical treatment which was maintained even out to four years.\textsuperscript{75} In the trial for spinal stenosis, the authors found a similarly high rate of cross-over to surgical treatment (43%), however, the intention-to-treat (as well as the as-treated) analysis found a significant benefit for surgery (i.e. decompression) over nonsurgical treatment.\textsuperscript{72}

Costs attributed to low back pain account for a substantial burden on society with indirect costs, resulting from lost work productivity, representing the majority of total costs.\textsuperscript{10} The economic burden of all musculoskeletal disorders in Canada was estimated to be over $25 billion in 1994 and disorders of the back and spine accounted for nearly one-third of this burden second only to injuries.\textsuperscript{11} A decade later in the United States, the total cost attributed to musculoskeletal diseases was estimated to be almost $850 billion, nearly 8% of the gross domestic product and accounted for the majority of both lost work and bed days due to health conditions compared with all other chronic medical conditions.\textsuperscript{12} Back pain accounted for more than 53 million health care visits in the US in 2004 and was the cause of over 300 million bed days and nearly 200 million lost work days.\textsuperscript{12}

Rates of spinal surgery, both in Canada and the US, have been steadily increasing over the past two decades.\textsuperscript{13-15} Rates of spinal fusion in the US increased 113% from 1996 to 2001, compared with only 13% and 15% in hip and knee replacement surgery, respectively.\textsuperscript{13} In the US, Medicare spending on inpatient back surgery more than doubled from 1992 to 2003. For lumbar fusions, the costs increased more than 5-fold to a total of $482 million and accounted for 47% of spending on inpatient back surgery. In Canada, rates of spinal fusion similarly increased by 63% from 1995 to 2001 while rates of decompressions dropped by 24%.\textsuperscript{14} With the aging population, the prevalence is expected to rise resulting in an increased burden on already strained healthcare resources.\textsuperscript{4, 76, 77}
Individual decisions surrounding the use of surgical treatment for DDLS are ultimately made by the patient and their surgeon which is facilitated by the FP.

3 Cast of Characters

3.1 The Family Physician

In the Province of Ontario, the FP serves as a gatekeeper for provincially insured specialty care. FPs, therefore, play an integral role in access to surgical treatment. Determining the ideal candidate for surgical referral is a difficult task for FPs since there is a wide spectrum of clinical presentations and only a minority of patients (<15%) that might benefit from surgery.\textsuperscript{17, 78-80}

In primary care, initial presenting symptoms for most patients with degenerative conditions of the lumbar spine are nonspecific. Often, the exact underlying cause is unknown, severe clinical conditions are only present in approximately 7% of patients, and diagnostic tests for these conditions have uncertain accuracy.\textsuperscript{78, 79} Katz and coworkers (1995) found that history and physical examination findings provided useful information in the appropriate diagnosis of spinal stenosis in patients with low back pain.\textsuperscript{81} Kent et al. (1992) in their meta-analysis, however, were unable to provide strong conclusions regarding the accuracy of magnetic resonance imaging (MRI) or computed tomography (CT) for the diagnosis of spinal stenosis.\textsuperscript{82} Kent and colleagues (2004, 2005) in their mailed surveys found that primary care physicians were least likely to recognize different conditions for patients complaining of low back pain and therefore were least likely to treat these patients differently.\textsuperscript{83, 84} In a systematic review of the literature, de Graff and co-authors (2006) concluded that making the correct diagnosis in patients with lumbar spinal stenosis is difficult since no single best diagnostic test exists.\textsuperscript{80} They suggested that proper diagnosis relies on a combination of radiographic (i.e. cross-sectional imaging such as MRI or CT) and clinical features (i.e. walking tolerance).

Because diagnosis and management of patients with DDLS is challenging in the primary care setting, ‘evidence-based’ clinical practice guidelines have been developed to improve the quality of care and assist FPs by providing recommendations on further investigation, treatment, and referral for a variety of degenerative lumbar spinal conditions. The first guideline was published by the Quebec Task Force in 1987.\textsuperscript{85} Following this, the most well known guideline was developed in 1994 by the Agency for Health Care and Policy Research (AHCPR).\textsuperscript{86} Since then,
fifteen other published guidelines from ten different countries for the treatment of low back pain have been identified by systematic reviews. Koes et al. (2001) found that all of the guidelines were generally similar since they were based on the same available evidence. Other systematic reviews found that overall the quality of the guidelines to be poor. van Tulder and colleagues (2004) found that the guidelines were similar with respect to diagnosis and treatment recommendations, however, more than half of them did not take patient preferences into account. These authors found that the guidelines put forth by the AHCPR scored the highest, not only on the inclusion of patient preferences, but obtained the highest total score on all domains of the AGREE (Appraisal of Guidelines, Research, and Evaluation in Europe) instrument for the assessment of quality. The AHCPR guidelines have since been used extensively as a ‘reference standard’ to assess physician guideline adherence.

Rossignol et al. (1996) performed a prospective observational study of 2147 low back pain patients identified by the worker’s compensation board in Quebec and found that although health service utilization for back pain was close to what was practiced elsewhere, referrals to specialists was far lower than the guideline recommendations. Little and co-workers (1996) performed a mailed questionnaire to 166 FPs in the UK to compare reported management with evidence-based guidelines and found that 15% of FPs would not refer to a specialist for neurological signs and that one-third of respondents rated their satisfaction with the management of back pain as 4 out of 10 or less. Negrini et al. (2001) performed a similar mailed questionnaire to 217 FPs in Italy and found that compared with the AHCPR guidelines, physicians tended to over-treat in acute back pain but under-treated in chronic back pain. More recently, however, Engers and colleagues (2005) performed a randomized trial on FPs in the Netherlands whereby physicians were randomized to a guideline educational session. These authors found that even with an intensive education program, there was only modest improvement in treatment, but no significant difference in rates of specialist referral.

Somerville and co-workers (2008) recently performed a systematic review of 53 randomized trials comparing any intervention to usual primary care for low back pain to describe the content of ‘usual care’ as it relates to guidelines. The authors found that treatments of back pain varied significantly and that they were not in-line with guideline recommendations.
Despite the widespread availability of guidelines, there exists a lack of adherence and inconsistent use.\textsuperscript{90,91,98-101} Physician knowledge, attitude and behaviour have been identified as barriers to guideline adherence and influence treatment decisions.\textsuperscript{101} In general, lack of familiarity or awareness of guidelines are the main reasons for incomplete knowledge. The main barriers related to physician attitude typically include a lack of agreement with the guidelines (or lack of confidence about the quality of evidence on which they are based) or a lack of motivation to follow them. The main behavioural barriers include reconciling patient preferences with guideline recommendations, specific characteristics of the guidelines or the presence of contradictory guidelines, and a lack of time, resources, or reimbursement to implement the recommendations.

Multiple factors have been found to influence guideline adherence and referral practices. Langley and co-workers (1992) found from their physician survey and interviews of 41 FPs and 20 consultants, that referral practices were highly influenced by patient and family factors (patient wishes), FP/consultant factors (capabilities of the FP) and the style of practice.\textsuperscript{102} A subsequent physician survey by Langley and colleagues (1997) found that regional practice characteristics (such as access to hospital facilities and remoteness from specialist care) and the relationship with the consultant were among the most important non-medical factors influencing decisions for referral.\textsuperscript{103} In a series of interviews with twenty patients with low back pain and their FPs, Schers et al. (2001) found that although FPs were well aware of and agreed with the guidelines, non-adherence to the guidelines was related primarily to patient demands.\textsuperscript{104}

Personal characteristics of physicians (i.e. gender, years in practice, location of training, the type of practice), and even the personal beliefs, attitudes and behaviour of physicians can influence physician practice patterns and guideline adherence in the management of patients with degenerative lumbar spine conditions.\textsuperscript{101,103,105-109}

Both over- and under-referral of patients with degenerative lumbar conditions have been observed.\textsuperscript{110-114} Roland and colleagues (1991) found that nearly 43\% of outpatient referrals to orthopaedic surgeons in the UK were inappropriate and that the rates were highest for referrals of back pain.\textsuperscript{112} Similarly, Fertig and co-authors (1993) found that there was wide variation in referral rates and 33\% of orthopaedic referrals were inappropriate, significantly higher than any other specialty investigated.\textsuperscript{111} Oldmeadow and co-workers (2007) found that 70\% of referrals for back pain were considered appropriate for non-surgical management.\textsuperscript{113} In comparing referral
rates with the best available evidence at the time, Schroth and colleagues (1992) found that 14% of referrals for back pain to surgeons were inappropriate, but 47% of patients with potential indications for surgery were not referred suggesting a problem of both over- and under-referral in this condition.\textsuperscript{114} Under-referral of patients may prevent appropriate patients from obtaining beneficial care, while over-referral burdens the system further lengthening waiting lists and prolonging the pain and disability for those patients appropriate for surgery.

In a systematic review, O’Donnell (2000) found that there was wide variation in referral rates from primary care and that patient characteristics explain less than 40% of the practice variation and physician and practice characteristics explain less than 10%.\textsuperscript{115} O’Donnell concluded that much of the variation is unexplained and that under-referral may be a larger problem than over-referral.

### 3.2 The Surgeon

Although there is general consensus about absolute indications for surgery (i.e. tumours, infections, fractures and dislocations, and cauda equina syndrome), for the majority of patients, the decision to undergo surgery is discrentional, based largely on the severity and pattern of their clinical symptoms and the preference of patients. Based on a systematic review of the literature, Aalto and coworkers (2006) found several clinical factors, such as walking ability, duration of symptoms, and leg pain predominance, to be important predictors of outcome following surgery.\textsuperscript{116}

There exist a multitude of surgical options for the treatment of DDLS: minimally invasive decompressions (e.g. laminotomy), standard open decompressions (e.g. laminectomy), spinal fusions which may be instrumented, non-instrumented, performed through a posterior approach, anterior approach or both, either with or without interbody fusion devices (i.e. cages). With advances in technology, alternative surgical options include interspinous spacer devices,\textsuperscript{117} non-fusion stabilization techniques,\textsuperscript{118} artificial facet\textsuperscript{119} or disc replacements.\textsuperscript{120, 121}

With such an abundance of surgical options, it is tempting to believe that no one surgical option provides a definitive solution for this clinical entity. Variation in surgeon indications for surgery and preferences for the surgical technique, therefore, vary widely. Within orthopaedics, surgical variation in opinion has been observed in indications for knee replacements and rotator cuff
surgery, for example. In spinal surgery, there exists variation in surgical preferences for the treatment of thoraco-lumbar trauma, and degenerative conditions of the cervical spine.

Irwin and colleagues (2005) performed a survey of 30 surgeons (22 orthopaedics, 8 neurosurgeons) to assess surgical preferences for five degenerative lumbar spinal conditions. They assessed the surgeons’ recommendations for the use of fusion, instrumentation, and surgical approach for (1) multilevel stenosis without deformity or instability, (2) degenerative spondylolisthesis with stenosis, (3) isthmic (spondylolytic) spondylolisthesis, (4) degenerative scoliosis with stenosis, and (5) recurrent stenosis for prior laminectomy. The authors found that there was wide variation in the recommendations for surgery and that surgeon-specific characteristics, such as training background (specialty, fellowship training) and years in practice had a profound influence on the choice of surgical procedure (i.e. fusion vs. decompression, the use of instrumentation, and anterior vs. posterior approach). Specifically, they found that younger surgeons preferred instrumentation and orthopaedic surgeons recommended fusion and instrumentation over neurosurgeons. For the case of degenerative scoliosis with stenosis, for example, 25 of 30 surgeons recommended surgery, however, half (13 of 25) recommended a decompression alone while half (12 of 25) recommended fusion in addition to decompression. Of those who recommended fusion, eight surgeons (67%) recommended a posterior-only procedure, one (8%) recommended an anterior-only procedure, and three (25%) recommended a combined anterior and posterior operation. Despite this lack of agreement on treatment, there was, however, consensus on the overall need for surgery in the majority of these cases.

Bederman and co-workers (2009) performed a population-based study of surgery for the degenerative lumbar spine in Ontario and found that surgeon specialty (i.e. neurosurgery vs. orthopaedic surgery) and surgical volume were highly associated with the types of procedures performed (e.g. decompression without fusion vs. fusion with/without decompression), suggesting that, “what surgical procedure you get depends on who you see”.

3.3 The Patient

In addition to differences in preferences among physicians, wide variation exists in patients’ preferences towards treatment for low back pain and degenerative lumbar spinal conditions. Perceived quality of life, an understanding of the underlying pathology, prognosis, and methods for secondary prevention or alternative therapy, and general attitudes such as help-seeking
behaviour and satisfaction with primary care management all influence patients’ perception of their condition and their preferences for treatment decisions.\textsuperscript{128, 129} The importance of including patient preferences in guiding clinical care has been well established.\textsuperscript{130-132} Even on a broader level, it has been argued that the economic value of health interventions requires a measure that accounts for patients’ desirability of various health outcomes.\textsuperscript{133}

In a commentary on patient preferences and their relationship with health disparities, Katz (2001) explains that although physicians clearly play active roles in medical decision making, the preferences of patients ultimately drive decisions.\textsuperscript{134} Moreover, he purports that variation in patient preferences may even explain some of the disparities in the use of health services. Robinson and Thomson (2001) suggest that highly variable patient preferences for treatment decisions may be associated with many factors including age, sex, and the type of clinical problem at hand (i.e. decisions with potential mortality vs. quality of life).\textsuperscript{135} They imply that patients may not fully appreciate how their preferences impact on decisions and therefore we need to better understand patient preferences and how they may be influenced.

In addition to variation in patient preferences, the expectations of patients are equally important because they directly relate to overall satisfaction with treatment.\textsuperscript{136} In a series of 84 patients followed prospectively following surgery for nerve root compression, McGregor and Hughes (2002) found that pre-operative expectations were higher than postoperative expectations despite the majority of patients satisfied with their decision for surgery.\textsuperscript{137} Patients’ expectations of surgery, and thus their satisfaction, may be influenced by previous experiences, observations of friends and relatives, influences from the medical team, or even the media.\textsuperscript{138} Preference-based scoring instruments that weight perceptions of pain highly best predict outcomes in patients with low back pain, suggesting that treatment success highly depends on patients’ values of their own pain.\textsuperscript{139} Owens (1998), therefore, advocated the need to include patient preferences into practice guidelines.\textsuperscript{131} He suggested that since differences in preferences lead to different preferred therapy, physician decisions, based on practice guidelines that do not consider patient preferences, will ultimately lead to suboptimal recommendations.

4 The Shared Decision Making Process

How patients value their own condition, their perception of outcome, and their attitude toward risk all guide their enthusiasm for seeking and accepting treatment recommendations. Patients often
possess values quite different from their physicians.\textsuperscript{140} Since physicians, from their own experiences and training, also possess variable enthusiasm for recommending treatment, what is the benefit to resolving preferences between physicians and their patients? It has been shown that greater agreement in patient-physician decisions is associated with improved patient satisfaction and overall health status.\textsuperscript{141} Thus, the shared decision making process that aligns preferences of patients and physicians, can not only reduce variation in the utilization of health services, but may also improve individual patient outcomes.

In their editorial, Kravitz and Melnikow (2001) explain that patients ought to be more involved in their own care which can ultimately produce better health outcomes.\textsuperscript{142} They suggest that the complexity in decision making surmount clinical uncertainty and variation in self-perceived health utility. Medical decision making also depends highly on the perceived risk of treatment as well as the interaction with family and culture. Even with a fully informed patient, shared decision making remains complex. Because greater agreement in patient-physician decisions is associated with improved patient satisfaction and overall health status, decision aids aimed at improving the decision making process for patients with DDLS have resulted in more patients opting for surgery and feeling better informed about their condition.\textsuperscript{141, 143} Despite the evidence to suggest that patient preferences may drive healthcare decisions, Weiner and Essis (2006), in a survey of 200 patients considering spinal surgery, found that although patients felt strongly about the need for complete risk information, the majority of patients felt that the treating physician, rather than the patient, should be making the basic treatment decisions.\textsuperscript{144}

Surgeons and FPs place different value on the importance of differing clinical factors in deciding on the referral of a patient to a surgeon and in deciding to recommend surgery for their patient. Because perceptions and attitudes can vary significantly, it is not surprising, therefore, that generalist and specialist physicians express different preferences for treatment.\textsuperscript{145, 146} Dreinhofer and colleagues (2006) found major differences in importance of clinical factors between referring FPs and surgeons concerning indications for total hip replacements.\textsuperscript{147} It remains unclear how referring physicians compare with surgeons concerning the importance of accepted clinical factors in influencing their decisions for referral and surgery of patients with degenerative spinal disorders. Physician characteristics, clinical uncertainty, and individual preferences all contribute to variation in treatment preference for patients.
Rainville et al. (2000) performed a mailed survey to 63 orthopaedic spine surgeons and 79 FPs to explore recommendations for activity and return to work for patients with chronic low back pain. The authors found that recommendations varied widely and that surgeons were less restrictive in their recommendations to return to work or activity. Specific recommendations were influenced mostly by the personal attitudes of the physicians in addition to patient clinical factors.

Harrold et al. (1999) explain that specialists and generalists possess differences in three main areas, namely, knowledge, patterns of care, and clinical outcomes of care. As one might expect, they found that specialists were more knowledgeable about their area and quicker to adopt new and effective treatments than generalists. Donohoe (1998) suggests that despite these differences, the quality of care could be improved through changes in training, optimizing referrals, and strategies for generalist-specialist comanagement.

5 Current Trends and Regional Variation in Surgical Rates

Regional variation in spinal surgery across geographic regions in the US has been shown to be greater than most other procedures. In the 1995 US Medicare population, Birkmeyer and colleagues (1998) found that the variation in rates of back surgery was high, second only to radical prostatectomy. Weinstein and coworkers (2006) found that the coefficient of variation (the ratio of the standard deviation to the mean rate) for lumbar decompression surgery and fusion was approximately 35 and 50, respectively, for Medicare patients (1992-2002) compared with only 12 and 25 for hip fracture and hip replacement surgery, respectively. They further found that the extremal quotient (the ratio of highest to lowest rate regions) was 12 for lumbar decompressions and over 20 for lumbar fusion. The authors theorized that physician-related factors are likely the explanation. They suggested that a lack of scientific evidence, financial incentives, and differences in clinical training and professional opinion are underlying causes and argued that patient preferences are unlikely to explain any major differences in rates.

Wennberg and Gittlesohn (1973, 1982) first introduced the concept that there exists wide variations in resource input, utilization, and expenditures across small areas and coined the term, small-area variation. Fisher and Wennberg (2003) defined three categories of medical care, namely, effective care, preference-sensitive care, and supply-sensitive care. Effective care refers to health care services with proven benefit with little tradeoff such that every patient with a specific need should receive it (e.g. breast/colon cancer screening, blood lipid monitoring for
diabetics, beta-blockers following myocardial infarctions, etc.). Preference-sensitive care refers to services with proven benefits that do involve tradeoffs and are therefore discretionary decisions that should be based on the preferences of patients (e.g. prostatectomy vs. radiation for early stage prostate cancer, lumpectomy vs. mastectomy for localized breast cancer, and surgery vs. non-operative treatment for degenerative conditions). Supply-sensitive care refers to the services provided in the absence of specific theories of benefit (e.g. hospital lengths of stay, frequency of physician follow-up or diagnostic tests, inpatient vs. outpatient treatment of chronic disease). For each category, significant variation poses a different problem. The authors believed that variation in effective care suggests a problem of underuse; variation in preference-sensitive care suggests a problem of misuse; and variation in supply-sensitive care suggests a problem of overuse.

Regional variation in care, for discretionary conditions, such as DDLS, can at least partially be explained by variation in individual physicians’ enthusiasm (or preference) for recommending treatment for their patients and predictions on who will benefit from this treatment. Incorrect perceptions often generate this variability, it is associated with undesirable error, even when the ideal surgical rate is unknown. This error can manifest itself in several ways. In areas of high utilization or low utilization, some patients may undergo surgery unnecessarily and some patients may not undergo surgery that may, in fact, benefit from surgery. In other discretionary procedures for degenerative disease, such as hip and knee replacement, it has even been shown that there exists greater unmet need in high-utilization areas suggesting that the ideal rate may, in fact, be higher than that in high-rate regions and under-utilization is the primary problem.

6 The Influence of Enthusiasm on Practice Patterns

This phenomenon of small-area variation, has been shown for some conditions to be better explained by differences in health care delivery, specifically, physician enthusiasm or physician uncertainty, rather than by differences in disease prevalence or availability of resources. Although not previously studied, it has been hypothesized that enthusiasm of patients also generate substantial regional variation in care. The influence of provider enthusiasm was first hypothesized by Chassin (1993). Conventional wisdom, at that time, suggested that inappropriateness and clinical uncertainty were the main influences on regional variation. Chassin showed that geographic differences in the utilization of health services may be caused by differences in the ‘enthusiasm’ of physicians for particular services.
Wright and co-authors (1999) examined the variation in knee replacement surgery across Ontario.\textsuperscript{158} The authors looked at measures of surgeon and referring physician (FP and rheumatologists) ‘enthusiasm’ as predictors of surgical rates. After controlling for the characteristics of the population, access to surgery (hospital beds, number of orthopaedic surgeons), and the prevalence of disease, they found that the enthusiasm of regional orthopaedic surgeons for surgery, relative to their peer group, was the dominant modifiable determinant of this variation.\textsuperscript{158}

Coyte and colleagues (2001) assessed variation in rates of middle-ear surgery across Ontario.\textsuperscript{159} The authors also measured the ‘enthusiasm’ of otolaryngologists and referring physicians (FPs and paediatricians) and after controlling for access to surgery (number of referring physicians and specialists) and population characteristics, found that the enthusiasm of referring physicians rather than surgeons that was the dominant driver of surgical rates.

Although not previously studied, it has been hypothesized that the enthusiasm of patients may also generate substantial regional variation in care.\textsuperscript{134, 157} On an ecologic basis, the enthusiasm of treating surgeons as well as referring physicians may all influence health service utilization. It therefore remains unclear whether, for discretionary spinal surgery, the system is being driven by providers, referrers, or patients.

7 Theoretical Perspectives and a Conceptual Framework

In formulating a conceptual framework for this dissertation to study variation in preferences towards spinal surgery and how varying enthusiasm may influence regional surgical rates, a theory explaining the behaviour of both patients and physicians on an individual and aggregate (population) level is required. Preferences are best explored on an individual basis while the influence of these preferences relative to the peer group (i.e. enthusiasm) on regional utilization rates is best studied on an ecologic basis over a population.

Andersen’s Behavioral Model of Health Services Use (1995) is a prominent and comprehensive framework for explaining patient behaviour in the utilization of health services that has been improved on and applied over a 40 year period (Figure 1-2).\textsuperscript{160} In this model, the patients’ health behaviour is primarily influenced by population characteristics including \textit{predisposing characteristics, enabling resources} and \textit{need}. Andersen further sub-classifies \textit{predisposing characteristics}, \textit{enabling resources} and \textit{need}.
characteristics into demographics (i.e. age, gender), social structure (i.e. occupation, income, education level), and health beliefs. Enabling resources encompass both personal resources (i.e. insurance coverage, waiting times) as well as the resources of the community (i.e. medical facilities and availability of personnel). Andersen dichotomizes need into patients’ perceived need, or their preference for the utilization of services, and evaluated need, the physician’s interpretation or recommendation for health service utilization.

In the Province of Ontario, universal insurance coverage under the Ontario Health Insurance Plan (OHIP) covers payments for all medically necessary consultations and treatments provided by physicians, such as surgical referral and surgery. In light of this, the personal resources of the patient play less of a role in Ontario than, for example, in the US. Furthermore, since we are interested in studying the preferences for health services, the overall health beliefs of patients are not independent of their specific preferences. That is, patients’ health beliefs are highly correlated with their preferences (perceived need) and even possibly correlated with the preferences of their physicians. Clinical factors most likely act in concert with patient demographics and socioeconomic status thereby explaining a significant amount of variation in the likelihood of considering surgical referral or surgery for DDLS.

Andersen’s model provides an excellent scaffold for the study of health service utilization from the patient perspective, and although a patient’s perceived need is clearly influenced by a physician’s evaluated need, one disadvantage of the model is that it falls short in fully explaining physician influences. Many factors influence physicians’ determination of need, particularly in the case of discretionary surgery for conditions such as DDLS.

The Physician Propensity Framework proposed by O’Neill and Kuder (2005) serves as a useful model for explaining the factors that influence physician behaviour in the recommendation of health service use (Figure 1-3). In this model there are three main ‘stages’ or influences on physician propensity which lie on a macro, meso, and micro level, respectively. First, from a clinical baseline heuristic, the clinical training and experience of the physician sets the background for medical decisions. Second, medical decisions may be modified through practice-specific factors such as the organizational setting (group vs. solo practice), the practice population, and remuneration. Third, individual physician decisions are further refined based on the specific clinical characteristics, the preferences of patients, as well as other characteristics
that might include static non-clinical factors such as employment status or perceived outcome as well as dynamic factors such as the patient-physician interaction itself.

Since reimbursement is provided by OHIP for nearly all physicians in Ontario, the financial rewards by procedure vary little. The practice-specific heuristic is thus better measured by the clinical volume of patients seen for this condition and the type of practice (university vs. community). Finally, in studying how clinical factors, along with physician characteristics influence physician decisions, we have not included the patient’s non-clinical factors (patient preferences, patient characteristics) in our model. Therefore, clinical factors of the individual patients alone account for the patient-specific strategy in our model.

Based on our interpretation of these two complimentary theoretical models, we speculate that, in the case of conditions with discretionary interventions, both patient and physician factors heavily influence the perceived and evaluated need (or preferences) for treatment, which, in turn, ultimately impacts on health service utilization. Our conceptual model is depicted in Figure 1-4.

On an ecologic level, the regional demographics (age, gender) and regional social structure (socioeconomic status) form the basis of the predisposing characteristics. Measures of community access (per capita supply of surgeons, FPs, and MRI scanners) characterize the enabling community resources whereas personal access may be better assessed by other barriers, such as lack of familiarity with official languages. The enthusiasm of patients and physicians portray the perceived and evaluated need for surgery, respectively. All of these components, along with a measure of regional disease prevalence, potentially explain significant variation in the rates of surgery for DDLS in Ontario. Although we have developed a practical conceptual framework for the investigation into preferences of patients and physicians and the influence on surgical rates, the relationships in the framework are linear and do not completely characterize the dialogue or interaction that occurs between patients and physicians. Acknowledging that this model simplifies the relationships, we feel that it incorporates the predominant factors that influence the utilization of services for DDLS and serves as a starting point for the investigation of influences on utilization rates.
8 Research Aims and Hypotheses

8.1.1 Variation in Patient and Physician Preferences

Research Aim: To evaluate how preferences of patients, FPs, and surgeons compare for decisions on referral and surgery for DDLS and if they are associated with patient demographics or physician characteristics.

We expect to see a wide range of opinions on the recommendations for surgery or surgical referral on an individual level but on the aggregate we expect less discrepancy. In regards to the specific clinical factors that are important in driving these decisions, we hypothesize that surgeons will place a high emphasis on the presence of spinal stenosis (i.e. leg-dominant pain) while patients will place more importance on symptoms that interfere more with their quality of life.

8.1.2 Primary Care Referral Practices

Research Aim: To investigate how referral practices in primary care compare with FP preferences and recommendations arising from the evidence for surgical referral of patients with DDLS.

Based on our review of the literature, we expect to find that actual referral practices are different from stated preferences of FPs as well as guideline recommendations since patient preferences and consultant relationships play major roles in decisions for consultant referrals.

8.1.3 The Influence of Patient and Physician Enthusiasm on Surgical Rates

Research Aim: To quantify the regional variation in surgery for the degenerative lumbar spine across Ontario and to determine the influence of the enthusiasm of patients, FPs, and surgeons, controlling for the regional demographics, community resources, and prevalence of disease.

We hypothesize that there exists wide variation in regional surgical rates even across a single-payer system such as in the Province of Ontario. As in the case of knee replacements, we expect that surgeon enthusiasm is a major driver of surgical rates, however, based on the hypotheses of others, patient enthusiasm may also likely be influential. Prevalence of disease and community resources seem unlikely to play significant roles.
9 Methodological Considerations

9.1 Selection of Variables

In selecting the variables to study each of the three research questions, we employed our conceptual model shown in Figure 1-4. Preferences of patients and physicians were represented by the perceived and evaluated need, respectively. Preference for surgery (and surgical referral) was measured on an ordinal Likert scale ranging from 1 (least prefer) to 6 (most prefer). Since patient clinical factors (characteristics) heavily influence the preference for surgery, we used a variety of hypothetical clinical vignettes to provide us with overall measures of preference based on different clinical severities of disease. For patients, preference for surgery in a theoretical model (like hypothetical vignettes) may be influenced by their predisposing characteristics, such as age, gender, and education level and less so by income, language, or community resource barriers. For physicians, their preference for surgery may be influenced by their clinical experience (years in practice, location of training, and any specialty training) as well as their practice characteristics (such as academic vs. community, and their typical volume of similar patients seen). These factors along with the specific clinical characteristics of the patients potentially influence the recommendations for surgery in hypothetical scenarios.

For the aggregate model, the use of health services included the regional rates of lumbar spinal surgery across Ontario for degenerative conditions (i.e. lumbar decompressions and fusions). In our model, these rates may be influenced by regional disease burden (prevalence of back pain), the perceived and evaluated need for surgery (‘regional enthusiasm’ of patients and physicians), the community enabling resources (regional per capita surgeon and FP supply, and regional supply of MRI scanners), personal enabling resources (regional percentages of familiarity with official languages), and the predisposing characteristics of the population (age, gender, income, and education level).

9.2 Definition of Terms

There are many terms used in common language to describe how individuals feel about taking certain actions (e.g. preference, enthusiasm, opinions, attitudes, etc.). Although definitions vary across disciplines of study, for the purpose of this dissertation, we use the term ‘preference’ to describe the affinity that an individual or group possesses for a particular course of action. In this thesis, preference has been measured on an ordinal scale from 1 to 6 whereby the qualifiers
depend on the individual responding and the question asked. In general, the lowest response indicates a strong preference against an action while the highest response indicates the strongest preference in favour of an action. In this regard, the range of preferences may vary between individuals depending on their attitudes towards taking action.

Alternatively, we reserve the term ‘enthusiasm’ to reflect the overall preference that an individual or group possesses for a course of action relative to the overall preferences of the entire peer group. Specifically, we measured enthusiasm as an average of the deviation of individual preference responses relative to the overall preference for that peer group. The mathematical formulae are given in detail in Chapter 4.

9.3 Analytical Methods

This dissertation employed a mixed-methods approach. Although this is predominantly a hypothesis-testing quantitative analysis, we have used qualitative methods, such as a consensus process, to assist with the analysis. Surveys are the main tool used in this research using a series of clinical vignettes within the framework of a conjoint analysis (discussed further in Chapter 2). Vignettes in the survey were used to obtain preferences towards spinal surgery from practicing surgeons and FPs as well as a purposeful sample of patients. The consensus process was used solely to obtain an interpretation of an evidence-based guideline recommendation for surgical referral using the same vignettes. Responses were then used to determine regional measures of ‘enthusiasm’ for patients and physicians. Population-based administrative databases were used to obtain hospital discharge data to quantify area variation and regression techniques were used to explain regional surgical rates using the survey-based measures of ‘enthusiasm’ as well as regional characteristics.

10 Potential Significance

By understanding preferences for surgery for DDLS among patients, FPs, and surgeons, and where significant sources of variation and disagreement exist, we may identify potentially modifiable determinants of regional variation. For example, we aim to identify which stakeholders (patients or physicians) are significantly different from each other and which ones significantly influence regional variation in care.
If physician enthusiasm is a determinant of regional variation which could be reduced then reducing variation in enthusiasm could have several beneficial effects. First, aligning preferences of patients and physicians would improve the shared decision making process itself and patients’ expectations about surgery would be more accurate, resulting in a significant improvement in patient satisfaction with the healthcare process and even overall health status following treatment.\textsuperscript{141}

Second, examining preferences in the referral process and understanding what modifiable factors influence why physicians deviate from the evidence, may lead to improvements in the sensibility and framing of guidelines to incorporate patient preferences. Additionally, this understanding can allow us to direct strategies to modify physician behaviour. Furthermore, the identification of other factors important in the decision making of patients and physicians can provide directions for future studies. Reducing this variation in the referral process for patients with DDLS may result in fewer referrals for less-appropriate patients and more referrals for more-appropriate patients. The net effect being a more efficient referral process, reduced overall waiting lists for surgeons and duration of pain and disability for patients.

Last, reducing regional variation in surgical rates for DDLS can result in either a reduction in utilization for less-severe patients and/or an increase in utilization for more-severe patients. Even without knowing what the ideal rate should be, we speculate that there is under-utilization in low-rate regions.\textsuperscript{154} Whether surgical rates in high-rate regions should be hampered remains unclear. In any event, both potential outcomes can reduce disparity and improve the overall economic value of healthcare.\textsuperscript{133}

As individual stakeholders, physicians, patients, and policymakers may all profit from answers to these questions. Surgeons stand to benefit by a better understanding of the preferences of patients which may improve the shared decision making process, the efficiency of care and even, potentially, the outcomes of treatment. By learning of the regional variation in surgical rates across the province, surgeons would gain a perspective of how their practice compares with that of their peers. Referring physicians would better understand the preferences of consultants in selecting appropriate patients for surgery and may also be able to reduce variation in referral practices. Patients can stand to benefit following a reduction in practice variation by under- and over-utilized regions regressing towards the most appropriate rate for that region thus reducing
access disparities. Policymakers, informed on the regional variation in utilization rates, could better plan resource use to increase surgical rates in low-rate regions and examine the resource investment in high-rate regions to further improve geographic disparities in access.

In summary, by investigating the preferences of patients and physicians and understanding how they relate to regional surgical rates, we can better address ways of reducing practice variation, improving the referral process, reducing long waiting lists, boosting patient satisfaction, and ensuring a more equitable and cost-effective delivery of healthcare for all patients.

## 11 Dissertation Outline

In Chapter 1, we have provided a general introduction to the importance of understanding preferences of patients and physicians towards discretionary health services such as surgery for the degenerative lumbar spine. The chapter has also included a detailed review of the literature and the formulation of a conceptual framework for investigating the influence of preferences on the utilization of health services in this condition and outlines our objectives for this dissertation. Chapter 2 explores the differences in preferences for referral and surgery between patients, FPs, and surgeons using the methods of conjoint analysis. Chapter 3 provides further insight into the specific case of preferences for referral within the primary care setting by comparing FP preferences for surgical referral and a consensus interpretation of the recommendation for surgical referral from a widely-used guideline with patient-reported physician referral practices. Chapter 4 examines the regional variation in utilization of surgery for the degenerative lumbar spine across Ontario and attempts to explain this variation ecologically by regional patient, FP, and surgeon enthusiasm for surgery controlling for regional demographics, access to care, and the prevalence of disease. Chapter 5 describes the overall conclusions from this work with a detailed discussion, an outline of limitations, and suggestions for future direction.

Ethics approval was obtained from the Research Ethics Boards of the Hospital for Sick Children and the University of Toronto.
Figure 1-1. The Spectrum of Degenerative Disease of the Lumbar Spine.

Figure 1-2. Andersen's Behavioural Model of Health Service Use.

Figure 1-3. O'Neill and Kuder's Physician Propensity Framework.

Figure 1-4. Conceptual Model for Preferences of Patients and Physicians in the Utilization of Health Services in Degenerative Disease of the Lumbar Spine
CHAPTER 2: In the Eye of the Beholder: How do the Preferences of Patients, Family Physicians and Surgeons Compare in Decision Making for Lumbar Spinal Surgery?

1 Introduction

1.1 Background

Systematic literature reviews as recently as 2005 have found insufficient evidence concerning the efficacy of surgical treatment for degenerative disease of the lumbar spine (DDLS).\textsuperscript{69, 70} Despite this lack of high-quality evidence, surgeons who treat these patients believe that surgical treatment confers significant benefit over nonoperative treatment.\textsuperscript{19, 20} Surgical rates, therefore, have increased dramatically over the past two decades.\textsuperscript{16} A better appreciation of the pathoanatomy, refined diagnostic modalities, enhanced surgical techniques, and improved patient selection may have contributed to better outcomes in spinal surgery.\textsuperscript{162} Only in the past three years have randomized clinical trials confirmed that surgery may benefit selected patients with spinal stenosis and degenerative spondylolisthesis.\textsuperscript{71, 72, 75}

Since surgical consultation can only be obtained following physician referral in the Province of Ontario, the family physician (FP) typically serves as the ‘gatekeeper’ to care. There is little appreciation among referring FPs on the factors, both clinical and non-clinical, that identify an appropriate surgical candidate for discretionary surgery.\textsuperscript{163} Furthermore, the need to include patient preferences in guiding medical decision making is now widely accepted.\textsuperscript{131, 134} Patients rarely make decisions for referral and surgery entirely on their own. The FP, in partnership with the patient, often decides on the benefits of surgical referral and, in turn, the surgeon, in partnership with the patient, decides on the appropriateness of surgery. Thus, the decision to undergo surgery really is a shared process involving patient, FP, and spinal surgeon.

The \textit{shared decision making} process incorporates both patient and physician preferences into the decision making, with the goal of reaching an appropriate decision.\textsuperscript{134, 164, 165} Patients and physicians, however, may have different preferences for care.\textsuperscript{140} Differences in preferences between patients and physicians have the potential to lead to wide variation in referrals and impede the shared decision making process. Greater agreement in patient-physician decisions is associated
with improved patient satisfaction and overall health status. Thus, aligning preferences for treatment among patients and physicians may not only streamline the shared decision making process by reducing variability in decisions, but may even result in improved outcomes.

The specific preferences of surgeons, FPs, and patients in surgical decisions for the degenerative lumbar spine are poorly understood. To the best of our knowledge, there exist no studies to assess preferences and quantify the importance of different clinical factors for surgery for DDLS. Conjoint Analysis (CA) is one such method for eliciting preferences that is being used more commonly in healthcare to understand how different factors influence patients’ decisions for the provision of health services. CA thus provides an ideal means to identify the dominant clinical factors influencing preferences for lumbar spinal surgery.

1.2 Specific Objective

The objective of this study was to determine the relative importance that surgeons, FPs, and patients place on presenting clinical symptoms when considering surgical referral and surgery for DDLS. We further aimed to investigate if other non-clinical factors such as physician characteristics or patient demographics are associated with these preferences.

2 Methods

2.1 Conjoint Analysis

Conjoint analysis (CA) is founded on the principle that any health state (or service/good) can be deconstructed by its clinical factors (attributes) and that the extent an individual values a certain action or decision depends on the nature and level of these clinical factors. This technique involves presenting multiple hypothetical clinical vignettes (scenarios) comprised of different levels of clinical factors (e.g. varying severities of clinical symptoms) and asking respondents to rate (on a fixed scale), rank (in order of preference), or make discrete choices between vignettes (e.g. choose A over B). Once the data have been gathered, regression techniques are used to establish the average relative importance of each clinical factor and determine how people are willing to trade between clinical factors. In healthcare, discrete choice responses have been used most commonly, however, rating methods are being used more frequently.

CA permits the measurement of the relative importance of each clinical factor (i.e. patient symptom) which simple explicit rating exercises cannot. In general, questionnaires can be
designed either to elicit explicit preferences on individual clinical factors of interest or preferences for treatment of hypothetical clinical vignettes with varying implicit clinical factors. The major advantage of the latter strategy is that clinical vignettes take a holistic approach, emphasizing the importance of the whole rather than the individual components thus simulating real scenarios better than explicit clinical factors. Furthermore, vignettes have been shown to be a valid method of measuring quality of care and have even been used to study physician recommendations for treatment in chronic low back pain.\textsuperscript{107, 171} The primary outcome measure of CA is the average relative importance that participants attach to each clinical factor, that is, which patient symptoms are most important in making the decision to proceed with surgical referral or surgery. Typically, CA is carried out using a random-effects probit regression model to calculate coefficients that estimate the value attached to the clinical factors that determine preferences. The probit function, similar to the logit function used in logistic regression, is a member of the general linear model family used to model binary or ordinal variables. Random effects are useful in studies with repeated measures (multiple observations per individual) because they allow for the measurement of the amount of variation between individuals as well as within individuals and permit each individual respondent to vary in their mean overall response.

2.2 Questionnaire Development

Our questionnaire was developed as a CA experiment using clinical vignettes in order to measure implicit preferences of patients and physicians.\textsuperscript{171} In addition to clinical vignettes, respondents were asked several questions pertaining to the referral process and surgery as well as a reporting of their demographics (age, sex) and professional characteristics for physicians (years in practice, university affiliation, medical training location, patient volume, and sub-specialty training in spine for surgeons).

In general, the four main steps in conducting a CA study are to (1) identify the characteristics, (2) assign levels to the characteristics, (3) choose the scenarios, and (4) establish preferences. Statistical analysis then follows the data collection.\textsuperscript{166}

2.2.1 Step 1: Defining the clinical factors

We first identified the relevant clinical factors that were used in the clinical vignettes by literature review. We identified a previously published systematic review of preoperative predictors for outcome in patients undergoing surgery for lumbar spinal stenosis.\textsuperscript{116} Twenty-one
relevant articles were retrieved in full. In order to maintain a similar questionnaire for patients and physicians, we selected all listed clinical factors from these articles that represented symptoms describable by patients, not requiring clinical or radiographic examination. Sixteen clinical factors were identified and presented to a pilot group of five surgeons, five FPs, and ten patients. The pilot group individually ranked all clinical factors from highest to lowest importance in influencing their decision towards surgery. Since it has been recommended that CA studies contain no more than six factors, after combining the group rankings, we selected the six most important clinical factors to use in our clinical vignettes (Table 2-1).172,173

### 2.2.2 Step 2: Assigning factor levels

In CA studies, each clinical factor may have multiple levels of severity or categories. That is, clinical symptoms may be mild, moderate, or severe. With unequal numbers of levels for each factor, an artificially higher importance may be attributed to a factor with more levels compared with fewer levels.172 In order to avoid this biasing effect, we chose to constrain our design to two levels for each factor. The specific levels were determined using the information from the literature review, the pilot survey, and an understanding of the variation in levels that would be plausible yet still actionable. We used effects coding in our design to standardize the parameters, whereby the severe level was coded as +1 and mild as -1. This strategy was employed to ensure that the regression parameters for each clinical factor are scaled in the same way and therefore are directly comparable to calculate relative importances.174 Clinical factors and associated levels are shown in Table 2-1. Since we were interested in preferences for discretionary surgery and not knowledge of absolute indications (i.e. for cauda equina syndrome), we limited the neurological symptoms to include ‘some motor weakness’ as the most severe level.

### 2.2.3 Step 3: Creating the vignettes

We next selected the vignettes to be presented to the participants. With six clinical factors and two levels each, a full-factorial model would yield a total of 64 different vignettes, which would place too large a burden on each participant to complete. In order to reduce the number of vignettes to a manageable size, we relied on a one-fourth fractional factorial experimental design yielding a total of 16 vignettes. In doing so, our goal was to maintain the overall efficiency of the design by ensuring the absence of colinearity between factors and maintaining balance of the factor levels and orthogonality to the design. Using the SAS Macro MktEx175 we created an
experimental design of 16 vignettes (Table 2-2) which included two vignettes for validation (best case scenario and worst case scenario). The results of this experimental design yielded a balanced, orthogonal set of vignettes with D-, A-, and G-efficiencies all of 100%, indicating goodness of the design relative to other hypothetical orthogonal designs. Furthermore, all vignettes were considered clinically sensible and plausible. We further pilot tested the vignettes with a group of five junior orthopaedic residents.

2.2.4 Step 4: Obtaining the preference data

In CA studies, responses may take the form of ratings, rankings, or even discrete choice. Most applications of CA in health care have used discrete choice methods, whereby each respondent has to choose between two profiles (vignettes) in order to simulate real-life decisions of one option over another. Ranking responses require respondents to consider all vignettes and rank order them relative to each other. This requires one to simultaneously consider all vignettes at the same time and always forces a choice between vignettes. Rating responses allows a respondent to rate different vignettes using the same scale such that two or more vignettes could be assigned the same response. Since our study attempts to measure preferences, rather than choices or behaviour, we chose a ratings outcome. Furthermore, CA rating studies have been shown to have high levels of consistency, reliability, and internal validity. This technique has also been shown to offer higher information efficiency compared with the ranking and choice-based techniques, likely because respondents are not artificially forced into a decision but rather can respond equivalently between two different scenarios. And, from a pragmatic perspective, the rating technique is also easier for elderly respondents to understand the tasks rather than choice-based methods.

From our mailed surveys, we established rating preferences from all of our three groups (surgeons, FPs, and patients). We determined the preference for surgical referral and for surgery for each of the clinical vignettes on a scale from 1 to 6. Questionnaires to surgeons, FPs, and patients are given in Appendix A.
2.3 Study Population

We used a mailed survey to measure preferences of surgeons, FPs, and patients. All mailing addresses of physicians practicing in Ontario were acquired from the College of Physicians and Surgeons of Ontario (CPSO).

2.3.1 Surgeons

In order to limit our survey to orthopaedic and neurosurgeons who manage adult spinal patients, we excluded those surgeons whose primary address was listed at paediatric hospitals or who were still in training. We mailed surveys to the remaining 519 orthopaedic and neurosurgeons in Ontario. With the assistance of a university-based organized survey research unit, we used a four-step mailing technique involving (1) an introductory letter, (2) the complete mail package, (3) a reminder/thank-you card, and (4) a second package for non-responders. Our introductory survey question for surgeons was aimed at identifying only those who manage patients (operatively or non-operatively) with DDLS in order to best reflect current practice. Only surgeons who actively see patients with these conditions were considered eligible for completion of our survey and those ineligible were asked to return the remainder of the questionnaire blank.

In order to ensure that our survey captured the majority of surgeons treating these patients across Ontario, we validated our sample using hospital discharge data for surgical treatment of DDLS from Ontario, 2002-6 obtained from the Canadian Institute for Health Information (details provided in Chapter 4). To distinguish counties with treating surgeons from those without treating surgeons, we determined the local treatment rate by calculating the proportion of patients residing in any county that underwent surgery in that county. This local treatment rate was then used to verify whether our surgeon response reflected actual practice patterns.

2.3.2 Family Physicians

FPs were surveyed with a stratified random sampling design in order to ensure we captured sufficient FP responses from all regions for the third objective of this thesis (Chapter 4). We first stratified our sampling initially by the 14 Ontario Local Health Integration Networks (LHINs) but then due to low response rates in rural areas, we repeated our mailing stratified by the 49 counties. Because we planned to sample patients via their FPs, we required a FP sampling frame that best reflected physicians managing a wide variety of patients and having a regular patient
population. For these reasons, we chose to exclude FPs who (1) were not designated as ‘Family Physician’ in the CPSO listing, (2) had primary practice addresses in universities, government agencies, hospitals, specialty or walk-in clinics, (3) graduated from medical school earlier than 1961 or later than 2002, and (4) had postal codes that were unable to be mapped to a specific Ontario county. In counties with fewer than 10 FPs left in the sampling frame, we merged that county with a neighbouring county with the fewest FPs in the sampling frame. We then stratified by 43 of 49 Ontario counties that contained 10 or more FPs in our sampling frame and conducted a similar four-step mailing technique. We subsequently followed-up with telephone calls to low response counties in our second round.

**Sample Size Calculation for CA**

For the purposes of CA, an estimate of the required sample size was proposed by Orme (2006) such that the product of the sample size \(n\), the number of tasks or vignettes \(t\), and the number of alternatives/choices per task \(a\) divided by the largest number of levels \(c\) should be greater than 500 (or preferably 1000)

\[
\frac{nta}{c} \geq 500 \text{ (or 1000)}
\]

In our case, using a ratings-based CA, there is a single numeric choice such that \(a=1\). Therefore, for our design with \(t=16\) and \(c=2\), and using the more conservative cutoff of 1000, our sample size estimate was 125 respondents

\[
\left( \frac{n = \frac{1000 \times c}{t \times a}}{16 \times 1} = \frac{1000 \times 2}{16} = 125 \right).
\]

Since we also planned to survey patients via their FPs, we overestimated our required sample size based on a response rate of 20% in a prior similar Ontario survey.181

2.3.3 Patients

In order to capture patient data at the primary care level and establish a direct link between FP and patient for the second objective of this thesis (Chapter 3), we asked FPs to request their own patients complete questionnaire packages.182 Each FP was supplied with five patient questionnaires and we aimed to obtain at least 2 patients per FP via a convenience sample of regularly scheduled prospective patients.181

Patients eligible for survey participation were those ages 50 and over with any history or current complaint of at least two months duration of back or leg pain (sciatica) likely consistent with
DDLS (with or without previous imaging, referral, or surgery). The rationale for including patients with symptoms consistent with DDLS is that a clear association exists between patients’ current health state and how they value hypothetical health states. Because of concerns about low response in patients, we further planned to over sample FPs (1000 FPs to capture at least 150 patients and 150 FPs across 43 of the 49 counties).

2.4 Statistical Analyses

Analysis of the conjoint data consists of two main components. The first component is an analysis of the validity of the experiment. Once shown to be valid, the second component is the analysis of the vignette ratings.

Since FPs were sampled using a stratified sampling design, we weighted each physician by the ratio of the number of total physicians divided by the number of responding physicians for the county relative to the entire province in order to reflect the preferences of Ontario FPs. Since patients were surveyed via their FPs, they cannot be considered independent but rather clustered within their FP’s practice. We therefore used a multilevel modeling approach to account for this clustering. Patient clusters were weighted in a manner similar to their FPs, since they were collected via the FPs stratified sample.

In order to determine whether differences existed in overall preference for surgical referral or surgery, we first analyzed all responses using a repeated measures linear regression. We further dichotomized responses into preference for (response ≤ 3) or against (response > 3) surgical referral and surgery and analyzed overall preferences using repeated measures logistic regression. The details of the conjoint validation and the formal conjoint analysis regression are provided below.

2.4.1 Conjoint Validation

The validity of the model was tested within the framework of CA in two ways. Internal (theoretical) validity was tested by exploring model consistency with a priori expectations. That is, belief that one or more of the clinical factors presented in the scenarios have strong directions for preference was tested by examining the direction and significance of the coefficient of that clinical factor. In the design of our analysis, none of the six clinical factors should have a
significant negative coefficient if the clinical symptoms are interpreted correctly, since our effects coding (-1 for mild level and +1 for severe level) ensured that the more severe level, with higher probability of requiring referral or surgery, was greater than the milder level. Since patients and FPs are not spinal specialists, this assumption may not necessarily hold true for all clinical factors in these groups. For example, the ‘severe level’ for the dominant location of pain (i.e. leg-dominant) or the typical onset of pain (i.e. at rest) may not be correctly interpreted by patients. For surgeons, however, we assumed that this was an acceptable validation of the experiment.

Second, internal consistency was assessed by determining the proportion of vignettes rated higher than another when they differ by one having more ‘severe’ levels than the other. In these vignette pairs, all attributes are the same with the exception of one or two considered, \textit{a priori}, to be more preferable than in the other vignette. In the design of our conjoint experiment we included two extreme scenarios (best case and worst case). Our first test of internal consistency was that the most extreme rated higher than the least extreme. From Table 2-2, we see that Vignette 9 is the most severe and Vignette 3 is the least severe (Test 1). Since the severity of the levels for all the clinical factors (such as location of pain) may not be intuitive for patients or all physicians, we chose to use two vignettes that were different only in their severity of pain and walking tolerance – two clinical factors whereby the levels can be easily interpreted by both patients and physicians as better (mild/moderate pain; able to walk at least 6 blocks) or worse (severe pain; unable to walk more than 2 blocks). Vignettes 5 and 16 differed only in walking tolerance and severity of pain (Test 2), with Vignette 16 having both severe levels (severe pain, <2 blocks). Therefore, to measure internal consistency, we looked at the proportion of respondents who rated Vignette 9 higher than 3, and Vignette 16 higher than 5.

Furthermore, since CA assumes that a respondent’s preference (or utility) is continuous in which one severe clinical factor can compensate for another mild clinical factor, it implies that respondents are willing to ‘trade’ between clinical factors. Dominant preferences occur when respondents are not willing to trade between clinical factors and the presence of the severe or mild level of one clinical factor always determines their preference, regardless of the other factors. There currently exists no consensus on how best to handle these dominant responders. In our analysis, we first eliminated all responders who reported the same response for all vignettes because the absence of any variation precludes any statistical inference.
We further analyzed the data with and without removing ‘dominant’ and internally inconsistent responders.

2.4.2 Analysis of the Vignette Ratings

For a more detailed description of the analysis of conjoint data, the reader is referred to Hensher, Rose, and Green (2005).\(^{174}\) For the purpose of this dissertation, however, we provide a brief overview of the statistical theory behind it. In order to relate responses to preferences, CA makes the assumption that the observed rating response, \(R_{ij}\), for the \(i^{th}\) vignette and the \(j^{th}\) respondent is generated by a function of an unobserved utility (or preference) \(U_{ij}\).\(^{190}\) This utility is further assumed to be a linear additive function of the six clinical parameters \(X_{1i}\) to \(X_{6i}\) for the \(i^{th}\) vignette (Equation [1]).

\[
R_{ij} = F(U_{ij}) \quad \text{where} \quad U_{ij} = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_6 X_{6i} + \nu_j + \epsilon_{ij} \tag{1}
\]

and \(\nu_j \sim \text{Normal} \{0, \sigma^2\}\), \(\epsilon_{ij} \sim \text{Normal} \{0, \sigma^2\}\), \(\text{Variance}\{\nu_j + \epsilon_{ij}\} = \sigma^2 + \sigma^2\)\(^{190}\)

In order to account for the potential heterogeneity between respondents, we added a random intercept term \(\nu_j\) for the \(j^{th}\) respondent (random-effects model) such that each respondent’s utility can have a varying baseline with a measurable between-respondent variance \((\sigma^2)\).

Since our responses were collected on an ordinal scale we used the proportional odds model for ordinal data.\(^{191}\) This model assumes that the responses \((R_{ij})\), given on an ordinal scale from 1 to 6, for the \(j^{th}\) respondent and the \(i^{th}\) vignette are given by the utilities \((U_{ij})\) with the different thresholds representing cutoffs for the ordinal responses:

\[
R_{ij} = \begin{cases} 
1 & \text{if } U_{ij} \leq \tau_1 \quad \text{with} \quad \Pr = \Phi(\tau_1 - U_{ij}) \\
\text{m if } \tau_{m-1} < U_{ij} < \tau_m \quad \text{with} \quad \Pr = \Phi(\tau_m - U_{ij}) - \Phi(\tau_{m-1} - U_{ij}) \\
6 & \text{if } U_{ij} \geq \tau_5 \quad \text{with} \quad \Pr = 1 - \Phi(\tau_5 - U_{ij})
\end{cases} \tag{2}
\]

for \(m = 2\ldots5\), thresholds \(\tau_1\ldots\tau_5\) where \(\tau_0 = -\infty\), \(\tau_1 = 0\), \(\tau_6 = \infty\), and \(\Phi\) representing the standard normal cumulative density function. In this model, Equation [2] takes the operationalized form of ordered probit regression with random effects, which has been advocated for use in CA to best account for multiple responses per respondent.\(^{192,193}\)

35
Using the above model, we performed a random-effects ordered probit regression using the SAS Procedure GLIMMIX.\textsuperscript{194,195} Using this regression model, we determined the magnitude of the coefficients in order to calculate the ‘relative importance’ of each clinical factor for each group. Since all factors were coded in a similar way, the regression coefficients are directly comparable. In our design, relative importance for any given clinical factor is defined as the ratio of the magnitude of that coefficient to the sum of all coefficients. Results were compared across all three groups. We analyzed the conjoint data using multiple regression techniques with the demographics and professional/clinical characteristics of each respondent captured from their surveys to determine if these factors had a significant influence on preferences for referral and surgery.

Overall goodness-of-fit statistics (McFadden’s R-square and Chi-square likelihood ratio [LR] test) were performed based on likelihoods calculated from the equivalent fixed-effects model but regression diagnostics were performed on the random-effects model. All statistical analyses were carried out using SAS v9.1 (Cary, NC).

3 Results

3.1 Survey Response and Participant Characteristics

Responses were obtained from 131 of 302 potentially eligible surgeons (43%) across 24 of the 49 Ontario counties. Only three counties had surgery performed locally that were missing surgeon responses. For patients who reside in the county, the ratio of the number of patients operated on in that county relative to the total number of patients operated from that county was 1 in 536 (0.02%), 58 in 427 (13.6%), and 104 in 1666 (6.2%) for the three counties without surgeon response, demonstrating the small number of patients treated in counties with no surgeon response (Figure 2-1).

For FPs, we initially obtained only 105 responses from 915 potentially eligible physicians (11%) and 61 patient responses from 33 FPs. We therefore sent out a second wave of surveys. In total, over both waves, we obtained responses from 202 of 1694 potentially eligible FPs (12%) and 164 patients from 87 FPs (mean 1.9 patients/FP). From our sampling of 43 counties with more than 10 FPs in our sampling frame, we gathered responses across 45 counties in total. There were no significant differences in sex, years in practice, or location between FP responders and...
non-responders. The characteristics of surgeons and FPs are shown in Table 2-3 and of patients in Table 2-4.

As an overall measure of the value in exploring preferences for referral and surgery in DDLS, all respondents were asked to rate how straightforward they felt that decisions on surgical referral and surgery are for patients with these conditions. Although the majority of surgeons and patients felt that decisions surrounding referral and surgery were straightforward (range 51.7% to 67.6%), less than half (44.5%) of FPs felt that decisions around surgical referral were straightforward and less than one-quarter of FPs (22.8%) felt that deciding on surgery was straightforward. Ninety-one patients (57.6%) expressed a preference for seeing a surgeon for their condition compared with all other specialists.

3.2 Conjoint Analysis

Validation results of our model are shown in Table 2-5. Although physicians generally understood the task well, a large proportion of patients (19.8% and 21.6%) reported the same response for all scenarios for referral and surgery, respectively. All of these respondents were subsequently excluded from the analysis. Comparing patients who responded with the same response for referral with the rest of the cohort demonstrated that they were older by 7.6 years (p<0.02) and reported more severe pain (74% vs. 53%, p<0.03) but otherwise had no significant differences with respect to sex, education level, location of residence, previous utilization of surgical services, and other clinical symptoms. No differences were observed for patients who reported the same response preference for surgery with the rest of the cohort.

Rates of internal inconsistency as assessed by our two tests remained low (≤6%). Dominant preferences were observed in all three groups. FPs and patients had rates up to 1% (2 respondents), however, six surgeons (5%) demonstrated a dominant preference for the location of pain (leg pain over back pain). We analyzed the response data with and without internally inconsistent and dominant respondents and found no significant differences between models.

3.2.1 Surgical Referral

Responses for the preferences for surgical referral for the 16 vignettes varied between surgeons, FPs, and patients. Median and quartile responses are shown in Figure 2-2. Least squares means
for surgical referral response was 4.25 (95% confidence interval [CI] 4.14 to 4.36), 4.35 (95% CI 4.31 to 4.39), and 4.11 (95% CI 4.05 to 4.18) for surgeons, FPs, and patients, respectively. Patients had a significantly lower mean response compared with surgeons (p<0.04) and FPs (p<0.0001) for surgical referral preference but no statistical differences between physician groups were found. The odds of preferring surgical referral relative to surgeons was 1.35 (95% CI 1.05 to 1.75, p<0.02) and 0.71 (95% CI 0.49 to 1.04, p<0.08) for FPs and patients, respectively.

Results from our random-effects probit regression models for surgical referral are shown in Table 2-6. Significant between-subject variance was observed in all three groups, indicating that significant heterogeneity existed between respondents within a group and that the chosen random-effects model was appropriate. All six clinical factors in the CA were found to be highly significant for all respondents (p<0.0002). With the exception of patients’ interpretation of the location of pain, all clinical factors were found to have positive direction (internal validity). Magnitudes of each of these factors, however, varied between surgeons, FPs, and patients.

Of the non-clinical factors, fewer surgeon years in practice was significantly associated with a higher preference for surgical referral (p<0.0008). Patients with at least post-secondary education and those who have had prior surgical consultation were both associated with a higher preference for surgical referral (p<0.02 and p<0.0005, respectively). No FP characteristics were significantly associated with preference for surgical referral.

The relative importance of these clinical factors in influencing preferences for surgical referral is shown in Figure 2-3. Surgeons placed the highest importance (33%) on the dominant location of pain (leg vs. back) compared with FPs and patients who both considered this factor to be of little importance (<10%). Surgeons also considered the severity of pain (20%) and walking tolerance (18%) to be secondarily important. FPs did not identify a single dominant factor, but rather considered walking tolerance (23%), severity of pain (22%), neurological symptoms (19%), and the onset of pain (19%) to all be of similar importance. Patients, on the other hand, considered the duration of pain to be the most important factor (32%). Walking tolerance (26%) and the severity of pain (22%) were considered of secondary importance to patients.
Responses for surgical preferences for the 16 vignettes varied between surgeons, FPs, and patients. Median and quartile responses are shown in Figure 2-4. Least squares means for surgery preference was 3.57 (95% confidence interval [CI] 3.45 to 3.68), 3.89 (95% CI 3.85 to 3.93), and 3.77 (95% CI 3.71 to 3.83) for surgeons, FPs, and patients, respectively. Both FPs and patients had a significantly higher mean response (p<0.0001 and p<0.002, respectively) compared with surgeons for surgery preference. The odds of preferring surgery for FPs relative to surgeons was 1.63 (95% CI 1.29 to 2.06, p<0.0006). No difference in the relative odds of preferring surgery was found for patients compared with surgeons.

Table 2-7 shows the results of the random-effects probit regression model for recommendation and consideration of surgery for surgeons, FPs, and patients. Significant between-respondent variation was again observed, indicating that the random-effects assumption was appropriate. Similar to surgical referral, all of the clinical factor regression coefficients except for patients’ interpretation of the dominant location of pain were positive and significant (internal validity).

Orthopaedic surgical specialty was statistically associated with a lower preference for recommending surgery compared with neurosurgical specialty (p<0.047). Older patient age (p<0.03) and previous surgical consultation (p<0.03) were both associated with a greater patient preference for considering surgical treatment. No FP characteristics were found to have any significant association with preference for surgery.

The relative importance of each clinical factor in deciding for surgery is shown in Figure 2-5. Surgeons again placed the highest importance on the location of pain (34%), followed secondarily by the severity of pain (19%) and walking tolerance (19%). FPs considered neurological symptoms (23%), walking tolerance (20%), severity of pain (20%), and onset of pain (16%) to be of similar importance without a single dominant factor. The severity of pain (29%), walking tolerance (29%), and the duration of pain (28%) were jointly considered to be the most important factors for patients in deciding for surgery.
4 Discussion

In this study, we aimed to explore the preferences of patients and physicians around decisions for the surgical treatment of these degenerative spinal conditions. We found that there was wide variation in how patients, FPs, and surgeons value different clinical factors in deciding for surgical referral and surgery. The majority of surveyed FPs felt that decisions around surgical referral and surgery in particular, were not straightforward. The majority of patients expressed a preference for seeing a surgeon for their condition over a non-surgical specialist. FPs expressed the highest preference for surgical referral and surgery. Patients had lower preference for surgical referral but higher preference for surgery compared with surgeons. Surgeons placed high importance on the dominant location of pain (leg-dominant over back-dominant) while patients placed high importance on quality of life symptoms such as pain severity, duration of pain, and walking tolerance. These large differences in preferences between patients and physicians may impede the shared decision making process since each stakeholder, namely patient, FP, and surgeon, considers different clinical factors to be more or less important in the overall decisions.

The recent Spine Patient Outcomes Research Trials (SPORT) for spinal stenosis and degenerative spondylolisthesis have demonstrated the benefit of surgical treatment of these degenerative spinal conditions.71, 72 The SPORT investigators conducted preference trials with both a randomized and observational cohort assessing the effectiveness of surgical treatment for degenerative spondylolisthesis and spinal stenosis. In their trial for degenerative spondylolisthesis, the authors found that there was tremendous cross-over (approximately 40% in each direction) within the randomized cohort, illustrating the effect of strong patient preferences and thus questioning the value of a typical intention-to-treat analysis.71 As expected, the intention-to-treat analysis failed to demonstrate any difference, however, the non-randomized as-treated comparison with careful control for confounders found a substantially greater improvement for surgical treatment (most commonly decompression and instrumented fusion) over nonsurgical treatment. In their trial for spinal stenosis, the authors found a similarly high rate of cross-over to surgical treatment (43%), however, the intention-to-treat (as well as the as-treated) analysis found a significant benefit for surgery (i.e. decompression) over nonsurgical treatment.72 Despite these recent findings confirming what practicing spinal surgeons have long felt, misperceptions about the benefits of surgery are widespread.
Inappropriate referrals to orthopaedic surgeons ranged from 25 to 43%.\textsuperscript{110-112} For back pain surgical referrals, Oldmeadow and co-workers (2007) found that 70% of referrals were considered more appropriate for non-surgical management by orthopaedic surgeons.\textsuperscript{113} Conversely, Schroth and colleagues (1992) found that only 14% of surgical referrals for back pain, according to the best available evidence, were clearly inappropriate, but that 47% of patients with potential indications for surgery were not referred indicating that problems exist in both over- and under-referral of these patients.\textsuperscript{114}

In a systematic review of the literature, Montgomery and Fahey (2001) also noted that different treatment preferences exist between patients and physicians over a wide range of clinical conditions. They concluded that the magnitude and direction of these differences are not consistent and that ultimately in the shared decision making process, the extent of patient participation depends on their desire for involvement and the clinical decision at hand.\textsuperscript{140}

The technique of CA to elicit preferences has been used increasingly in health service research. Variation in preferences between patients and physicians has previously been demonstrated using CA in exploring choice of dental materials, prenatal screening for Down syndrome, management of postoperative nausea, and choice of treatment products in hemophilia, to name a few.\textsuperscript{196-199} In orthopaedics, this technique has been used to determine patient preferences for the treatment of adolescent idiopathic scoliosis.\textsuperscript{200}

In our study, surgeons, on average, placed high importance on the dominant location of pain (i.e. leg-dominant over back-dominant) in their decisions for surgical referral and surgery. Surgical outcomes for conditions with nerve root involvement, such as spinal stenosis, over back-dominant conditions, such as degenerative disc disease or facet arthrosis, are significantly better.\textsuperscript{201} Surgeons want the best outcomes for their patients, and therefore, knowingly or unknowingly, may prefer or ‘select’ those who they feel will benefit most, namely, those with leg-dominant symptoms. These distinctions may not necessarily be well-known to FPs or patients and therefore may not be appreciated as a significant influence on the decision to operate.

Irwin and colleagues (2005) found that surgeon-specific characteristics, such as training background (specialty, fellowship training) and years in practice had a profound influence on the choice of surgical procedure (i.e. fusion vs. decompression, the use of instrumentation, and
anterior vs. posterior approach) for a variety of degenerative lumbar spinal conditions. The authors found that younger surgeons preferred instrumentation and orthopaedic surgeons recommended fusion and instrumentation over neurosurgeons. Similarly, in this study, we found that surgical referral was preferred among younger surgeons over older surgeons and that preference for surgical treatment was higher for neurosurgeons over orthopaedic surgeons. In our sample, neurosurgeons performed an average of 2 more cases per month than their orthopaedic colleagues (p<0.02) although there were no significant differences in the number of patients seen in consultation. This attests to the reality that although many orthopaedic surgeons, like neurosurgeons, consult on these patients, more spinal surgery is being performed in a general neurosurgical practice rather than a general orthopaedic practice. This is not surprising since Dvorak and co-authors (2006) found that neurosurgical residents spend more than twice the amount of time in spine training than orthopaedic residents. The implication of this finding is that the type of specialist seeing a patient may influence the probability of a recommendation for surgical (or nonsurgical) treatment.

FPs had the highest overall preference for surgery and, on average, considered neurological symptoms to be the most important clinical factor, only slightly more important than pain severity, and walking tolerance. We speculate that a worsening neurological condition or the presence of motor weakness may signal a more urgent need for surgical intervention from the primary care perspective regardless of the prognosis for recovery or the degree of improvement. No specific physician characteristics were associated with higher preferences for surgery. On the other hand, however, patients were found to have higher preferences for duration of pain, walking tolerance, and severity of pain. All of these symptoms are highly related to quality of life and have little direct bearing on outcomes following surgery. Not surprisingly, the dominant location of pain, for example, does not play a major role in their decision to proceed with surgery. The longer and more severe their pain and the greater their disability (walking tolerance) the more patients would consider surgery. Patients who have had more education and previous consultations had higher preferences for referral and surgery, likely because they are better informed about potential outcomes.

Shared decision making is a process in which patient and physician reach mutual agreement after considering all of the information with respect to treatment options and outcomes. The goal of this process is to come to a decision that integrates the preferences of patients with the
recommendations of physicians. The main advantages are that it allows for better information gathering for both patient and physician and it forces the physician to present all treatment options. Frosch and Kaplan (1999) reviewed the literature on shared decision making and noted several studies in which patients who engaged in shared decision making had a greater sense of control, less concern over their disease, more satisfaction with treatment, and even improved clinical outcomes. Decision supports have been developed to facilitate this process which have resulted in better aligning patient and physician preferences. Deyo and co-authors (2000), in a randomized trial, found that a video program on back surgery enhanced the decision making process for patients with spinal stenosis and even resulted in more patients opting for surgery and feeling better informed about their condition. Despite these attempts at improving the shared decision making process, there is some evidence to suggest that patients may prefer decisions around lumbar spinal surgery to be left up to their surgeons. Weiner and Essis (2006) found that spine surgery patients preferred to defer decision making, such as basic treatment decisions as well as technical decisions, to their surgeons.

This study has several limitations. First, we had relatively low response rates from our survey. We obtained responses from 43% of practicing orthopaedic and neurosurgeons across Ontario. Although this number is in keeping with several other surveys of surgeons in Ontario, we believe that it is an underestimation since the number of eligible surgeons (those who see patients with DDLS in consultation) is likely lower than that of the non-responders. From a population-based study of spinal surgery in Ontario, Bederman et al. (2009) identified 147 surgeons performing spinal surgery. Although our inclusion criteria were surgeons who see these patients in consultation (and do not necessarily operate on them), we believe that our sample of 131 surgeons captured the majority of those treating DDLS patients. Response rates from FPs, on the other hand, were significantly lower (12%). In order to obtain enough physicians, this necessitated a second round of mailings. Although we obtained results from 202 physicians, well above our specified number, we may not be able to generalize to all FPs across the province. Nonetheless, we are reassured by another similar survey capturing patient data via physicians with low response. Furthermore, we observed no significant differences between responders and non-responders in sex, years in practice, and location of practice. FPs who responded to our survey may, however, be different from those that did not respond in that they may have more clinical interest in DDLS than non-responders. We would therefore speculate
that responders would have more familiarity with the management of patients with DDLS than non-responders and that the exclusion of non-responders may have resulted in more variability and greater disagreement with surgeons in preferences for surgical decision making. Our multivariate analysis failed to identify any significant physician characteristics that were associated with higher preference for surgery or surgical referral. The inclusion of a broader sample of FPs would likely not have provided us with any different results. Patients were sampled purposefully through their FPs, and therefore may not represent the true preferences of patients across Ontario. There is a trade-off between sampling patients that have undergone surgery (i.e. from a surgical database) and general patients with no experience with the disease state. Patients having undergone surgery would likely have very different opinions about decisions for surgery while asymptomatic patients would have a lesser understanding of the degree of pain and disability that patients who suffer from DDLS possess. Therefore, our sample was selected to represent the full spectrum of disease and balance the experiences between those who have had surgery with those without symptoms and thus reported a wide range of clinical severities and experiences with the health care system.

Second, eliciting preferences based on clinical vignettes may not determine the true preferences physicians have for real clinical encounters. However, using holistic vignettes with implicit clinical factors instead of explicit lists of clinical parameters can simulate real clinical scenarios more accurately.\textsuperscript{107, 171, 207} We recognize that this technique comes at the expense of a limited number of clinical factors with which to test preferences (six factors), but because this is generally agreed upon as the number of important factors clinically, it is a practical limit to the number of factors respondents can consider in a vignette. Asking patients to give their preferences for treatment based on theoretical clinical vignettes rather that their actual situation does not necessarily provide a true account of their attitudes in reality. However, it would be extremely difficult to obtain enough patients that fit the same exact clinical scenarios presented to physicians. Alternatively, surveying patients without any clinical symptoms would be easier but less valuable since they have not experienced the disease state.\textsuperscript{183}

Third, CA is a relatively new technique and may require further work to establish it as a rigorous methodological tool.\textsuperscript{169} Despite these methodological concerns, CA has been highly recommended as a quantitative technique to elicit preferences.\textsuperscript{177} Results from our analysis demonstrated generally high rates of internal consistency and internal validity. A relatively high
number of surgeons exhibited dominant preferences for the location of pain (6%), however, this finding is consistent with the high importance found in the overall analysis and clinical experience. Nonetheless, no significant differences were observed in our analysis with dominant responders removed. One major concern was that approximately 20% of patients reported the same response for all vignettes presented. This might be explained by either a frustration with the exercise or rather that these patients failed to understand the task. With the exception of older age and more severe pain, no other differences were noted between patient respondents suggesting this problem may be the latter. Using a ranking or choice-based technique would doubtlessly have resulted in higher rates of task misunderstanding. Furthermore, we used a linear additive model although CA permits the use of interaction terms. Since we limited the number of vignettes to 16 to decrease responder burden, we were limited in how many parameters we were able to estimate. We therefore did not estimate the effect of interaction terms in our model.

Fourth, it is important to note that differences may exist between individual’s preferences and average preferences over a population. While our random-effects model allowed for variation in individual intercept terms (and therefore overall preference), without including random-effects coefficients for each of the clinical factors, we did not model individual preferences of respondents for each factor. Rather, we determined the average preferences for each factor over each group. Although this represents a means of interpreting the data and not truly a study limitation, it is worth mentioning. The implication of this strategy is that although individual preferences are responsible for individual decision making, and therefore referral practices and surgical rates, average preferences may be more informative for the population as a whole.

Finally, this survey was confined to clinical practice in Ontario and the results may not necessarily generalize to other Canadian provinces, or to other areas outside Canada. However, our results are worth considering in any health system whereby referring physicians play an intermediary role with patients and surgeons regarding decisions about referral and treatment. Even in jurisdictions with self-referral, surgeon understanding of patient preferences may facilitate the shared decision making process.

CA provides a useful technique to elicit preferences from patients and physicians for decisions around surgical treatment for DDLS. Although obtaining responses from patients via their FPs
was used to directly link patients to FPs, other strategies to survey patients and physicians may have improved our response rates. Furthermore, direct patient contact and patient piloting of vignettes may have allowed us to reduce the number of all same response data.

Understanding preferences for surgery on the degenerative lumbar spine among patients, FPs, and surgeons and where significant sources of variation and disagreement exist, can have several beneficial effects. Firstly, aligning opinions of patients and physicians would improve the shared decision making process itself and patients’ expectations about surgery would be more accurate. This can directly result in a significant improvement in patient satisfaction with the healthcare process and even overall health status following treatment.141 Second, with FPs having a better appreciation for symptoms related to nerve root involvement (i.e. leg-dominant pain) the referral process may be more efficient. This allows the most appropriate surgical candidates to obtain timely assessments. Finally, with FPs and surgeons realizing the importance that patients place on quality of life symptoms, such as duration, severity, and walking tolerance, they may enhance the shared decision making process and improve the care delivered to these patients.

5 Conclusions

Surgeons, FPs, and patients place differing importance on various clinical factors for surgical treatment of DDLS. Surgeons considered the location of pain to be the dominant factor determining their decision for surgery. FPs placed nearly equal weight on patients’ walking tolerance, severity of pain, presence of neurological symptoms, and the typical onset of pain when initiating surgical referral. Dominant location of pain accounted for no more than 10% of their decisions. Patients, on the other hand, placed a large emphasis on quality of life symptoms including the duration of pain, pain severity, and their walking tolerance in considering having surgery. FPs may reduce over- and under-referrals by appreciating surgeons’ importance on location of pain (leg vs. back). Surgeons and FPs may improve the shared decision making process by being mindful of patients’ importance on quality of life symptoms.
Table 2-1. Clinical Factors and Levels for Conjoint Experiment.

<table>
<thead>
<tr>
<th>Clinical Factors</th>
<th>Mild Level (-1)</th>
<th>Severe Level (+1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of Pain</td>
<td>Less than 4 months</td>
<td>More than 8 months</td>
</tr>
<tr>
<td>Severity of Pain</td>
<td>Mild/Moderate</td>
<td>Severe</td>
</tr>
<tr>
<td>Location of Pain</td>
<td>Back-dominant</td>
<td>Leg-dominant</td>
</tr>
<tr>
<td>Onset of Pain</td>
<td>Only with walking</td>
<td>At night or at rest</td>
</tr>
<tr>
<td>Neurological Symptoms</td>
<td>Mild numbness/tingling</td>
<td>Moderate leg weakness</td>
</tr>
<tr>
<td>Walking Tolerance</td>
<td>More than 6 blocks</td>
<td>Less than 2 blocks</td>
</tr>
</tbody>
</table>
Table 2-2. Table of Presented Clinical Vignettes.

<table>
<thead>
<tr>
<th>Vignette</th>
<th>Duration of Pain</th>
<th>Severity</th>
<th>Location of Pain</th>
<th>Onset of Pain</th>
<th>Neurological Symptoms</th>
<th>Walking Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8 months</td>
<td>severe</td>
<td>back</td>
<td>at rest</td>
<td>mild numbness/tingling</td>
<td>2 blocks</td>
</tr>
<tr>
<td>2</td>
<td>4 months</td>
<td>moderate</td>
<td>back</td>
<td>at rest</td>
<td>mild numbness/tingling</td>
<td>2 blocks</td>
</tr>
<tr>
<td>3</td>
<td>4 months</td>
<td>moderate</td>
<td>back</td>
<td>with walking</td>
<td>mild numbness/tingling</td>
<td>6 blocks</td>
</tr>
<tr>
<td>4</td>
<td>4 months</td>
<td>severe</td>
<td>legs</td>
<td>with walking</td>
<td>moderate weakness</td>
<td>6 blocks</td>
</tr>
<tr>
<td>5</td>
<td>8 months</td>
<td>moderate</td>
<td>legs</td>
<td>with walking</td>
<td>mild numbness/tingling</td>
<td>6 blocks</td>
</tr>
<tr>
<td>6</td>
<td>4 months</td>
<td>moderate</td>
<td>legs</td>
<td>with walking</td>
<td>moderate weakness</td>
<td>2 blocks</td>
</tr>
<tr>
<td>7</td>
<td>8 months</td>
<td>severe</td>
<td>back</td>
<td>with walking</td>
<td>moderate weakness</td>
<td>6 blocks</td>
</tr>
<tr>
<td>8</td>
<td>4 months</td>
<td>severe</td>
<td>back</td>
<td>at rest</td>
<td>moderate weakness</td>
<td>6 blocks</td>
</tr>
<tr>
<td>9</td>
<td>8 months</td>
<td>severe</td>
<td>legs</td>
<td>at rest</td>
<td>moderate weakness</td>
<td>2 blocks</td>
</tr>
<tr>
<td>10</td>
<td>4 months</td>
<td>severe</td>
<td>legs</td>
<td>at rest</td>
<td>mild numbness/tingling</td>
<td>6 blocks</td>
</tr>
<tr>
<td>11</td>
<td>4 months</td>
<td>severe</td>
<td>back</td>
<td>with walking</td>
<td>mild numbness/tingling</td>
<td>2 blocks</td>
</tr>
<tr>
<td>12</td>
<td>4 months</td>
<td>moderate</td>
<td>legs</td>
<td>at rest</td>
<td>moderate weakness</td>
<td>2 blocks</td>
</tr>
<tr>
<td>13</td>
<td>8 months</td>
<td>moderate</td>
<td>legs</td>
<td>at rest</td>
<td>mild numbness/tingling</td>
<td>6 blocks</td>
</tr>
<tr>
<td>14</td>
<td>8 months</td>
<td>moderate</td>
<td>back</td>
<td>at rest</td>
<td>moderate weakness</td>
<td>6 blocks</td>
</tr>
<tr>
<td>15</td>
<td>8 months</td>
<td>moderate</td>
<td>back</td>
<td>with walking</td>
<td>moderate weakness</td>
<td>2 blocks</td>
</tr>
<tr>
<td>16</td>
<td>8 months</td>
<td>severe</td>
<td>legs</td>
<td>with walking</td>
<td>mild numbness/tingling</td>
<td>2 blocks</td>
</tr>
</tbody>
</table>
### Table 2-3. Physician Characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Surgeons (n=131)</th>
<th>Family Physicians (n=202)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age (range)</td>
<td>52.8 (33-82)</td>
<td>50.2 (29-84)</td>
</tr>
<tr>
<td>Female (%)</td>
<td>8 (6.1%)</td>
<td>69 (34.3%)</td>
</tr>
<tr>
<td>Mean years in practice (range)</td>
<td>21.6 (1-48)</td>
<td>23.2 (2-58)</td>
</tr>
<tr>
<td>North American medical training (%)</td>
<td>110 (85.3%)</td>
<td>173 (86.5%)</td>
</tr>
<tr>
<td>University Affiliation (%)</td>
<td>78 (60.0%)</td>
<td>37 (18.4%)</td>
</tr>
<tr>
<td>Median monthly volume of patients (range)</td>
<td>10 (1-300)</td>
<td>4 (0-100)</td>
</tr>
<tr>
<td>Specialty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthopaedic Surgeons</td>
<td>100 (76.3%)</td>
<td></td>
</tr>
<tr>
<td>Neurosurgeons</td>
<td>31 (23.7%)</td>
<td></td>
</tr>
<tr>
<td>Spinal subspecialty fellowship (%)</td>
<td>43 (33.1%)</td>
<td>N/A</td>
</tr>
<tr>
<td>Median monthly volume of operations (range)</td>
<td>5 (1-20)</td>
<td>N/A</td>
</tr>
</tbody>
</table>
### Table 2-4. Patient Characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Patients (n=164)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age (range)</td>
<td>62.4 (30-88)</td>
</tr>
<tr>
<td>Female (%)</td>
<td>86 (52.8%)</td>
</tr>
<tr>
<td>Highest Level of Education</td>
<td></td>
</tr>
<tr>
<td>Did not complete High School</td>
<td>39 (24.5%)</td>
</tr>
<tr>
<td>Completed High School</td>
<td>38 (23.9%)</td>
</tr>
<tr>
<td>Post-secondary Studies</td>
<td>49 (30.8%)</td>
</tr>
<tr>
<td>University (undergraduate)</td>
<td>19 (12.0%)</td>
</tr>
<tr>
<td>University (professional or graduate)</td>
<td>14 (8.8%)</td>
</tr>
<tr>
<td>Surgical Consultation</td>
<td>96 (58.5%)</td>
</tr>
<tr>
<td>Orthopaedics</td>
<td>42 (46.2%)</td>
</tr>
<tr>
<td>Neurosurgery</td>
<td>38 (41.8%)</td>
</tr>
<tr>
<td>Both</td>
<td>11 (12.1%)</td>
</tr>
<tr>
<td>Undergone Surgery (%)</td>
<td>36 (22.0%)</td>
</tr>
</tbody>
</table>
Table 2-5. Validation of Conjoint Analysis Responses.

<table>
<thead>
<tr>
<th>Respondents Exhibiting:</th>
<th>Surgeons (n=131)</th>
<th>FP (n=202)</th>
<th>Patients (n=164)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Referral Surgery</td>
<td>Referral Surgery</td>
<td>Referral Surgery</td>
</tr>
<tr>
<td>All Same Response</td>
<td>4 (3.1%) 2 (1.5%)</td>
<td>3 (1.5%) 2 (1.0%)</td>
<td>35 (21.6%) 32 (19.8%)</td>
</tr>
<tr>
<td>Internal Inconsistency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 1*</td>
<td>1 (0.8%) 1 (0.8%)</td>
<td>2 (1.0%) 1 (0.5%)</td>
<td>4 (2.8%) 1 (0.7%)</td>
</tr>
<tr>
<td>Test 2†</td>
<td>8 (6.2%) 7 (5.4%)</td>
<td>3 (1.5%) 4 (2.0%)</td>
<td>8 (5.2%) 6 (3.9%)</td>
</tr>
<tr>
<td>Dominant Preferences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Severity</td>
<td>0</td>
<td>0</td>
<td>1 (0.5%) 0</td>
</tr>
<tr>
<td>Location</td>
<td>6 (4.6%) 6 (4.6%)</td>
<td>1 (0.5%) 0</td>
<td>1 (0.6%) 2 (1.2%)</td>
</tr>
<tr>
<td>Onset</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Neurological</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Walking</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* Extreme Scenarios
† Walking and Severity
Table 2-6. Regression Results for Preferences about Surgical Referral.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Surgeons* Estimate (SE)</th>
<th>P-value</th>
<th>FPs† Estimate (SE)</th>
<th>P-value</th>
<th>Patients‡ Estimate (SE)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept 1</td>
<td>-0.81 (0.30)</td>
<td>0.008</td>
<td>-1.34 (0.24)</td>
<td>&lt;0.0001</td>
<td>-2.62 (0.67)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Intercept 2</td>
<td>0.38 (0.30)</td>
<td>0.20</td>
<td>-0.12 (0.24)</td>
<td>0.60</td>
<td>-1.56 (0.67)</td>
<td>0.02</td>
</tr>
<tr>
<td>Intercept 3</td>
<td>1.15 (0.30)</td>
<td>0.0002</td>
<td>0.83 (0.24)</td>
<td>0.0006</td>
<td>-0.68 (0.67)</td>
<td>0.32</td>
</tr>
<tr>
<td>Intercept 4</td>
<td>1.83 (0.30)</td>
<td>&lt;0.0001</td>
<td>1.47 (0.24)</td>
<td>&lt;0.0001</td>
<td>0.09 (0.67)</td>
<td>0.89</td>
</tr>
<tr>
<td>Intercept 5</td>
<td>2.89 (0.31)</td>
<td>&lt;0.0001</td>
<td>2.40 (0.24)</td>
<td>&lt;0.0001</td>
<td>1.21 (0.67)</td>
<td>0.07</td>
</tr>
<tr>
<td>Duration</td>
<td>0.11 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.15 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.42 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Location</td>
<td>0.43 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.10 (0.003)</td>
<td>&lt;0.0001</td>
<td>-0.11 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Neuro</td>
<td>0.15 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.27 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.06 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Onset</td>
<td>0.10 (0.03)</td>
<td>0.0002</td>
<td>0.26 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.11 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Severity</td>
<td>0.26 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.31 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.29 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Walking</td>
<td>0.24 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.33 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.34 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Female</td>
<td>0.33 (0.33)</td>
<td>0.31</td>
<td>-0.07 (0.14)</td>
<td>0.63</td>
<td>0.10 (0.23)</td>
<td>0.65</td>
</tr>
<tr>
<td>Years in Practice</td>
<td>-0.02 (0.006)</td>
<td>0.0008</td>
<td>-0.003 (0.005)</td>
<td>0.63</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NA Trained</td>
<td>0.15 (0.22)</td>
<td>0.49</td>
<td>0.29 (0.18)</td>
<td>0.12</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Monthly volume of consultations</td>
<td>0.006 (0.005)</td>
<td>0.24</td>
<td>0.006 (0.005)</td>
<td>0.23</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Orthopaedics</td>
<td>-0.16 (0.18)</td>
<td>0.39</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Age</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.009 (0.01)</td>
<td>0.36</td>
</tr>
<tr>
<td>Post-secondary Education</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.54 (0.23)</td>
<td>0.02</td>
</tr>
<tr>
<td>Consultation</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.79 (0.23)</td>
<td>0.0005</td>
</tr>
<tr>
<td>Between-subject Variance</td>
<td>0.38 (0.08)</td>
<td>N/A</td>
<td>0.50 (0.08)</td>
<td>N/A</td>
<td>0.81 (0.19)</td>
<td>N/A</td>
</tr>
<tr>
<td>Chi-square LR Test§</td>
<td>2186.2</td>
<td>&lt;0.0001</td>
<td>5307.7</td>
<td>&lt;0.0001</td>
<td>1433.2</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

SE = Standard Error  NA = North American  LR = Likelihood Ratio
* Random effects probit regression model.
† Random effects weighted probit regression model.
‡ Multilevel random effects weighted probit regression model.
§ Chi-square Likelihood Ratio test based on fixed effects models.
Table 2-7. Regression Results for Preferences about Surgery.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Surgeons*</th>
<th>P-value</th>
<th>FP§</th>
<th>P-value</th>
<th>Patients‡</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td></td>
<td>Estimate (SE)</td>
<td></td>
<td>Estimate (SE)</td>
<td></td>
</tr>
<tr>
<td>Intercept 1</td>
<td>-1.43 (0.26)</td>
<td>&lt;0.0001</td>
<td>-1.94 (0.20)</td>
<td>&lt;0.0001</td>
<td>-3.17 (0.63)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Intercept 2</td>
<td>-0.41 (0.26)</td>
<td>0.12</td>
<td>-0.60 (0.20)</td>
<td>0.003</td>
<td>-2.12 (0.63)</td>
<td>0.0009</td>
</tr>
<tr>
<td>Intercept 3</td>
<td>0.36 (0.26)</td>
<td>0.17</td>
<td>0.40 (0.20)</td>
<td>0.05</td>
<td>-1.29 (0.63)</td>
<td>0.04</td>
</tr>
<tr>
<td>Intercept 4</td>
<td>1.07 (0.26)</td>
<td>&lt;0.0001</td>
<td>1.10 (0.20)</td>
<td>&lt;0.0001</td>
<td>-0.47 (0.63)</td>
<td>0.45</td>
</tr>
<tr>
<td>Intercept 5</td>
<td>2.13 (0.26)</td>
<td>&lt;0.0001</td>
<td>2.08 (0.20)</td>
<td>&lt;0.0001</td>
<td>0.93 (0.63)</td>
<td>0.14</td>
</tr>
<tr>
<td>Duration</td>
<td>0.09 (0.03)</td>
<td>0.0005</td>
<td>0.14 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.38 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Location</td>
<td>0.43 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.15 (0.003)</td>
<td>&lt;0.0001</td>
<td>-0.016 (0.006)</td>
<td>0.003</td>
</tr>
<tr>
<td>Neuro</td>
<td>0.19 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.34 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.030 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Onset</td>
<td>0.07 (0.03)</td>
<td>0.01</td>
<td>0.24 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.14 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Severity</td>
<td>0.24 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.30 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.39 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Walking</td>
<td>0.23 (0.03)</td>
<td>&lt;0.0001</td>
<td>0.30 (0.003)</td>
<td>&lt;0.0001</td>
<td>0.39 (0.006)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Female</td>
<td>0.40 (0.28)</td>
<td>0.17</td>
<td>0.005 (0.12)</td>
<td>0.97</td>
<td>0.06 (0.21)</td>
<td>0.78</td>
</tr>
<tr>
<td>Years in Practice</td>
<td>-0.009 (0.005)</td>
<td>0.10</td>
<td>0.003 (0.005)</td>
<td>0.55</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NA Trained</td>
<td>0.24 (0.19)</td>
<td>0.22</td>
<td>0.01 (0.16)</td>
<td>0.94</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Monthly volume of consultations</td>
<td>-0.0008 (0.004)</td>
<td>0.86</td>
<td>0.003 (0.004)</td>
<td>0.45</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Orthopaedics</td>
<td>-0.31 (0.15)</td>
<td>0.05</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Age</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.019 (0.009)</td>
<td>0.03</td>
</tr>
<tr>
<td>Post-secondary Education</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.17 (0.22)</td>
<td>0.43</td>
</tr>
<tr>
<td>Consultation</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.47 (0.21)</td>
<td>0.03</td>
</tr>
<tr>
<td>Between-subject Variance</td>
<td>0.30 (0.06)</td>
<td>N/A</td>
<td>0.37 (0.06)</td>
<td>N/A</td>
<td>0.79 (0.17)</td>
<td>N/A</td>
</tr>
<tr>
<td>Chi-square LR Test§</td>
<td>2254.8</td>
<td>&lt;0.0001</td>
<td>5333.4</td>
<td>&lt;0.0001</td>
<td>1295.4</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

SE = Standard Error  NA = North American  LR = Likelihood Ratio
* Random effects probit regression model.
† Random effects weighted probit regression model.
‡ Multilevel random effects weighted probit regression model.
§ Chi-square Likelihood Ratio test based on fixed effects models.
Figure 2-1. In and Out of County Surgical Discharges by Surgeon Response.
Figure 2-2. Surgical Referral Preference Quartile Responses.

Symbols represent median responses.
Upper and lower tails represent 75th and 25th percentile responses, respectively.
Figure 2-3. Relative Importance of Clinical Factors in the Decision for Surgical Referral.

Error bars represent 95% confidence intervals.
Figure 2-4. Surgical Preference Quartile Responses.

Symbols represent median responses. Upper and lower tails represent 75th and 25th percentile responses, respectively.
Figure 2-5. Relative Importance of Clinical Factors in the Decision for Surgery.

Error bars represent 95% confidence intervals.
CHAPTER 3: How do Family Physicians’ Referral Practices Compare with their Preferences and Guidelines for Degenerative Lumbar Spinal Surgery?

1 Introduction

1.1 The Degenerative Lumbar Spine in Primary Care and the Use of Clinical Practice Guidelines

Degenerative disease of the lumbar spine (DDLS) is a common condition presenting to family physicians (FPs). Surgery for specific clinical conditions, such as spinal stenosis and degenerative spondylolisthesis, can benefit appropriately selected patients.71, 72 For most patients with back pain, initial presenting symptoms are vague and the underlying cause is often not determined.79, 80

In Ontario, since FPs serve as ‘gatekeepers’ to care, initial consultation is typically obtained in the primary care setting. For patients with DDLS, it is therefore primarily the FP who must discern who is the most appropriate candidate for surgical referral. Exposure to musculoskeletal disorders in medical school and family practice residency is disproportionately low relative to the knowledge and skill requirements.208, 209 Because the management of patients with back pain in primary care is challenging, many FPs experience frustration and there is often a divergence of treatment preferences between patients and physicians.210 In Chapter 2, we found that less than half of FPs felt that decisions around surgical referral were straightforward and less than one-quarter felt that deciding on surgery for their patients was straightforward. The implication is that under-referral of appropriate patients may prolong pain and disability for patients who might stand to benefit from surgery, while over-referral may exacerbate existing waiting lists for surgical consultations.115, 211

At least seventeen ‘evidence-based’ clinical practice guidelines (CPGs) have been developed to improve the quality of care and assist FPs by providing recommendations on investigation, treatment, and referral for a variety of DDLS conditions.87, 88 Lack of guideline awareness and adherence, as well as physician attitudes and behaviour, however, have limited their use.90, 91, 98, 101, 212
1.2 Specific Objectives

The objectives of this thesis chapter were to (1) obtain consensus from an expert-panel on recommendations for surgical referral from a widely used published CPG for patients with DDLS and (2) compare primary care surgical referral practices with individual FP preferences and CPG recommendations based on patient clinical factors.

2 Methods

2.1 Selection of Guidelines

We identified seventeen different CPGs from systematic reviews. From all the available guidelines, we selected for this study those put forth by the Agency for Health Care and Policy Research (AHCPR) in 1994. First, the guidelines have been adopted by many professional societies and organizations and have been used extensively to assess physicians’ guideline adherence. Second, while the specific treatment recommendations may have changed since that time, the evaluative and specialist referral component of these guidelines remains current. Third, recent systematic reviews have concluded that the AHCPR guidelines for referral continue to be valid. Finally, all of these guidelines were based on the available published evidence and the recommendations for referral are therefore similar, however, the AHCPR guidelines were found to score highest on quality compared with all other guidelines.

2.2 Questionnaire Design and Conjoint Analysis

In this chapter, we use the terms ‘preference’ to describe the attitude that individual FPs have for the decision to initiate surgical referral and ‘recommendation’ for the panel’s interpretation of the guideline. The rationale is that since we are using conjoint analysis (CA), a method designed to elicit preferences, for individual FPs that ‘preferences’ best describe their responses. Since our responses from the expert-panel are based on their interpretation of guidelines via a consensus process, the individual ‘preferences’ matter less and the overall consensus opinion towards referral is more accurately described as a ‘recommendation’.

CA, as explained in Chapter 2, is a five-step method used to elicit preferences. First, six clinical parameters (duration of pain, location of pain [i.e. leg vs. back], severity of pain, onset of
pain, neurological symptoms, and walking tolerance) were selected by a pilot group of surgeons, FPs, and patients from a total of sixteen factors identified by systematic review of the literature.116 Second, the clinical parameters were dichotomized into two levels each to reflect typical variation in clinical severity (Table 2-1). Third, sixteen vignettes were developed using a one-fourth fractional-factorial experimental design that reduced the total number down from 64 vignettes (2⁶). This experimental design method was used to reduce the number to a manageable size while maintaining a high efficiency for estimation whereby the levels of the six clinical factors were balanced and uncorrelated (Table 2-2).175 Fourth, responses were obtained from each respondent to establish preferences based on a six-point rating scale. Fifth, the data were analyzed using an ordered probit regression model.

We validated the analysis by measures of internal validity (that the direction of each regression coefficient made clinical sense) and internal consistency (comparison between two vignettes which differ in factor levels clearly identifying ‘more severe’ and ‘less severe’ scenarios). The estimated parameters from the regression model were used to determine the relative importance of each clinical factor in the decision for surgical referral as well as an estimated predicted likelihood of referral for varying combinations of factor levels.166

### 2.3 Delphi and the Expert-Panel

Delphi is an anonymous iterative group consensus process.213, 214 Despite its traditional use in establishing criteria for appropriateness of treatment, it has also been used to establish consensus for diagnostic criteria.215 Delphi, however, has not previously been used in the setting of CA to interpret guideline recommendations. In the Delphi process there is no contact between participants and, at each iteration, summary scores are calculated. A priori criteria for consensus are used to determine when no further iterations are required. Group summary responses are reported back to participants prior to each iteration.

Other methods of consensus include the nominal group and methods developed by the National Institutes of Health and Glaser.213 Of all these consensus processes, Delphi was considered the most appropriate because it lacks the requirement for participants from diverse geographic locations to meet in person. Furthermore, because it is anonymous, it prevents a single participant from influencing the entire group as is possible with other consensus processes, however, at the same time prevents the opportunity for clarification.216
The multidisciplinary expert panel consisted of ten members including four FPs, two rheumatologists, two neurosurgeons, and two orthopaedic surgeons. We identified potential panel members by a combination of (a) recommendations from local clinical experts and professional associations and (b) internet searches of academic physicians’ websites and relevant publication history. Potential members were asked to participate by email.

Panelists were asked to provide a strength of recommendation for surgical referral on a scale from 1 (very inappropriate) to 6 (very appropriate) for a series of 16 clinical vignettes. We chose to use an even scale to force participants into some preference (albeit mild) in favour or in disfavour of surgical referral. Participants were instructed to base their recommendation entirely on their interpretation of the CPGs and not on their individual opinions. Each participant was provided with the questionnaire as well as the relevant section of the AHCPR guidelines via email (including a direct internet link to the full publication). In this iterative process, there was no contact between participants. Median summary scores for each vignette were calculated and reported back to panelists prior to each iteration.

Of the many ways of measuring consensus, we chose to use Cronbach’s alpha ($\alpha$) measured across panelists. Cronbach’s $\alpha$, although traditionally used as a measure of internal consistency across items for a scale, has been used to measure consistency across panelists in Delphi, and thus can be considered a measure of consensus. It is given by:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{1}{n} \sum_{i=1}^{k} s_i^2 \right)$$

where $k$ is the number of panelists, $s_i^2$ is the variance of the $i^{th}$ panelist and $s_T^2$ is the variance of the total score formed by summing over all the panelists. As the covariance among panelists increases, the ratio of the sum of individual variances to the total variance decreases and $\alpha$ approaches one. Bland and Altman (1997) suggested that an alpha of 0.9 serve as a minimum cutoff for consistency. Using this criterion \textit{a priori} as our threshold, we repeated the iteration process until (1) either consensus was achieved ($\alpha \geq 0.9$) or (2) no change was observed in the responses.
2.4 Surgical Referral in Primary Care

A random sample of Ontario FPs, stratified by county, was surveyed (refer to Chapter 2 for details). FPs received the identical 16 vignettes designed from the CA experiment, based on the six clinical parameters. For each vignette, FPs were asked to rate their likelihood of surgical referral on a scale from 1 (very unlikely) to 6 (very likely). (“Based on [the given] information, how likely are you to initiate surgical referral for this patient?”)

In order to establish a direct link between patient and FP, we surveyed patients via their FPs (refer to Chapter 2 for details). Each FP was asked to identify patients with any history or current complaint with a minimum of two months of symptoms likely consistent with DDLS (with or without previous imaging, referral, or surgery). FPs were also asked in an open-ended question about other factors, apart from those clinical factors in the vignettes, that heavily influence their recommendations for surgical referral. (“Are there any other important factors not included in these vignettes, apart from further imaging, that you consider equally or more important in deciding about surgical referral for patients with lumbar spinal stenosis?”)

Each patient was asked to describe their own characteristics (symptoms) using the most appropriate level of each clinical factor used in the hypothetical vignettes from the CA design. Patients were also asked if they had been referred to a spinal surgeon for their condition. Only responses from FPs with complete patient responses (characteristics and referral) were included in the analysis. Questionnaires to FPs and patients are provided in Appendix A.

2.5 Statistical Analyses

For the Delphi panel, we calculated Cronbach’s $\alpha$ across panelists as well as Pearson correlation coefficients between each panelist and the group at each iteration. Once consensus was achieved ($\alpha \geq 0.9$) the survey responses from the final iteration were used for regression analysis. These results from the expert-panel’s interpretation of the CPG’s recommendations for referral were aggregated and analyzed first for validity and then using a random-effects ordered probit regression model (refer to Chapter 2 for details). Parameter estimates from the regression analysis were used to determine the relative importance the CPGs place on each clinical factor in the decision for surgical referral. Using these regression coefficients along with the patient-
reported characteristics (symptoms), we calculated a predicted recommendation of referral (from 1 to 6) based on the consensus-panel’s interpretation of the CPGs for each patient.²¹⁹

Collectively, responses from each individual FP were analyzed using a random-effects ordered probit regression model. Relative importance was calculated based on the regression coefficients. Individual (fixed-effects) regression models were also derived for each FP such that regression coefficients were similarly used with the clinical characteristics of their own patients to derive a physician-specific predicted preference for referral (from 1 to 6) for each individual patient.

In comparing actual referral with predicted CPG recommendation and FP-specific preference for referral, we considered actual referral to be the ‘reference standard’ and compared it with the predicted likelihood of referral obtained from both physician preferences and guideline recommendations. Receiver Operating Characteristic (ROC) curves were constructed and we calculated the area under the curve (AUC) as a measure of concordance.²²⁰ AUC represents the probability that a random pair of referred and non-referred patients would be correctly predicted (Null hypothesis: AUC = 0.5). Bootstrap standard errors and 95% confidence intervals were calculated using the SAS Macro BOOT1UAC.²²¹
3 Results

3.1 Panel Consensus

The panel reached consensus after two iterations. After the first iteration, Cronbach’s $\alpha$ was 0.86 and the panelist-group Pearson correlation coefficients ranged from 0.39 to 0.72 (Table 3-1). The three highest correlations from the first iteration were all FPs while the three lowest were surgeons. In the second iteration, after the median scores from the first iteration were reported back to panelists, Cronbach’s $\alpha$ increased to 0.96 and the panelist-group correlations ranged from 0.39 to 0.98. Similarly, the three highest correlations were all in non-surgeons (rheumatologists and FPs) while the three lowest were surgeons. The panelist-group correlation for one of the two orthopaedic surgeons increased (0.39 to 0.88) while for one of the two neurosurgeons it remained low (0.42 to 0.39). The responses obtained from this neurosurgeon only ranged between 4 and 6 on the 6-point scale in the final round. In contrast, with the exception of two physicians whose responses ranged from 2 to 6, the remaining 7 physician responses all ranged between 1 and 6.

3.2 Patient Participants

Of the 164 patients from 87 FPs that responded to the survey, there were 107 patients (65%) with complete responses on their typical symptoms from 61 FPs (70%) with valid CA responses. Thirty-three FPs (54%) had a single patient respond, 15 FPs (25%) had two patients respond, 9 FPs (15%) had three patients, 3 FPs (5%) had four patients, and only a single FP (1.6%) had responses from five patients. The characteristics of the patients included in this analysis are given in Table 3-2. In comparing our patient sample with Ontario patients who underwent surgery for DDLS obtained from hospital discharge data 2002-6 (refer to Chapter 4), there were no differences in mean age (p<0.11) or sex (p<0.7).

3.3 Importance of Clinical Factors and Concordance with Referral Practices

In the analysis of the conjoint data from the consensus panel, we found no violations of validity, 100% internal consistency, and no dominant preferences exhibited among the panelists. Validation of the survey responses from all 202 FPs in the entire survey is shown in Table 2-5. Results from the random-effects ordered probit regression analysis for the consensus panel and aggregated FPs are shown in Table 3-3. All clinical factor regression coefficients were highly significant for both groups (p<0.0004) and the direction was consistent with a priori
expectations (internal validity) suggesting that the clinical factors were interpreted correctly by respondents. Little between-panelist variation was observed demonstrating that the responses from the panelists were relatively homogeneous as a result of our consensus process. FPs, on the other hand, demonstrated significant between-physician variation and heterogeneity.

Magnitudes of the coefficients for each clinical factor are better compared using their relative importance. As shown in Figure 3-1, the panel considered the CPGs to place the highest importance on the duration of pain (26%) followed by the dominant location (22%) while FPs considered these factors to be of lowest importance (≤12%). On the other hand, FPs placed the highest importance on the severity of pain (23%), walking tolerance (22%), and onset of pain (18%). These three clinical factors were ranked lowest for the panel’s interpretation of the guidelines (≤13%).

In order to assess the accuracy of the predictive model for FPs, we identified all patients with self-reported characteristics that matched any one of the 16 vignettes presented in the CA questionnaire. Thirty-one patients (29%) reported symptoms with an exact match to one of the presented vignettes. Of these 31 patients, the predicted likelihood of referral based on the CA regression coefficients accurately matched the reported preference for that matching vignette in 29 cases (94%). In the remaining two cases, the prediction differed by a single point on the 6-point scale.

ROC curves, based on the 107 individual patients, showing the accuracy of prediction of actual referral by individual FP preference and the consensus panel’s interpretation of the CPG are depicted in Figure 3-2. The solid 45 degree line represents the information attributed to chance alone. As shown, the concordance between predictions from the CPG recommendations is only slightly better than the individual FP preferences, however, both curves are relatively shallow and close to the diagonal reference line. For FP preferences, AUC was 0.57 (95% confidence interval 0.49 to 0.67; [null = 0.5]) while for CPG recommendations, AUC was 0.64 (95% confidence interval 0.53 to 0.74) showing only a modest amount of concordance with actual referral practices.
3.4 Other Important Factors in Primary Care

From the initial survey of 202 FPs, 108 indicated that other factors apart from the six clinical factors strongly influence their decisions to initiate surgical referral for these patients in an open-ended question. The major themes in order of most to least frequently cited included (1) patient age/comorbidities – 66 responses, (2) the presence of other clinical factors (‘red flags’, worsening symptoms, clinical signs, imaging results, etc.) – 30 responses, (3) patient demand or willingness to consider referral/surgery – 21 responses, (4) availability/relationship with the consultant – 19 responses, (5) failure to respond to other treatments – 17 responses, (6) patient occupation or compensation status – 9 responses, (7) social supports/caretaker responsibilities – 4 responses, and (8) patient personality or presence of psychiatric undertones – 2 responses.

Apart from patient age and presence of medical comorbidities, the two most frequently cited factors influencing the decision for surgical referral include patient demand or willingness to consider surgery and the availability of a consultant (i.e. the proximity for patients and waiting time to be seen).

4 Discussion

In this study, we aimed to evaluate the relationship between the recommendations of widely available CPGs for spinal surgical referral and actual surgical referral practices. We found that it was possible to obtain consensus on how an expert-panel recommends referral for hypothetical vignettes based on their interpretation of guidelines. The panel found that the AHCPR guidelines placed highest importance on the duration of pain and the dominant location of pain (i.e. leg-dominant over back-dominant) in recommending surgical referral for patients with DDLS. Individual FPs, however, considered these factors to be of least importance suggesting that practicing FPs disagree with the guidelines in determining appropriateness of surgical referral. In comparison with the province-wide survey (Chapter 2), we observed that practicing surgeons placed the highest importance on the dominant location of pain, while patients placed highest importance on duration of pain suggesting that the best evidence balances the preferences of both surgeons and patients. FPs, on the other hand, had preferences that were contrary to those obtained from the CPGs.
Furthermore, there was poor concordance between actual referral practices and the preferences of FPs and recommendations of CPGs, implying that referral practices do not match individual FP preferences or guideline recommendations based on clinical factors alone. In the context of surgical referral for DDLS, individual physician opinion and evidence did not match behaviour.

Several prior studies have investigated the relationship between physician opinion and behaviour. Jones et al. (1990) reviewed the literature on criterion validity for written case simulations and found no clear relationship with physician behaviour.222 Langley et al. (1991) found that, in general, written scenarios were comparable to physician referral practices when assessed across geographic regions and are therefore useful in assessing physician practices.223 In a conjoint analysis, Bouma and colleagues (2004), found high agreement between cardiologists’ predicted preference for surgery in aortic stenosis based on written case vignettes and actual treatment decisions with an AUC of 0.81.170 They concluded that the discriminative ability of their model based on clinical vignettes was high in the context of risky medical decisions. The current study found considerably lower values of AUC (0.57 and 0.64 for FP and CPG, respectively). Only the CPG had an AUC statistically significantly different than the null (0.5). The differences between these studies may be explained by the focus on specialists, rather than primary care physicians and more risky and timely decisions (i.e. aortic valve surgery) rather than more discretionary surgery for DDLS. Discrepancies between FP preferences for referral and referral practices may also be biased by social desirability in which responses are more in keeping with what one believes is the right answer rather than what one does in reality. Alternatively, other factors, not included in our vignettes, may also play a significant role in directing referrals to surgeons.

In assessing individual FP comments, we found that the two most common non-clinical factors stated by FPs to influence their referral decisions included the availability of a surgeon and the preferences of patients. One FP, remarking on the lack of available consultants, stated the determining factor was, “[an] available surgeon to accept [the] referral. Most local orthopedic surgeons will not accept a referral about back pain. The regional Spine Clinic has a very long (approx. 10-12 month) waiting time.” Another FP expressed their frustration, “I've practiced 30 yrs - I've had one patient get surgery for spinal stenosis despite many referrals of more than 8 months leg pain; I've sent folks to neurologist/surgeon and despite [symptoms and imaging] I am told they are OK so I don't know who is a candidate!” One FP who reported on the influence of
patient demand stated, “patients usually have gone to chiropractors or physiotherapists first without success – they initiate referral (demand it)”. Another FP, conveying a sense of clinical uncertainty associated with patient demand, explained, “I am not sure of my referral. Physical exam and sometimes patient demand to send will influence my [decision].”

Results from this study suggest that there is wide discrepancy between evidence, individual primary care physician opinion and their behaviour. Given that guidelines for referral were evidence-based and widely endorsed, greater compliance with guidelines should improve the referral process. However, this study illustrates the complexity of medical decision making and that clinical factors alone are insufficient in explaining physician behaviour. One potential method to improve the referral process would be to develop strategies to educate FPs and adopt CPGs in their practices. However, since individual FP preferences for clinical factors did not match their behaviour it is unlikely that new knowledge and even different attitudes would result in changes in behaviour. A second approach would be to investigate non-clinical factors that influence referral practices, such as patient preferences or the FP’s relationship with the surgeon and to focus on strategies that improve communication among patients, FPs, and specialists.

Decision aids are a potential strategy to influence patient preferences and improve communication between physicians and patients. Decision aids for spinal stenosis have been shown to change patients’ opinions about their condition and the benefit of surgery. This method would enhance the shared decision making process among patients and FPs and reduce the expectation on the FP for a thorough explanation of the patient’s condition and the potential outcomes of surgical management.

Another strategy would be to focus on improving communication between individual FPs and consulting surgeons. Facilitating this dialogue for an individual consultant’s referral base has been identified as an important area for improving the understanding between generalist and specialist physicians. The inclusion of explicit indications for requesting referral by FPs in referral request letters and recommending surgery for surgeons in consultant response letters would, over time, result in a better appreciation and ultimately more agreement in indications for surgical referral. The end result is higher efficiency referral practices by using the consultant’s own preferences for what constitutes an appropriate referral.
This study had several limitations. Although potential limitations exist in any consensus process, including the composition of the panel members and their potential influence by group responses, the high consensus after only two iterations suggests that the size of the panel was adequate. Furthermore, the inclusion of both referring physicians (FPs and rheumatologists) as well as consulting surgeons (orthopaedic and neurosurgeons) strengthens the validity of our choice of panel composition. However, a different panel may have led to different results.

Second, the questionnaire to patients inquired about actual prior referral and not requested referral (which may have been refused). We cannot speculate on whether this difference may be confounded by surgeon preferences. Factors preferred by surgeons, such as dominant location of pain (i.e. leg-dominant over back-dominant), which are highly recommended by CPGs but not necessarily by FPs, may directly influence the likelihood of a patient having been seen by a surgeon.

Third, survey response rates from FPs were low (12%). Furthermore, only 65% of patients had complete responses on their characteristics and indication of prior referral. The large burden placed on FPs by completing individual surveys in addition to selecting and requesting their own patients to respond likely accounts for an overall low response, as demonstrated in another similar survey. Patients were purposefully selected via their FPs, not randomly selected, and therefore may not be truly representative of Ontario patients. However, the sample of 107 patients from across the province with a range of clinical severities can provide us with sufficient insight into the variability of surgical referral across Ontario. Furthermore, no significant differences were found in age and sex between our surveyed patient sample and Ontario patients having undergone surgery for DDLS. Thus, we have no reason to believe predictions of referral from nonresponders would show higher agreements.

Finally, predictions of referral based on the regression results from the CA are subject to error. Formal testing would have necessitated separating the data into separate derivation and testing subsamples. In the case of FPs, each individual physician had only sixteen observations to derive their coefficients. Determining a prediction model on fewer than this would have led to even lower precision (and possibly accuracy) of prediction. We did, however, validate the FP prediction by comparing the predicted referral with the reported likelihood of referral for patients with symptoms identical to those in the vignettes and found high accuracy (94%). One of the
primary goals of CA is to derive prediction models for marketing purposes and it has been shown that even individual-level CA is effective and robust for models of prediction. Moreover, areas under ROC curves have been used previously in other studies to measure agreement between implicit preferences and actual behaviour. Furthermore, the inclusion of additional clinical factors or non-clinical factors in our prediction model may have increased the predictive ability of our model and perhaps the concordance with actual referral rates. However, the feasibility of this approach was limited by physician response burden in the questionnaire.

By understanding how FP’s preferences differ from recommendations obtained from evidence-based CPGs, individual FPs’ may appreciate the importance these CPGs place on clinical factors (duration and location of pain) in recommending surgical referral for patients with DDLS. Further research into other non-clinical factors that influence referral practices, such as surgeon availability and patient preferences may help us reduce practice variation. Knowledge dissemination strategies designed to better align the preferences of practicing physicians with the evidence and account for influential non-clinical factors may ultimately improve the efficiency of referrals thus reducing waiting lists and the duration of pain and disability for many patients.

5 Conclusions

Referral practices are poorly predicted by individual FP preferences based solely on clinical factors alone. There may be some minimal concordance between actual referral practices and CPG recommendations, as interpreted by an expert-panel. Non-clinical factors may be more influential in guiding FPs’ referral practices. Understanding these other non-clinical factors may be important in reducing wide variation in referrals and improving the process of care.
Table 3-1. Panelist-Group Correlations over Two Iterations

<table>
<thead>
<tr>
<th>Panelist</th>
<th>Iteration 1</th>
<th>Iteration 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Physician 1</td>
<td>0.68</td>
<td>0.86</td>
</tr>
<tr>
<td>Family Physician 2</td>
<td>0.62</td>
<td>0.98</td>
</tr>
<tr>
<td>Family Physician 3</td>
<td>0.72</td>
<td>0.87</td>
</tr>
<tr>
<td>Family Physician 4</td>
<td>0.69</td>
<td>0.89</td>
</tr>
<tr>
<td>Rheumatologist 1</td>
<td>0.60</td>
<td>0.84</td>
</tr>
<tr>
<td>Rheumatologist 2</td>
<td>0.68</td>
<td>0.91</td>
</tr>
<tr>
<td>Neurosurgeon 1</td>
<td>0.42</td>
<td>0.39</td>
</tr>
<tr>
<td>Neurosurgeon 2</td>
<td>0.67</td>
<td>0.83</td>
</tr>
<tr>
<td>Orthopaedic Surgeon 1</td>
<td>0.41</td>
<td>0.74</td>
</tr>
<tr>
<td>Orthopaedic Surgeon 2</td>
<td>0.39</td>
<td>0.88</td>
</tr>
<tr>
<td>Cronbach’s $\alpha$</td>
<td><strong>0.86</strong></td>
<td><strong>0.96</strong></td>
</tr>
</tbody>
</table>
Table 3-2. Patient Characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Patients (n=107)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age (sd)</td>
<td>63.5 (12.9)</td>
</tr>
<tr>
<td>Female (%)</td>
<td>56 (52.8%)</td>
</tr>
<tr>
<td>Highest Level of Education</td>
<td></td>
</tr>
<tr>
<td>Did not complete High School</td>
<td>29 (28.4%)</td>
</tr>
<tr>
<td>Completed High School</td>
<td>25 (24.5%)</td>
</tr>
<tr>
<td>Post-secondary Studies</td>
<td>28 (27.5%)</td>
</tr>
<tr>
<td>University (undergraduate)</td>
<td>12 (11.8%)</td>
</tr>
<tr>
<td>University (professional or graduate)</td>
<td>8 (7.8%)</td>
</tr>
<tr>
<td>Surgical Consultation (%)</td>
<td>64 (59.8%)</td>
</tr>
<tr>
<td>Undergone Surgery (%)</td>
<td>25 (23.4%)</td>
</tr>
<tr>
<td>Self-Reported Clinical Symptoms</td>
<td></td>
</tr>
<tr>
<td>Duration 8 months or more (%)</td>
<td>101 (94.4%)</td>
</tr>
<tr>
<td>Severe Pain (%)</td>
<td>65 (60.8%)</td>
</tr>
<tr>
<td>Leg-dominant location of pain (%)</td>
<td>57 (53.3%)</td>
</tr>
<tr>
<td>Pain onset at night or at rest (%)</td>
<td>81 (75.7%)</td>
</tr>
<tr>
<td>Moderate leg weakness (%)</td>
<td>58 (54.2%)</td>
</tr>
<tr>
<td>Inability to walk 2 blocks (%)</td>
<td>69 (64.5%)</td>
</tr>
</tbody>
</table>

sd = Standard Deviation
Table 3-3. Regression Results for Consensus Panel (Guideline) Recommendations and Family Physician Preferences for Surgical Referral.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Consensus Panel</th>
<th>Family Physicians</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>P-value</td>
</tr>
<tr>
<td>Intercept 1</td>
<td>-2.13 (0.26)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Intercept 2</td>
<td>-0.82 (0.21)</td>
<td>0.004</td>
</tr>
<tr>
<td>Intercept 3</td>
<td>0.43 (0.21)</td>
<td>0.07</td>
</tr>
<tr>
<td>Intercept 4</td>
<td>1.62 (0.24)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Intercept 5</td>
<td>3.02 (0.35)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Duration</td>
<td>0.84 (0.10)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Location</td>
<td>0.72 (0.10)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Neuro Sx</td>
<td>0.51 (0.09)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Onset</td>
<td>0.33 (0.09)</td>
<td>0.0004</td>
</tr>
<tr>
<td>Severity</td>
<td>0.42 (0.09)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Walking</td>
<td>0.41 (0.09)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Between-subject Variance</td>
<td>0.28 (0.17)</td>
<td>N/A</td>
</tr>
<tr>
<td>Chi-square LR Test</td>
<td>133.9</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

SE = Standard Error
LR = Likelihood Ratio
Figure 3-1. Relative Importance of Clinical Factors in the Decision for Surgical Referral.

FPs = Family Physicians
CPGs = Clinical Practice Guidelines (Panel)
Error bars represent 95% confidence intervals
Figure 3-2. Receiver Operating Characteristic Curves for Prediction of Referral by Family Physician Preference and Guideline Recommendation.

FPs = Family Physicians
CPGs = Clinical Practice Guidelines (Panel)
Numbers represent cutoff points on the Likert scale (1-6)
Solid 45 degree diagonal line represents the null (no information); AUC for the null = 0.5
CHAPTER 4: Who’s in the Driver’s Seat? The Influence of Patient and Physician Enthusiasm on Regional Variation in Degenerative Lumbar Spinal Surgery

1 Introduction

1.1 Background

Direct and indirect costs attributed to musculoskeletal conditions now approaches 8% of gross domestic product in the US and disorders of the back account for a large proportion of these costs.\textsuperscript{12} Surgery for specific conditions of the degenerative lumbar spine, such as spinal stenosis or degenerative spondylolisthesis, can result in significant improvement over nonoperative treatment in appropriately selected patients.\textsuperscript{71, 72} However, regional variation in spinal surgery across geographic regions in the US has been shown to be greater than most other procedures.\textsuperscript{60, 148} Rates of spinal decompression (i.e. laminectomy/laminotomy) vary up to 12-fold among US counties, and in the case of lumbar fusion, there is more than a 20-fold range in rates.\textsuperscript{16} Variation in procedure utilization exists in concert with variation in health expenditures.\textsuperscript{225} This phenomenon, termed small-area variation,\textsuperscript{151} has been hypothesized to be better explained by differences in health care delivery, specifically, physician uncertainty or enthusiasm rather than by differences in disease prevalence.\textsuperscript{155, 156} The ‘surgical signature’, whereby the regional variation remains constant over time, has been demonstrated for surgical treatment of degenerative disease of the lumbar spine (DDLS).\textsuperscript{16, 148} It has also been hypothesized that preferences of patients also generate substantial variation in care.\textsuperscript{134, 154, 157, 226}

The implication of small-area variation is that in areas of high or low utilization, some patients may undergo surgery unnecessarily and some patients may not undergo surgery that may, in fact, benefit. In discretionary procedures, such as hip or knee replacement for degenerative joint disease, it has even been shown that there exists greater unmet need in high-utilization areas suggesting that the ideal rate may, in fact, be higher than in high-rate regions and that, overall, under-utilization may be the primary problem.\textsuperscript{154} While variation has been extensively investigated for many procedures, little research has focused on potential explanations.
Furthermore, although patient enthusiasm is thought to influence regional variation, this has not been studied.

### 1.2 Specific Objective

The objective of this part of the thesis was to examine the regional variation in rates of surgery for DDLS in Ontario and to determine the influence of the enthusiasm of patients, family physicians (FPs) or surgeons, controlling for regional demographics, healthcare resources, and disease prevalence on the regional variation in rates of surgery for DDLS.

### 2 Methods

#### 2.1 Defining the Cohort

Using administrative databases, we identified all patients ages 50 or older who were admitted to hospital for surgical treatment of DDLS in Ontario, Canada between 2002 and 2006. Data were obtained from the Discharge Abstracts Database (DAD) of the Canadian Institute for Health Information (CIHI) which maintains the records of all discharges within the Province of Ontario. Physician billing data from the Ontario Health Insurance Plan (OHIP) was linked to the DAD.

Although several population-based studies evaluating trends in DDLS have limited inclusion to patients ages 65 years or older, we chose to include patients ages 50 and over since more than 90% of patients with spinal stenosis are older than 50 and degenerative instability patterns are frequently seen in patients as young as 50.

The Province of Ontario contains 49 counties (Figure A 1) which have been used extensively in small-area variation analysis. In April, 2006, the province created 14 Local Health Integration Networks (LHINs). These LHINs are geographically unique districts (Figure A 2) established as not-for-profit corporations that are responsible for planning and funding health services in their own district. Therefore, establishing the utilization of surgical treatment for DDLS for each LHIN is a priority from a policy perspective. However, tracking the utilization by county allows us to better demonstrate and potentially explain the overall amount of variation across the province.

Data was collected from April 1, 2002 to March 31, 2007 using a modification of a cohort identification algorithm originally developed by Cherkin et al. (1992) and updated by Deyo et
al. in 2005. We aimed to identify all hospital discharges of patients, ages 50 or older, for the surgical treatment of DDLS, excluding patients treated for fractures, tumours, infections, or cervical/thoracic disease. The algorithm was further updated from International Classification of Diseases, 9th Revision (ICD-9) to 10th Revision (ICD-10) diagnostic codes and from Current Procedural Terminology (CPT) to Canadian Classification of Interventions (CCI) and OHIP billing procedural codes. Details of the algorithm are provided in Appendix B.

Discharge counts were recorded by sex, age groupings (50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80 or older), fiscal year (2002/3-2006/7), and region (county and LHIN). We classified each patient encounter as either a decompression alone without fusion or as a fusion with or without decompression. Each surgical hospitalization was attributed to the county where the patient resided and not where the surgery took place.

Since OHIP, the provincial government-funded insurance plan pays 95% of physicians on a fee-for-service basis (cf. alternate funding plans, private insurance, worker’s compensation) and provides universal healthcare coverage to all residents for medically necessary surgical procedures, there is little variation in patient insurance coverage or physician fees, making Ontario an ideal setting for the investigation of population-based regional variation in utilization. After removing multiple admissions, we determined the total number of discharges for each county-age group-sex-year stratum based on the residence of each patient.

### 2.2 Predictors of Regional Variation

In selecting the appropriate predictors of regional variation we used a conceptual framework based on Andersen’s Behavioral Model of Health Services Use (1995) shown in Figure 1-4.

#### 2.2.1 Perceived and Evaluated Need (Patient and Physician Enthusiasm)

From a province-wide survey of all orthopaedic and neurosurgeons, a random sample of FPs (stratified by county), and a purposeful sample of patients surveyed via their FPs (i.e. clustered within a FP practice), we calculated measures of patient and physician enthusiasm. The details of the survey design, implementation, and results are provided in Chapter 2. Briefly, a portion of the survey was designed, as part of a conjoint analysis (explained in more detail in Chapter 2), whereby surgeons were asked to rate their recommendations for surgery, FPs were asked to rate
the *appropriateness* of surgery, and patients were asked to rate the *likelihood of considering* surgery for each of sixteen hypothetical clinical vignettes, on a scale from 1 (strongly against surgery) to 6 (strongly in favour of surgery). In addition to ratings on these clinical vignettes, respondents were also asked about demographics and practice characteristics (for physicians).

As a measure of regional surgeon enthusiasm, we calculated a regional *enthusiasm to perform surgery* as previously described by Wright and colleagues.\textsuperscript{158} For each surgeon, the individual *enthusiasm to perform surgery* was calculated as the mean of the difference between each of their responses and all surgeons’ median response.

\[
E(S_j) = \frac{1}{16} \sum_{r=1}^{16} [R_r - \text{Median}(R_S)] \quad \text{(for the } j^{th} \text{ surgeon)}
\]

Therefore, the *county enthusiasm to perform surgery* was the mean of the individual enthusiasm scores, weighted by their reported volume of patients seen for this condition.

\[
E(C_i, S) = \frac{\sum_{j=1}^{J} E(S_j) \times Vol_j}{\sum_{j=1}^{J} Vol_j} \quad \text{(for the } i^{th} \text{ county)}
\]

In counties where there were no surgeons (or no surgeon response), we used the cross-boundary flow to impute a weighted measure of enthusiasm. For example, if County A had no surgeons and half of the patients residing in County A were treated in County B and the other half in County C then the estimated enthusiasm to perform surgery for County A would be 50% of the enthusiasm from County B and 50% from County C. In the case where no surgeon response was captured from County B, the imputed enthusiasm for County A would be entirely derived from County C. Despite only capturing surgeon response from 24 counties, only a single county had more than 10% of local patients operated on in that county without any surgeon response. The distribution of in-county and out of county surgical discharges by county surgeon response is shown in Figure 2-1 (refer to Chapter 2 for details about surgeon responses per county). Since we were able to capture surgeon responses proportional to the surgical volume in Ontario, this provided us with further justification for the use of imputing surgeon enthusiasm in counties with no response to simulate the experience of a patient residing in a county without a spinal surgeon.
As a measure of regional FP enthusiasm, we calculated a regional enthusiasm to recommend surgery in a similar manner, taken from their responses to the appropriateness of surgery. For each FP, their individual enthusiasm to recommend surgery was calculated as the mean of the difference between each of their responses and all FPs’ median response.

\[
E(FP_j) = \frac{1}{16} \sum_{r=1}^{16} [R_r - \text{Median}(R_{FP})] \quad \text{(for the } j^{th} \text{ FP)}
\]

Therefore, the county enthusiasm to recommend surgery was the mean of the individual enthusiasm scores, weighted by their reported volume of patients seen for this condition.

\[
E(C_i, FP) = \frac{\sum_{j=1}^{J} E(FP_j) \times Vol_j}{\sum_{j=1}^{J} Vol_j} \quad \text{(for the } i^{th} \text{ county)}
\]

In counties where the survey sampling frame had fewer than 10 FPs, we combined that county with the nearest county with the smallest FP population (more than 10 FPs). In these counties, we averaged the enthusiasm score from the two counties and applied them to both counties (refer to Chapter 2).

As a measure of regional patient enthusiasm, we calculated a regional enthusiasm to consider surgery. For each patient, nested within a FP practice cluster, the individual enthusiasm to consider surgery was calculated as the mean of the difference between each of their responses and all patients’ median response.

\[
E(P_{tk}) = \frac{1}{16} \sum_{r=1}^{16} [R_r - \text{Median}(R_{Pt})] \quad \text{(for the } k^{th} \text{ patient in the } j^{th} \text{ practice cluster)}
\]

The practice cluster enthusiasm to consider surgery was the mean of all individual enthusiasm scores within that cluster.

\[
E(P_{t\_cluster_j}) = \frac{1}{K_j} \sum_{k=1}^{K_j} E(P_{tk}) \quad \text{(for the } j^{th} \text{ practice cluster and } K \text{ represents the number of patients from that cluster)}
\]
Therefore, the *county enthusiasm to consider surgery* was the mean of the practice enthusiasm scores, weighted by the reported volume of patients seen by that FP.

\[
E(C_i, Pt) = \frac{\sum_{j=1}^{J} E(Pt_{-\text{cluster}_j}) \times Vol_j}{\sum_{j=1}^{J} Vol_j}
\]

(for the \(i^{th}\) county)

In counties with no patient response, we imputed the average response from all the directly neighbouring counties (with patient response) to provide an estimate of that county’s patient enthusiasm to consider surgery.

### 2.2.2 Predisposing Characteristics (Demographics and Social Structure)

We obtained population counts for each county-age-sex stratum from the 2006 Canadian Census.\(^{233}\) County socioeconomic status was measured by two variables. Using the Canadian 2006 census, we obtained the median earnings\(^ {234}\) from 2005 for full-year, full-time earners by county, sex and age groups (45-64, 65 and over) and the percent with a high-school certificate or higher\(^ {235}\) by county, sex, and age groups (45-54, 55-64, 65-74, 75 and over).

### 2.2.3 Personal Enabling Resources

As one measure of individual barriers to health services, we used the percentage of the population with a non-official language (English or French) native tongue obtained from the 2006 Canadian Census by county, sex, and age groups (50-54, 55-59, 60-64, 65-69, 70-74, 75 and over).\(^ {236}\)

### 2.2.4 Community Enabling Resources

We hypothesized that two predominant health service resources were associated with surgical rates, namely the availability of MRI scanners and the supply of physicians (FPs and spinal surgeons). Although the utilization of MRI has been shown to explain a significant amount of variation in spinal surgical rates,\(^ {237}\) we chose to measure the per capita supply of MRI scanners in order to avoid any confounding with our outcome.\(^ {238}\) We obtained the number of MRI...
scanners by county across Ontario for 2002 from a previously published report. Using county population counts we then calculated the MRI supply per million population for each county.

The supply of FPs was measured using the number of full-time equivalent FPs per 10,000 population in 2001-2002 for each county, previously published. These measures were population-adjusted for differential physician utilization rates by age group and sex and for flow of patients between counties. Furthermore, to best quantify FPs in general practice, quasi-specialists (i.e. FPs who have more than 50% of their billings in one focused area of practice such as, obstetrics, psychotherapy or surgery) were excluded from this measure of supply.

Since orthopaedic and neurosurgeons may change their scope of practice (i.e. specialize in spinal surgery) which could be confounded with our outcome of interest (utilization of spinal surgery), we chose to use the per capita supply of both orthopaedic and neurosurgeons instead of the supply of self-reported spinal surgeons. Surgeon numbers were obtained from a survey mailing list provided by the College of Physicians and Surgeons of Ontario. County population counts were used to calculate the surgeon supply per 100,000 population.

### 2.2.5 Disease Prevalence

As a measure of disease prevalence, we used data obtained from the Canadian Community Health Survey (CCHS) made available by Statistics Canada. The CCHS was designed to obtain estimates of health determinants, health status and health system utilization at a sub-provincial level. One component of the survey, focused on chronic health conditions, addressed the presence of health-professional diagnosed “back problems” lasting six months or longer. Using this as a proxy for DDLS, we determined the prevalence of disease by sex, age group, and county. In an attempt to obtain more stable stratum-specific rates, we combined the numbers over all 3 available cycles of the survey (2000-1, 2003, and 2005) to obtain age-sex-county prevalence rates. Despite the high prevalence of “back problems” and the relatively infrequent need for surgical intervention in this cohort, we used this to measure the relative regional prevalence of DDLS and not as a true estimate of disease (i.e. spinal stenosis) prevalence.

### 2.3 Statistical Analyses

Discharges were allocated to the region in which the patient resided and not the region where the hospitalization took place. We used age-sex-year direct standardization to determine the
regional utilization rates relative to the regional population ages 50 or older. Chi-square statistics were calculated for the comparison of each standardized regional rate to the provincial rate, adjusting for multiple comparisons. We considered a p-value of 0.004 (=0.05/14) and 0.001 (=0.05/49) to be statistically significant for a comparison across LHINs and counties, respectively.

We performed multiple logistic regression analysis to determine the association of age and sex on the rates of spinal surgery as well as trends over time with types of procedures (i.e. decompressions vs. fusions). We further calculated quantitative measures of small-area variation, namely the extremal quotient (EQ), the coefficient of variation (CV), and the systematic component of variance (SCV). The EQ, calculated as the ratio of the highest rate region to the lowest, is the most widely reported measure and likely the most understandable. However, since the quotient relies on only the highest and lowest rate, it is highly sensitive to extremes. The CV, calculated as the ratio of the standard deviation of the rates to the mean rate multiplied by 100, is based on all rates. The CV, however, may overestimate the amount of variation if rates are low or in small populations. Finally, the SCV is a measure of the variation in rates (multiplied by 1000) after adjusting for the random variation within regions. The SCV, like the CV, is based on all rates, however, it is stable across a range of rates and population sizes.

For each county-age-sex-year stratum, we collected discharge counts after removing multiple admissions. We used an over-dispersed random-effects multivariate Poisson regression to model the count data and reported the incidence rate ratios with associated 95% confidence intervals and p-values rather than regression coefficients. For count data, such as surgical discharges within a county-age-sex-year stratum, the Poisson model is most commonly used, however, the Poisson distribution assumes that the mean is equivalent to the variance. For a more robust fit of the data, the over-dispersed Poisson model is frequently utilized to account for greater variability than would be expected from a simpler model. Furthermore, we selected a random-effects analysis, with a variable county-specific intercept, to model the between-county variation (in addition to the within-county variation) since we hypothesized that there would be significant variation in utilization rates across counties a priori. We used the SAS Procedure GLIMMIX for the regression analysis. All statistical analyses were carried out using SAS v9.1 (Cary, NC).
3 Results

3.1 Small Area Variation

We identified 10,318 surgical discharges for DDLS in Ontario from April 1, 2002 to March 31, 2007 after removing multiple admissions (4.2% multiple admission rate). The overall provincial rate was 58.6 procedures per 100,000 population ages 50 and over. Demographics of the cohort are shown in Table 4-1. The overall prevalence of disease was 28.9%. Chronic back pain was more common in females than males (odds ratio 1.14, 95% confidence interval [CI]: 1.09 to 1.19) but not different across age groups. Significant variation in prevalence of chronic back pain was found across counties (p<0.0001) ranging from 24% to 35%.

Surgical rates for each age-sex category are illustrated in Figure 4-1. The highest provincial rate was seen in male patients ages 70-74 while the lowest was in females ages 80 and over. Logistic regression demonstrated that both age and sex were highly significant in predicting spinal surgical rates (Likelihood ratio Chi-square 649.6, p<0.0001). Females had a 9% lower odds of having spinal surgery compared with males (odds ratio 0.91, 95%CI: 0.87 to 0.94) despite reporting a greater prevalence of disease. Patients ages 70-74 and 75-79 had higher than twice the odds of having spinal surgery compared with those ages 80 and over (odds ratio 2.1 and 2.0, 95%CI: 1.9 to 2.3 and 1.8 to 2.2, respectively).

We found no significant time trend in overall surgical rates over the study period (Figure 4-2). There were nearly three times more decompressions performed compared with fusions (odds ratio 2.71, 95%CI: 2.60 to 2.83, p<0.0001). Over the 5-year study period, however, the rate of fusions increased on average by 5% per year (odds ratio 1.05, 95%CI: 1.03 to 1.08, p<0.0002) while the rate of decompression remained nearly unchanged.

Variation in surgical rates across LHINs is shown in Figure 4-3. The EQ, CV, SCV, and adjusted Chi-square for overall LHIN variation was 2.3, 23.3, 65.3, and 110.3 (p<0.0001), respectively. The three LHINs with the lowest rates (Toronto Central, Central West, and Waterloo Wellington), situated predominantly in urban areas, were all significantly lower than the province-wide rate (measured across LHINs) of 56.2 while the two LHINs with the highest rates (Erie St. Clair and Northeast) were both significantly higher (Figure A 2). At the county level, there was more overall variation compared with LHINs. The EQ, CV, SCR, and adjusted
Chi-square for overall county variation was 5.0, 28.0, 95.4, and 148.2 (p<0.0001). In comparison with other commonly performed procedures, lumbar spinal surgery was more variable than rates of hip and knee replacement, cholecystectomy, and hysterectomy (Table 4-2). Due to smaller numbers of procedures performed within each individual county, fewer individual counties were significantly different from the province-wide rate, after adjusting for multiple comparisons (Figure 4-4). The distribution of relative rates across Ontario counties is shown in Figure 4-5.

3.2 Explanation of Variation

Characteristics of the Ontario counties are shown in Table 4-3. Counties demonstrated a wide variation in the enthusiasm to perform surgery even among the 24 counties with surgeon responses. In sampling 43 counties, we gathered FP responses across 45 counties in total. In the six counties with fewer than 10 FPs in the sampling frame, we imputed the average responses from the pre-specified neighbouring county into that county. We obtained responses from patients across 35 counties. Using the directly neighbouring counties, we imputed the average response into the missing 14 counties to estimate that county’s enthusiasm to consider surgery.

We observed significant correlation between the enthusiasm of patients and surgeons towards surgery, both with the directly measured counties as well as imputed measures (Pearson correlation 0.46, p<0.05 and 0.39, p<0.006). Furthermore, there was an inverse relationship between the enthusiasm of FPs and patients for direct and imputed measures (Pearson correlation -0.36, p<0.04 and -0.32, p<0.03). No significant correlation was identified between the enthusiasm of surgeons and FPs for surgery (p<0.4). County standardized surgical rates were correlated with the supply of FPs (Pearson correlation 0.42, p<0.003) and the prevalence of back pain (Pearson correlation 0.53, p<0.0001) and inversely correlated with median income (Pearson correlation -0.46, p<0.001) and percentage with post-secondary education (Pearson correlation -0.33, p<0.02).

There were significant differences between the directly measured surgeon enthusiasm for surgery (24 counties) in urban and rural counties (rural -0.96, urban -0.29, t-test p<0.04). Furthermore, urban counties compared with rural counties had higher mean per capita MRI scanners (mean difference 2.3 per 1 million population, 95% confidence interval [CI] 0.4 to 4.3), per capita supply of surgeons (mean difference 2.7 per 100,000 population, 95%CI 0.9 to 4.5), median
income (mean difference $5014, 95%CI $3113 to $6915), and percentage with non-official native language (mean difference 16.1%, 95%CI 11.6% to 20.6%). Standardized surgical rates were not significantly different among urban and rural counties (p<0.5).

Results from the random-effects Poisson regression model are shown in Table 4-4. In keeping with our small-area variation analysis, we found significant between-county variation (variance 0.054, standard error 0.015) as well as a significant amount of over-dispersion (1.068, standard error 0.026). Overall, our regression model explained a small amount of the overall variance (Adjusted R-squared = 0.28) but the model was highly significant (Chi-square 1092.6, p<0.0001). From our regression analysis, we found the following factors significantly increased the incidence rates of surgery: the county surgeon enthusiasm to perform surgery, lower median income, lower proportion of non-official native language, age less than 80 years, and male gender.

4 Discussion

In this study, we aimed to examine the regional variation in rates of surgery for the degenerative lumbar spine in Ontario and to determine if it was associated with the enthusiasm of patients or physicians. We found that there exists wide variation in surgical rates for treatment of the degenerative lumbar spine. Our study represents the first study that evaluates how the enthusiasm of patients, in addition to physicians, directly relates to utilization rates in a population-based study and provides insight into the drivers of DDLS surgery. Although patients and FPs have variable enthusiasm for surgery, surgeon enthusiasm was found to be the dominant modifiable factor influencing surgical rates. Despite a significantly higher surgeon enthusiasm in urban counties, no significant differences in standardized surgical rates were observed between urban and rural counties. In addition to surgeon enthusiasm to perform surgery, lower income, higher proportion of native English or French, age less than 80 years, and male gender, were all associated with a higher incidence in surgical rates. No community resources (i.e. supply of MRI scanners, FPs, or surgeons) nor the prevalence of disease were found to be significantly related to surgical rates, however, the degree to which our model explained a large amount of the variation was only modest.
Prior research reported tremendous variation in lumbar spinal surgical rates across US counties in the Medicare population. Weinstein et al. (2006) noted overall lumbar surgical rates from 1992-2003 increased steadily but the most significant increase was in rates of lumbar fusion (30 per 100K Medicare enrollees in 1992, 60 in 1998, and 110 in 2003) while rates of lumbar decompressions remained largely unchanged (210 per 100K enrollees in 2003). Regional variation was quantified with an EQ of 7.96 and 21.0 and CV of 34.6 and 49.5 for lumbar decompressions and fusions, respectively. In the current study, overall rates of fusions increased from 13 per 100K population age 50 and over in 2002 to 17 in 2006 while rates of decompressions also remained largely unchanged (at about 41 per 100K). Weinstein and colleagues hypothesized that the underlying causes for this variation include a lack of scientific evidence, financial incentives, and differences in clinical training and professional opinion.

Bederman and co-authors (2009), in a study of reoperation rates for degenerative lumbar spinal surgery in Ontario, found that overall lumbar surgical rates were 28.3 per 100K population age 50 and over. The authors similarly found that fusion rates increased from 8.2 per 100K in 1995 to 11.5 per 100K in 2001 while rates of decompressions slightly decreased from 21 per 100K in 1995 to 16.8 per 100K in 2001. They noted that these trends occurred during a period where the supply of spinal surgeons decreased by 14%.

Wright and colleagues (1999) found, after controlling for the characteristics of the population and access to surgery, the enthusiasm of orthopaedic surgeons was the dominant modifiable determinant of variation in knee replacement surgery in Ontario. Conversely, for middle-ear surgery in Ontario, Coyte and colleagues (2001) reported that it was the enthusiasm of the referring physicians rather than surgeons that most significantly contributed to area variation. The present study found that surgeon enthusiasm, rather than that of referring physicians or patients was the main driver of surgical rates. Unlike these previous studies, the present study included the enthusiasm of patients as well as a proxy for disease prevalence as potential determinants of regional variation, noted to be major limitations in prior studies. Neither of these factors, however, was found to be significantly associated with surgical rates.

The influence of provider enthusiasm was first hypothesized by Chassin (1993). He argued against the prevailing beliefs that inappropriateness and clinical uncertainty were the main influences on regional variation and suggested that the ‘enthusiasm’ of physicians was the main
driver of utilization variation. In this study, we found that surgeon enthusiasm was highly associated with surgical rates.

The relationship between less knowledge of official languages and lower surgical rates found in this study suggests that language/ethnic barriers may exist for spinal surgery in Ontario. Ethnic disparities in musculoskeletal health are now being reported with increasing frequency. Similarly, in this study, despite a 14% higher prevalence of chronic back pain among females, we found that rates of spinal surgery were 9% lower for women compared with men. Hawker and colleagues (2000) similarly identified gender disparities in hip and knee replacement rates in Ontario. They found that the prevalence of hip or knee arthritis was 76% higher for women but surgical rates were 22% lower than men. In an attempt to explain this discrepancy, Borkhoff and co-workers (2008) determined that physician gender bias plays some role in this access disparity. Disparity in access to spinal surgery has not, however, been studied. Further research into patient, provider, and healthcare factors may help us better understand reasons behind these inequalities. Culturally competent care and increasing the diversity amongst spinal providers may be strategies to reduce potential access disparity.

This study has several potential limitations. First, a limitation inherent in all administrative database research is that the data, acquired from CIHI, were not designed for research purposes. We cannot be certain that we identified patients correctly and that coding is performed accurately across the entire province. Nonetheless, prior work has demonstrated that primary diagnoses and surgical procedures are coded accurately. Furthermore, our algorithm for cohort identification was modified from one well-used in Medicare claims data research and therefore is likely to be identifying patients appropriately.

Second, our measures of enthusiasm for physicians and patients may not properly characterize the true enthusiasm of these groups. Since surgeons do not practice in all 49 Ontario counties, we imputed measures of enthusiasm into ‘empty’ counties by determining the cross-boundary flow of patients, likely representing the experience of Ontario patients. Although this strategy introduces some colinearity within the independent variable (surgeon enthusiasm), it more accurately reflects the experience of a patient who resides in a particular county since the aggregate surgical assessment is comprised of a combination of opinions from consulting surgeons. For FPs, the survey design merged six counties together such that there were 43
independent measures of enthusiasm. Patient enthusiasm was measured across 35 counties and empty counties were imputed based on all neighbouring counties with patient response. Additionally, the volume of consultations for DDLS obtained from our physician surveys, which was used to weight physician enthusiasm, was based on physician recall and may not necessarily reflect actual practice. The absence of 49 independent measures of enthusiasm for surgeons, FPs, and patients should, if anything, result in a more conservative estimate of enthusiasm for all empty counties, particularly for surgeons thus lowering the likelihood of finding a significant effect.

Third, the relationships we have modeled between surgical discharges and other predictors are ecologic and cannot all be obtained for individual patients. A patient-level analysis would not be feasible since regional resources, such as MRI or physician supply, cannot be measured at the patient level.

Finally, our study was limited to the province of Ontario for 2002-2006 and may not necessarily generalize to other areas either within Canada or abroad. However, our findings are worth considering in any health system since patients and surgeons must generally agree on proceeding with surgery with or without an intermediate referring physician.

Regional variation in the utilization of surgery raises important concerns for medical providers, policymakers, and the public. Under-utilization can create access disparities and prevent ideal surgical candidates from obtaining beneficial care. Over-utilization saddles our health care resources by increasing waiting lists for other who stand to benefit more and is wasteful of scarce and expensive resources. Despite not knowing what the ‘ideal’ utilization rate should be, wide regional variation nonetheless implies a problem of either under-utilization, over-utilization, or both. We can speculate that the ideal rate may lie closer to that of the high-rate regions as is the case for hip and knee replacements. Therefore, strategies to understand why surgeons are not ‘enthusiastic’ for surgery despite good quality evidence for its effectiveness may reduce under-utilization and access to care.

5 Conclusions

Significant regional variation in surgical rates for DDLS was observed. Counties with higher rates of surgery had older male patients with lower income, and more knowledge of official
languages. Because surgeon enthusiasm appears to be driving surgical rates for DDLS surgery, strategies to understand why surgeons have variable ‘enthusiasm’ for surgery, despite high-quality evidence for its effectiveness with specific indications (i.e. spinal stenosis and degenerative spondylolisthesis), may help us target ways at informing current evidence with the aim of reducing variation, costs, and improving access to care.
Table 4-1. Characteristics of DDLS Surgical Discharges in Ontario, 2002-6

<table>
<thead>
<tr>
<th></th>
<th>All Procedures</th>
<th>Decompressions</th>
<th>Fusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>10,318</td>
<td>7538</td>
<td>2780</td>
</tr>
<tr>
<td>Mean Age (sd)</td>
<td>65.0 (9.6)</td>
<td>64.8 (9.8)</td>
<td>65.4 (9.3)</td>
</tr>
<tr>
<td>Female (%)</td>
<td>5222 (50.6)</td>
<td>3601 (47.8)</td>
<td>1621 (58.3)</td>
</tr>
<tr>
<td>Year (Rate per 100K)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>1866 (53.7)</td>
<td>1434 (41.3)</td>
<td>432 (12.5)</td>
</tr>
<tr>
<td>2003</td>
<td>2025 (56.7)</td>
<td>1498 (42.0)</td>
<td>527 (14.8)</td>
</tr>
<tr>
<td>2004</td>
<td>2197 (59.7)</td>
<td>1555 (42.3)</td>
<td>642 (17.5)</td>
</tr>
<tr>
<td>2005</td>
<td>2071 (54.8)</td>
<td>1543 (40.9)</td>
<td>528 (14.0)</td>
</tr>
<tr>
<td>2006</td>
<td>2159 (55.3)</td>
<td>1508 (38.8)</td>
<td>651 (16.7)</td>
</tr>
</tbody>
</table>

sd = Standard Deviation
Table 4-2. Comparison of Small-Area Variation Measures with Other Procedures in Ontario.

<table>
<thead>
<tr>
<th>Procedure/Condition</th>
<th>Extremal Quotient</th>
<th>Coefficient of Variation</th>
<th>Systematic Component of Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hip Replacement*</td>
<td>1.6</td>
<td>11.8</td>
<td>7.6</td>
</tr>
<tr>
<td>Cholecystectomy*</td>
<td>1.9</td>
<td>12.5</td>
<td>23.0</td>
</tr>
<tr>
<td>Knee replacement*</td>
<td>2.4</td>
<td>22.9</td>
<td>63.2</td>
</tr>
<tr>
<td>Hysterectomy*</td>
<td>2.5</td>
<td>26.1</td>
<td>94.4</td>
</tr>
<tr>
<td><strong>Lumbar spinal surgery</strong>§</td>
<td><strong>5.0</strong></td>
<td><strong>28.0</strong></td>
<td><strong>95.4</strong></td>
</tr>
<tr>
<td>Antireflux surgery†</td>
<td>5.7</td>
<td>52.1</td>
<td>175.8</td>
</tr>
<tr>
<td>Cardiac endarterectomy*</td>
<td>9.1</td>
<td>36.1</td>
<td>185.8</td>
</tr>
<tr>
<td>Tympanostomy tubes‡</td>
<td>9.6</td>
<td>37.2</td>
<td>267.9</td>
</tr>
<tr>
<td>Routine circumcision*</td>
<td>11.9</td>
<td>44.9</td>
<td>216.0</td>
</tr>
</tbody>
</table>

Based on Ontario discharge data (1992-1994), as part of the work done for *Patterns of Health Care in Ontario: ICES Practice Atlas, 2nd ed.* (1996)\textsuperscript{248}

\textsuperscript{†} Based on Ontario discharge data (1991-2002), adapted from Lopushinsky et al., *Surg Innov* (2007)\textsuperscript{230}

\textsuperscript{‡} Based on Ontario discharge data (1996-1999), adapted from Coyte et al., *JAMA* (2001)\textsuperscript{159}

\textsuperscript{§} Based on Ontario discharge data (2002-2006)
<table>
<thead>
<tr>
<th>Predictor</th>
<th>Counties (N=49)</th>
<th>Mean (sd)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgeon Enthusiasm to Perform Surgery</td>
<td>24</td>
<td>-0.63 (0.81)</td>
<td>-3.00 – 1.29</td>
</tr>
<tr>
<td>FP Enthusiasm to Refer for Surgery</td>
<td>45</td>
<td>-0.16 (0.44)</td>
<td>-1.02 – 0.88</td>
</tr>
<tr>
<td>Patient Enthusiasm to Consider Surgery</td>
<td>35</td>
<td>-0.31 (1.00)</td>
<td>-3.00 – 2.00</td>
</tr>
<tr>
<td>Population Size (per 1000 adults ≥ 50)</td>
<td>49</td>
<td>232.9 (391.9)</td>
<td>12.7 – 2481.5</td>
</tr>
<tr>
<td>Median Income (per $1000)</td>
<td>49</td>
<td>26.4 (3.7)</td>
<td>19.1 – 37.6</td>
</tr>
<tr>
<td>Completed High School (%)</td>
<td>49</td>
<td>65.5 (6.3)</td>
<td>49.1 – 80.0</td>
</tr>
<tr>
<td>Non-Official Native Language (%)</td>
<td>49</td>
<td>13.8 (10.1)</td>
<td>3.6 – 47.7</td>
</tr>
<tr>
<td>MRI Supply (per 1M)</td>
<td>49</td>
<td>2.0 (3.2)</td>
<td>0.0 – 11.7</td>
</tr>
<tr>
<td>FP Supply (per 10K)</td>
<td>49</td>
<td>7.0 (0.8)</td>
<td>5.5 – 9.1</td>
</tr>
<tr>
<td>Surgeon Supply (per 100K)</td>
<td>49</td>
<td>4.0 (3.1)</td>
<td>0 – 11.3</td>
</tr>
<tr>
<td>Chronic Back Pain (%)</td>
<td>49</td>
<td>29.1 (3.1)</td>
<td>23.9 – 36.1</td>
</tr>
</tbody>
</table>

sd = Standard Deviation
Table 4-4. Regression Results for County Spinal Surgical Rates

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Incidence Rate Ratio</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (per 100,000)</td>
<td>48.1</td>
<td>(21.4, 107.8)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Surgeon Enthusiasm to Perform Surgery</td>
<td>1.29</td>
<td>(1.07, 1.56)</td>
<td>0.008</td>
</tr>
<tr>
<td>FP Enthusiasm to Recommend Surgery</td>
<td>1.17</td>
<td>(0.98, 1.39)</td>
<td>0.09</td>
</tr>
<tr>
<td>Patient Enthusiasm to Consider Surgery</td>
<td>1.04</td>
<td>(0.95, 1.14)</td>
<td>0.4</td>
</tr>
<tr>
<td>Median Income (per $10,000)</td>
<td>0.90</td>
<td>(0.84, 0.96)</td>
<td>0.002</td>
</tr>
<tr>
<td>Completed High School (per %)</td>
<td>0.90</td>
<td>(0.48, 1.68)</td>
<td>0.7</td>
</tr>
<tr>
<td>Non-Official Native Language (per 10%)</td>
<td>0.91</td>
<td>(0.86, 0.97)</td>
<td>0.004</td>
</tr>
<tr>
<td>MRI Supply (per 1000K)</td>
<td>1.02</td>
<td>(0.99, 1.05)</td>
<td>0.15</td>
</tr>
<tr>
<td>FP Supply (per 10K)</td>
<td>1.08</td>
<td>(0.99, 1.19)</td>
<td>0.09</td>
</tr>
<tr>
<td>Surgeon Supply (per 100K)</td>
<td>0.99</td>
<td>(0.96, 1.02)</td>
<td>0.5</td>
</tr>
<tr>
<td>Chronic Back Problems (per %)</td>
<td>0.82</td>
<td>(0.56, 1.19)</td>
<td>0.3</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-54</td>
<td>1.25</td>
<td>(1.01, 1.55)</td>
<td>0.04</td>
</tr>
<tr>
<td>55-59</td>
<td>1.37</td>
<td>(1.15, 1.63)</td>
<td>0.0004</td>
</tr>
<tr>
<td>60-64</td>
<td>1.70</td>
<td>(1.43, 2.02)</td>
<td>0.0001</td>
</tr>
<tr>
<td>65-69</td>
<td>1.84</td>
<td>(1.65, 2.04)</td>
<td>0.0001</td>
</tr>
<tr>
<td>70-74</td>
<td>2.14</td>
<td>(1.93, 2.38)</td>
<td>0.0001</td>
</tr>
<tr>
<td>75-79</td>
<td>2.00</td>
<td>(1.82, 2.20)</td>
<td>0.0001</td>
</tr>
<tr>
<td>80 and over</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Female Gender</td>
<td>0.85</td>
<td>(0.80, 0.90)</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Between-County Variance 0.054 (standard error 0.015)
Over-dispersion Parameter 1.068 (standard error 0.026)
Chi-square Likelihood Ratio test 1092.6 (p<0.0001)
Adjusted R-squared 0.28
Figure 4-1. Province-wide Age and Sex Specific Rates of DDLS Surgery, Ontario, 2002-6
Figure 4-2. Total Discharges for DDLS Surgery by Procedure, Ontario, 2002-6.
Figure 4-3. Variation in LHIN DDLS Surgical Rates

Solid bars significantly different from overall Provincial rate (P<0.004).
Figure 4-4. Variation in County DDLS Surgical Rates

Solid bars significantly different from overall Provincial rate (p<0.001).
Figure 4-5. Regional variation in DDLS surgical rates across Ontario.
CHAPTER 5: Conclusions and Discussion

In this chapter, we first outline the major findings of the study. Second, we provide a discussion of how our findings compare with respect to our conceptual framework (Figure 1-4). Third, we discuss the patterns of regional variation in spinal surgery, the preferences of patients and physicians (perceived and evaluated need), and how the enthusiasm of patients and physicians influences regional variation in surgical rates. Finally, we explain the major limitations of the research and discuss the implications and directions for future work.

1 Summary of Conclusions

The goals of this dissertation were to (1) evaluate how preferences for surgical treatment for degenerative disease of the lumbar spine (DDLS) differ between patients and physicians, (2) investigate how referral practices in primary care compare with recommendations of family physicians (FPs) and clinical practice guidelines (CPGs), and (3) characterize the regional variation in utilization of surgery for this condition and explain how it may be influenced by the enthusiasm of patients and physicians.

From these analyses, we found that FPs, over patients and surgeons, express the highest overall preferences for surgical referral and surgery. Surgeons, FPs, and patients place differing importance on various clinical factors for surgical treatment of DDLS. Surgeons, for instance, considered the location of pain (i.e. leg-dominant over back-dominant) to be the major factor determining the decision for surgery which comprised only 10% of the importance of FPs’ decisions. Patients, on the other hand, placed large emphasis on quality of life symptoms, including the duration of pain, severity, and walking tolerance, in considering having surgery.

Actual referral practices were poorly predicted by both individual FP and evidence-based CPG recommendations, based solely on clinical factors alone. We hypothesize that non-clinical factors, such as the availability of a surgeon or strong patient preferences, may be more influential than clinical factors in guiding FPs’ referral practices. Significant regional variation in surgical rates for DDLS was observed across Ontario from 2002 to 2006. Counties with higher rates of surgery had older male patients with lower income, and more knowledge of official languages. Although patients and FPs had variable enthusiasm for surgery, surgeon enthusiasm was found to be the dominant factor influencing surgical rates.
2 Discussion

2.1 Current Trends and Regional Variation in Spine Surgery

The utilization of spine surgery has been increasing dramatically over the past two decades. In addition, Weinstein and colleagues (2006) found that there was tremendous variation in lumbar spinal surgery across US counties in the Medicare population. In Ontario, although there was not as much variation compared with the US, in our research spinal surgery was found to be more variable than hip and knee replacement.

Variation in physician preferences for managing patients with DDLS has been previously studied. From the primary care perspective, over and under-referral of these patients has been demonstrated with up to 70% referred inappropriately and nearly 50% not referred with potential surgical indications. For surgeons, variable preferences for surgical treatment have been shown and attributed, in part, to physician characteristics, such as specialty and years in practice.

If regional variation is due to variable preferences or variable perceptions of outcome on the part of physicians and patients, this could result in undesirable error. Even without knowing what the ideal surgical rate should be, one could speculate, presuming variation is not due to disease prevalence, that in areas of high utilization some patients may undergo surgery unnecessarily and in areas of low utilization some patients may not undergo surgery that may, in fact, benefit from surgery, or some combination of these two. As seen in other discretionary procedures, such as total joint replacement, greater unmet need can even exist even in high-utilization areas. This demonstrates that the ideal rate, for at least one procedure, may even be higher than that of high-rate regions and overall under-utilization may be the primary problem. Weinstein and colleagues hypothesized that the underlying causes for regional variation include a lack of scientific evidence, financial incentives, and differences in clinical training and professional opinion.

In our conceptual model (Figure 1-4) adapted from Andersen’s Behavioral Model of Health Services Use (1995), many factors, including the perceived and evaluated need for surgery, predisposing characteristics, and enabling resources may potentially influence the use of health services. Below is a discussion of how our study findings relate to the conceptual framework.
2.2 Perceived and Evaluated Need

2.2.1 Family Physician Preferences

Previous research has shown inappropriate surgical referrals for musculoskeletal disorders ranging from 25-40%.\textsuperscript{110-112} And, in the case of referrals for patients with back pain, 70% were considered more appropriate for non-surgical management.\textsuperscript{113} On the other hand, 47% of patients with back pain managed by their FP had potential indications for surgery but were not referred to surgeons illustrating the problem with both over- and under-referral of these patients.\textsuperscript{114}

We hypothesized that there would be wide variation in preferences between patients and physicians and that from the physician perspective, variation in their characteristics, beyond specialty and years in practice, would influence variation in preferences for surgical referral and surgery. Based on O’Neill and Kuder’s Physician Propensity Framework (2005), other important factors in clinical training, such as the practice patterns of the training facility (postgraduate and/or fellowship), and the clinical experience of their own practice, including the types of patients commonly seen, may well play a role.\textsuperscript{161}

In our study, we found major differences in preferences between surgeons, FPs, and patients in their assessment of appropriateness for surgical referral and surgery. FPs, on average, did not identify a single dominant clinical factor for preferring surgical referral or surgery but based their decisions on relatively equal importance of neurological symptoms, pattern of onset, pain severity, and walking tolerance. Dreinhofer and colleagues (2006) similarly found major differences in importance of clinical factors between referring FPs and surgeons concerning indications for total hip replacements.\textsuperscript{147} In our study, no specific FP characteristics were associated with higher preferences for referral or surgery. Thus suggests that the variation in physician preferences is not well explained by the characteristics of the physician. Furthermore, we found that neither FPs’ recommendations based on clinical factors alone nor CPG recommendations for referral accurately characterized the patterns of referral. Thus, when it comes to primary care referral practices, not only is there no clear dominant factors influencing referral, but what FPs think they would do, based on our six clinical factors, does not match what they are actually doing in practice.
New evidence from the SPORT trials have demonstrated the effectiveness of surgical treatment for spinal stenosis and degenerative spondylolisthesis.\textsuperscript{71, 72} Because previous systematic reviews failed to demonstrate any significant benefit of surgery for spinal stenosis, the SPORT trial was designed as a randomized trial with an observational preference arm to determine the benefit of surgery for spinal stenosis, degenerative spondylolisthesis, and even disc herniation.\textsuperscript{69} In their trial for degenerative spondylolisthesis, the authors found that due to strong patient preferences there was tremendous cross-over (approximately 40\% in each direction) within the randomized cohort questioning the value of a typical intention-to-treat analysis.\textsuperscript{71} In the trial for spinal stenosis, the authors found a similarly high rate of cross-over to surgical treatment (43\%).\textsuperscript{72} For lumbar spinal stenosis, the authors found that patients who underwent surgical treatment had a nearly 14 point greater improvement than nonoperative patients on the Bodily Pain domain of the SF-36 (out of 100 points) at two years following treatment while for degenerative spondylolisthesis the improvement was over 18 points greater. Although the trial was designed as a preference randomized trial and analyzed as an observational trial, it demonstrates the difficulty in performing a purely randomized trial with the availability of surgical treatment in conditions with strong patient preferences. A randomized trial without any crossover is therefore not feasible and these trials therefore represent the best available evidence for the benefit of spinal surgery in DDLS.

While evidence with the passage of time should become known for physicians, it seems unlikely that simply having that information alone will translate into a behavioural change. Therefore, imparting knowledge alone, on an aggregate level to FPs will be unlikely to result in any meaningful improvement and given the influence of surgical enthusiasm on surgical rates, this may be less important. Theoretically, it would seem that targeting physician behaviour may best be done at a more ‘upstream’ time-point, such as during physician training. Although this has a greater potential yield, this would take a generation of physicians to bring about change. For physicians currently in practice, knowledge transfer strategies that can change current attitudes are needed. Multiple strategies exist, such as clinical practice guidelines (CPGs) and continuing medical education (CME) courses, however, attempts at changing behaviour even with these strategies have been inconsistent and they lack ongoing feedback loops. As stated above, information even from better designed guidelines may not have a significant impact on referral
practices and perhaps other strategies helping patients and physicians, such as through decision aids, discussed further below, may be more useful.  

Other non-clinical factors and clinical factors not included in the vignettes, as suggested above, may play a significant role in directing referrals to surgeons. We found that individual FP preferences, based solely on our six clinical factors (dominant location of pain, walking tolerance, pain severity, pain duration, pattern of onset, and neurological symptoms) poorly predicted actual referral practices. Open-ended questions probing into individual FP explanations for their choice of referrals, identified that the most common non-clinical factors influencing referral decisions were the availability of a surgeon (proximity and waiting time) and the preferences of patients. Langley and co-authors (1997), from their physician survey, also found that the relationship with the consultant was one of the most important non-medical factors influencing decisions for referral. As stated above, we have shown that there exists a disconnect between the opinions of FPs and surgeons. Anecdotal information would suggest that surgeons may refuse a consultation request if their waiting list is too long and the information provided by the FP does not convey any urgency or appropriateness of the consultation to the surgeon. As a strategy to circumvent this latter problem, some surgeons require FPs to obtain a MRI or CT prior to consultation in order to demonstrate a potential surgical indication. Although this may reduce the number of inappropriate referrals coming through a surgical office, it raises concerns for inappropriate use of these modalities which may unnecessarily increase healthcare costs and even possibly surgical rates.

As an aside, it is important to recognize the different roles of FPs and surgeons in the management of patients with DDLS. FPs must make sure that they do not miss a referral for an urgent indication. Similarly, FPs must manage patients over the long term and in order to nurture the relationship with their patients, they must be empathic to their situations and request for referral. Surgeons, on the other hand, must be very selective in their indications for surgery and only operate on those who clearly meet appropriate indications. Furthermore, the surgeon may discharge a patient who they believe will not benefit from surgical management even if their symptoms are severe. Because of these varying roles, it is not entirely surprising that FP over-referral and surgeon under-treatment was observed.
We speculate that improved communication between individual FPs and their consultants may provide one avenue for bringing about change that has not been well explored. With a better dialogue between referrer and consultant, the surgical indications and preferences for referrals would become better understood by referrers. For surgeons, the reasons behind the initiation of individual referrals by FPs would also become clearer such that unnecessary use of expensive imaging modalities could be prevented. Furthermore, unique indications for surgical referral and surgery for each individual surgeon could be learned by the referring FPs. Those patients who do not require any further imaging or referral could be better managed by the individual FPs or other non-surgical specialists and other patients who might benefit from surgery or a surgical opinion could be seen more expeditiously.

Another strategy that could directly improve the efficiency of the referral process for FPs, surgeons, and even patients is through the use of centralized assessment and referral centres. This method, although not directly influencing physician knowledge or attitude can redirect inappropriate referrals and expedite appropriate and urgent referrals. One further controversial strategy is to link physician reimbursement to appropriate care. In the US, pay-for-performance is an emerging strategy to get physicians to meet certain performance measures for quality and efficiency.\textsuperscript{250-254} We discuss these strategies in more detail below.

\subsection*{2.2.2 Surgeon Preferences}

In the current study, surgeons were found to place the highest importance on the dominant location of pain (i.e. leg-dominant over back-dominant) in their decisions for surgical referral and surgery which were not found to be highly important for FPs or patients. Surgical outcomes for conditions with nerve root involvement, such as spinal stenosis, over back-dominant conditions, such as degenerative disc disease or facet arthrosis, are significantly better.\textsuperscript{201} This suggests that surgeons seeking the best outcomes for patients may select out the ‘leg-dominant’ patient over a ‘back-dominant’ patient to maximize the clinical benefit and patient satisfaction. Knowledge of this particular preference imparted to FPs may, in itself, reduce referral variation.

That clinical experience and specialty for surgical training influence surgical preferences for surgery fits nicely into O’Neill and Kuder’s (2005) Physician Propensity Framework.\textsuperscript{161} These factors shape a surgeon’s clinical baseline heuristic on a macro level. Higher preference for surgical referral among younger surgeons and neurosurgeons was observed in our study. We
hypothesize that younger surgeons are more willing to see more patients (including non-operative) earlier in their practice as they develop their referral base since they possess more capacity to take on new patients while older surgeons may have stronger preferences for selection of patients for consultation since their practice would be better established. Additionally, our survey sample included neurosurgeons which, on average, perform more cases than orthopaedic surgeons, but maintain a relatively equal consultation volume. Again, we hypothesize that more neurosurgeons perform spinal surgery as part of their general neurosurgical practice compared with orthopaedic surgeons and, thus, general neurosurgeons may have more experience with DDLS in clinical and surgical practice and also may be more ‘up-to-date’ with the current literature on spinal disease and techniques. Therefore, if a surgically appropriate patient is referred to a general neurosurgeon over orthopaedic surgeon, there may be a higher likelihood of having surgery with a neurosurgeon if neither posses a subspecialty interest in spinal surgery. Obviously, an opinion from a spinal subspecialist (either orthopaedic or neurosurgery) would be ideal, however, with few subspecialists available and the presence of general orthopaedic and neurosurgeons providing spinal care, it may be preferable to obtain a consultation from a general neurosurgeon over an orthopaedic surgeon. Again, an improved dialogue between surgeons and their referring physicians could clarify the reasons for a surgeon to deem one patient appropriate and another inappropriate. In turn, referring FPs would better understand the unique preferences of their consultants.

Another modality for knowledge transfer would be through the use of decision aids. Methods in which patients receive the information regarding their condition, the natural history, management options, and potential outcomes following surgery, could empower the patient to make a more informed decision regarding the need for surgical referral and potentially surgery taking a large proportion of the decision out of the hands of FPs. Furthermore, FPs could benefit themselves either directly from becoming familiar with these decision aids or indirectly through discussions with their patients who have been exposed to them. From a surgeons’ perspective, decision aids derived through peer-consensus and the best available evidence could better align patient perceptions of surgical outcome with those held by most surgeons. In turn, widespread use of decision aids whereby patients from many regions consider similar information could further reduce regional variation in surgical rates. Decision aids, in a systematic review, have
been shown to have some impact on the decisional conflict, patient satisfaction, and quality of life, however, their influence on the utilization of care is less clear.\textsuperscript{249}

2.2.3 Patient Preferences

Patients, on average, were found to have higher preferences for duration of pain, walking tolerance, and severity of pain – all symptoms related to quality of life. Patients with more education and those who have had prior consultation had higher preference for surgical referral and surgery. Whether this indicates that better educated patients have a higher need for information gathering, or whether those patients with higher education and some experience with spinal specialists are just more demanding is less clear. Differences between physicians’ evaluated need for surgery and patients’ perceived need have been previously documented.\textsuperscript{144} In a prior systematic review, Montgomery and Fahey (2001) noted that different treatment preferences exist between patients and physicians over a wide range of clinical conditions. They concluded that the magnitude and direction of these differences are not consistent across different clinical situations and that ultimately in the shared decision making process, the extent of patient participation depends on their desire for involvement and the clinical decision at hand.\textsuperscript{140}

It is interesting to note that from our consensus process, the multi-specialty panel interpreted the CPGs such that the highest importance was placed on the duration of pain and the dominant location of pain. Duration of pain was considered one of the most important factors for patients while the dominant location was most important for surgeons. This demonstrates that the best evidence can successfully balance the preferences of both surgeons and patients. The importance of considering patient preferences is now well accepted.\textsuperscript{134} Because of this, Owens (1998) strongly advocated for the inclusion of patient preferences into CPG development, particularly for conditions with discretionary treatments such as DDLS.\textsuperscript{131} Although patient preferences clearly influence the decision for surgical referral and surgery, we are reassured that clinical guidelines, based entirely on the best available evidence at hand, are somewhat aligned with patient preferences for referral and surgery. However, current referral practices do not seem to be well explained by guideline recommendations and further guideline development may not bring about significant change.
Many of our findings were based on preferences measured over the aggregate, however, decisions for surgery are made at the individual level. Strategies to direct decisions based on aggregate findings may ultimately hamper our efforts to enhance the shared decision making process. Shared decision making aligns the perceived need with the evaluated need by a process in which patient and physician reach mutual agreement after considering all of the information with respect to treatment options and outcomes. The ultimate goal is to arrive at a decision that integrates the preferences of patients with the recommendations of physicians. Its main advantages are that it allows for better information gathering for both patient and physician and it forces the physician to present all treatment options. Patients who engaged in shared decision making with their surgeons were found to have a greater sense of control, less concern over their disease, more satisfaction with treatment, and even improved clinical outcomes.\textsuperscript{203} Decision supports have been developed to facilitate this process which have resulted in better aligning patient and physician preferences.\textsuperscript{143, 204} One such support was found to have enhanced the decision making process for patients with spinal stenosis and even resulted in more patients opting for surgery and feeling better informed about their condition.\textsuperscript{143} The end result of enhancing the shared decision making process is that patients are better informed and more satisfied with their decisions and ultimately may have better outcomes whether they decide for or against surgery for any particular condition. Therefore, decision aids for patients, either with their FPs or with surgeons, may better align preferences and perceptions between patients and physicians.

2.3  Explanation of Regional Variation

2.3.1  Patient and Physician Enthusiasm

In explaining the influence of patient and physician enthusiasm on regional variation in utilization of spinal surgery, we found that the surgeons’ enthusiasm to perform surgery, along with lower income, higher proportion of official native language spoken, age less than 80 years, and male gender, were all associated with a higher incidence in surgical rates. Interestingly, neither community enabling resources (i.e. supply of MRI scanners, FPs, or surgeons) nor the prevalence of disease were found to be significantly related to regional variation in surgical rates. We chose to use the per capita supply of MRI scanners rather than the utilization of spinal MRI since the utilization of MRI has been shown to be closely related to surgical rates.\textsuperscript{237}
Wright and co-workers (1999) found, after controlling for the characteristics of the population, access to surgery, and the prevalence of disease, that the enthusiasm of orthopaedic surgeons was the dominant modifiable determinant of variation in knee replacement surgery in Ontario. Conversely, for middle-ear surgery in Ontario, Coyte and colleagues (2001) determined that it was the enthusiasm of the referring physicians rather than surgeons that most significantly contributed to this area variation.

The influence of provider enthusiasm was first hypothesized by Chassin (1993) who argued against the prevailing beliefs that inappropriateness and clinical uncertainty were the main influences on regional variation. He suggested that the ‘enthusiasm’ of physicians was the main driver of utilization variation. In this study, we found that surgeon enthusiasm, rather than that of FPs or patients, was highly associated with surgical rates. Unlike these previous studies, we included the enthusiasm of patients as a potential determinant of regional variation, hypothesized to be a significant contributor. Despite this hypothesis, patient enthusiasm was not found to be significantly associated with surgical rates. Even though patients ultimately decide for or against surgery if recommended, it appears that the main driver of surgical rates for spinal surgery is the surgeon.

Therefore, in attempting to reduce practice variation and address unmet need, it appears that strategies targeting surgeon preferences, rather than those of referring FPs or patients, may possess the largest impact. On the other hand, the referral from the FP comes at a point more ‘upstream’ from the surgeon’s decision and although there is more variability and a less clear relationship between FP enthusiasm and surgical utilization, interventions at the level of the FP may have a greater impact on overall utilization rates.

### 2.3.2 Ethnic and Gender Barriers in Spine Surgery

Although not a primary objective of this study, the ecologic relationship between knowledge of official languages and higher surgical rates found in this study suggests that language/ethnic barriers may exist for spinal surgery in Ontario. Ethnic disparities in musculoskeletal health are now being reported with increasing frequency. In this study, despite a 14% higher prevalence of back pain among females, we found that female gender was associated with a lower rate of surgery. Hawker and colleagues (2000) similarly identified gender disparities in hip and knee replacement rates in Ontario. They found that despite a higher prevalence of arthritis, more
severe symptoms, equal willingness to consider surgery with men, and more joint replacements, unmet need was still higher in women. Disparity in access to spinal surgery has not, however, been well studied. Further research into patient, provider, and healthcare factors may help us better understand reasons behind these inequalities. Encouraging culturally competent care and increasing the diversity among spinal providers are among some of the strategies aimed at reducing potential access disparity. Reducing disparities has now become a major focus of many organizations within orthopaedics (J. Robert Gladden Society, American Academy of Orthopaedic Surgeons [AAOS]) as well as in medicine in general (The Cobb Institute, Institute of Medicine). For example, the AAOS has recently published a guidebook for practicing orthopaedic surgeons and maintains a website (www.aaos.org/diversity) with an advertising campaign to address diversity in orthopaedics and better understand the cultural needs of their patients.\textsuperscript{257}

3 Limitations

This dissertation has several limitations. First, we had relatively low response rates from our survey and this is discussed in more detail in Chapter 2. We obtained responses from less than half of practicing surgeons. Our response rate was in keeping with several other Ontario surgeon surveys.\textsuperscript{205, 206} Furthermore, we believe that it underestimates the true proportion of eligible surgeons (i.e. those who treat spine patients). The FP response rate was significantly lower (12\%) and necessitated a second round of mailings in order to obtain responses from over 200 FPs. Nonetheless, our response rate was similar to another survey that attempted to capture patient data via FPs.\textsuperscript{181} No significant differences between FP responders and non-responders in sex, years in practice, and location of practice were observed. Since patients were sampled purposefully via their FPs, they may not truly represent the preferences of patients across Ontario, however, the patient demographics encompass a wide range of clinical severities and experiences with the health care system.

Second, conjoint analysis (CA) is a relatively new technique and its role as a rigorous methodological tool is evolving.\textsuperscript{169} Despite this, CA has been highly recommended as a quantitative technique to elicit preferences.\textsuperscript{177} We demonstrated high rates of internal consistency and internal validity. Whether eliciting preferences based on clinical vignettes to simulate the true preferences physicians have for real clinical encounters remains unclear. However, using holistic vignettes with implicit clinical factors instead of explicit lists of clinical parameters can simulate
real clinical scenarios more accurately. Langley and colleagues (1991) found that, in general, written scenarios were comparable to physician referral practices when assessed across geographic regions and are therefore are useful in assessing physician practices. We recognize that this technique comes at the expense of a limited number of clinical factors with which to test preferences, but, it is a practical limit to the number of factors respondents can consider in a vignette. One specific concern was that approximately 20% of patients reported the same response for all vignettes presented. This might be explained by either a frustration with the exercise or rather that these patients failed to understand the task. With the exception of older age and more severe pain, no other differences were noted between patient respondents suggesting this problem may be the latter. Since surveying patients having experienced clinical symptoms is more valuable, this may have come at the expense of less understanding of the rating task. Using a ranking or choice-based technique would doubtlessly have resulted in higher rates of task misunderstanding.

Third, limitations exist in all consensus processes, such as Delphi, including the composition of the panel members and their potential influence by group responses. However, the high consensus after only two iterations suggests that the size of the panel was adequate. Furthermore, the inclusion of both referring physicians (FPs and rheumatologists) as well as consulting surgeons (orthopaedic and neurosurgeons) strengthens the validity of our choice of panel composition. Nonetheless, a different panel may have led to different results. Although the results from this process may not truly reflect the ‘correct’ interpretation of these or other CPGs, nor can they be directly comparable with other surveys, the results are directly comparable with the physician survey utilizing the same series of vignettes. The use of the CPGs from the AHCPR may not reflect the most up-to-date information, in light of the new SPORT trials, however, they remain a well-known and well-studied guideline that is relevant for the indications for surgical referral.

Fourth, the relationship between patient and physician enthusiasm and regional utilization rates is ecologic and discussed further in Chapter 4. Furthermore, our measures of enthusiasm for physicians and patients may not properly characterize the true enthusiasm of these groups since we imputed measures of enthusiasm into ‘empty’ counties. However, the absence of 49 independent measures of enthusiasm for surgeons, FPs, and patients should, if anything, result in a more
conservative estimate of enthusiasm for all empty counties, particularly for surgeons thus lowering
the likelihood of finding a significant effect.

Finally, this survey was confined to the province of Ontario and our results may not necessarily
generalize to other jurisdictions. However, our results are worth considering in any health system
whereby referring physicians play an intermediary role with patients and surgeons regarding
decisions about referral and treatment. Even in health systems with self-referral, surgeon
understanding of patient preferences and practice patterns may facilitate the shared decision
making process and improve access disparities.

4 Significance and Future Directions

Understanding preferences for surgery on the degenerative lumbar spine among patients, FPs,
and surgeons and where significant sources of variation and disagreement exist, can have several
beneficial effects. Aligning the preferences and perceptions of patients with physicians would
have a direct improvement on the patient-physician shared decision making process. Patients’
expectations about surgical treatment for the degenerative lumbar spine would be more in
keeping with physician expectations. Physicians may better focus on the needs of their patients
rather than on physician-preferred outcomes. Aligning these preferences and expectations, can
directly result in a significant improvement in patient satisfaction with the healthcare process and
even overall health status following treatment.141 In preference-sensitive care, such as surgery
for DDLS, patients should be allowed to have variable preferences for surgery. Ideally,
however, preferences should be better aligned between surgeons and referring physicians.

From the primary care standpoint, by having a better appreciation that some surgeons place high
importance on symptoms related to nerve root involvement (i.e. leg-dominant pain), the referral
process may become more efficient allowing the most appropriate surgical candidates to obtain
timely assessments. Even with an appreciation for the importance that patients place on quality
of life symptoms, such as duration, severity, and walking tolerance, FPs may enhance the shared
decision making process and allow for patients with higher preferences for surgical treatment to
gather further information from surgeons.

Clearly, our conceptual framework depicted in Figure 1-4 oversimplifies the interactions
between patients and physicians. The dialogue between patient and FP, FP and surgeon, and
finally surgeon and patient is complex and cannot easily be incorporated into the current model.
Further work to develop a useful theory to model the communication between these three groups is needed.

Regional variation in the utilization of surgery raises important concerns for medical providers, policymakers, and the public. Under-utilization creates access disparities and prevents ideal surgical candidates from obtaining beneficial care. Over-utilization saddles our health care resources by increasing waiting lists for others who stand to benefit more and is wasteful of scarce and expensive resources. Therefore, strategies to understand why surgeons in low-rate regions are not ‘enthusiastic’ for surgery, despite high-quality evidence for its effectiveness, may help us target ways at informing current evidence with the aim of reducing under-utilization and improving access to care.

4.1 Influencing Physician Behaviour

In order for these benefits to be realized, however, strategies to affect physician behaviour are needed. Although knowledge translation was not a main focus of this thesis, we believe that there are areas worth exploring and work with specialists in the field of knowledge translation may be fruitful. Simplistically, we employ the notion that knowledge can lead to a change in attitude which can, in turn, influence practice behaviour (knowledge-attitude-behaviour). New knowledge concerning the benefit of operative treatment of the degenerative lumbar spine has been recently obtained from the SPORT trials. Furthermore, surveying patients and physicians regarding their preferences for referral and surgery in this condition and exploring factors that impact regional surgical rates has given us further knowledge into patterns of health service use.

As previously mentioned, the challenge of changing the behaviour of FPs requires considerable attention. In theory, strategies to disseminate knowledge, that can directly impact on the efficiency of practice, may influence attitudes and, in turn, behaviour. Unfortunately, we have shown that physician attitudes are not necessarily aligned with physician behaviour. Therefore, despite a multitude of available guidelines, further dissemination, endorsement, and even development of new ones are unlikely to change practice in primary care. We should, therefore, focus on understanding the information needs and barriers in order to support FP decision making through the use of decision aids, referral communication feedback, and centralized intake centres.
Surgeon practice patterns exist on a continuum with two extremes. On the one hand, some orthopaedic and neurosurgeons with or without fellowship training engage in a nearly-exclusive practice of spinal surgery (‘subspecialty surgeons’) while on the other hand, many surgeons maintain a practice of general orthopaedics or neurosurgery and may perform some spinal operations or consult on patients with spinal disorders (‘non-subspecialty surgeons’). For the latter group, maintaining an up-to-date knowledge of recent studies, emerging trends, and new technology is difficult and therefore influencing practice behaviour will be challenging. There is some evidence to suggest that current trends indicate that the prevalence of general practice orthopaedic or neurosurgeons is decreasing and that higher-volume surgeons (i.e. subspecialty surgeons) are performing most of the work. Although this may make the care provided by fewer surgeons more homogeneous, it may also limit the availability of care across regions. Older surgeons and those with orthopaedic over neurosurgical training may possess systematic differences in opinions regarding the effectiveness of surgery, particularly for surgeons not managing these patients. Mentoring trainees in residency programs, instructional course lectures at general meetings, and continuing medical education strategies to inform surgeons about the effectiveness of surgical treatment of the degenerative lumbar spine may modestly improve the perception, particularly for older orthopaedic surgeons. Encouraging mixed group surgical practices with both older and younger surgeons and spinal groups with orthopaedic and neurosurgeons may also foster a better collegiality and understanding of the benefits of surgical treatment for the degenerative lumbar spine. Indications rounds and case discussions in practice groups can further facilitate this dialogue in a collegial transparent and ongoing way.

Analogous to levels of prevention (primary, secondary, and tertiary), different strategies may take place at different levels and affect different stakeholders (patients, FPs, and surgeons). For example, decision aids are a form of primary intervention that can affect patients, FPs, and in turn, surgeons. At the secondary level, communication feedback strategies work between FPs and surgeons. And finally, centralized intake centres are analogous to tertiary intervention, in that they affect the referral process downstream from the primary care level. These strategies are discussed in more detail below.
4.2 The Role of the Patient

From the perspective of the patient, there is a concern that ‘who you see determines what you get’.\textsuperscript{14} At the primary care level, the FP determines who requires referral based on their perceived need. As we have shown, however, FP recommendations for referral are very different from those obtained from guidelines. Therefore, one patient who may be an ideal surgical candidate may not get referred while another who needs no surgical intervention (or even a surgical opinion) may be referred. Moreover, once a patient is referred, their fate may lie in the hands of the consultant and may depend on the specialty and surgical volume.\textsuperscript{14, 127} A general orthopaedic surgeon who does not perform spinal surgery but occasionally evaluates patients with degenerative spinal conditions, may provide a different recommendation to another surgeon with spinal fellowship training who engages in nearly full-time spinal work.

Decision support tools, as discussed above, can empower patients to ‘do their own homework’ and in some way determine if the symptoms they experience may be amenable to surgical intervention. If so, the patient, as their own advocate, can discuss their preferences and rationale for requesting referral to see a spinal specialist and bring the conversation with FP or surgeon to a higher level of understanding and more realistic expectations.

Patients clearly have preferences for considering surgery. Although much focus has been placed on patient selection from the surgeon’s perspective and changing patient expectations, little attention has focused on addressing the specific needs of patients. Surgery, aimed at correcting an anatomical abnormality, has historically focused on measures of outcome directly related to that change (e.g. change in the amount of pain, or change in a joint’s range of motion). For example, numerous instruments have been developed for the assessment of hip replacements from the 1950’s to the 1980’s.\textsuperscript{260-267} These scores have included such components as pain, movement/function, range of motion, and radiographic assessment. Over more recent decades, newer functional outcome tools that center around patient-oriented parameters, such as activities of daily living, social functioning, and mental health, have been more widely used such as the Western Ontario and McMaster osteoarthritis scale (WOMAC)\textsuperscript{268} as a disease-specific health measure, and the Short Form-36 (SF-36),\textsuperscript{269} a more general measure. These measures are commonly used in the assessment of function following total hip replacement.\textsuperscript{270} Although these instruments focus on patient factors on an aggregate level, they fail to address the individual
patient’s specific concern. Wright and Young (1997), for example, developed the Patient-Specific Index to assess outcome according to the preferences of the individual patient.\textsuperscript{271} Recently, even the WOMAC has been adapted to account for the patient-specific expectations and requirements of outcome.\textsuperscript{272}

In spinal disorders, a multitude of outcome measures have similarly been reported,\textsuperscript{273} however, few patient-specific outcome measures exist and none have been widely adopted into clinical practice for assessment following surgical treatment.\textsuperscript{274} Although there is criticism for the development of further outcome measures in spinal disorders, there is a clear need to include the specific preferences and expectations of patients into measures of outcome such that outcome following surgical treatment more accurately reflects what is important to the individual patient, not the aggregate or the treating surgeon.

One other finding that has significant implications for the patient is that ethnic/language and gender barriers may exist in spinal surgery. Disparities in access to care have been demonstrated for other musculoskeletal procedures. We have explained the findings in more detail above and described strategies to reduce disparities from the physician’s perspective. However, from the patient’s perspective, minority and women’s health groups may also serve to reduce disparities. Medical societies and government health offices can provide information to these focused groups to help educate the patients such that even in a medical system that may not deal well with other languages or cultures, patients themselves may be more empowered with an understanding of the course of the disease and options for treatment.

4.3 Novel Strategies

4.3.1 Decision Supports

At the most basic level of intervention (primary), a decision support tool for patients may possess the most influence. The use of decision aids in the surgical decision making of patients with DDLS has been studied. Deyo and co-authors (2000) found that a video program on back surgery enhanced the decision making process for patients with spinal stenosis in a randomized trial, and even resulted in more patients opting for surgery and feeling better informed about their condition.\textsuperscript{143} The use of other techniques such as pamphlets, or websites could also provide information on common degenerative spinal conditions. For example, the Canadian Spine Society (www.spinecanada.ca) is particularly well-suited to develop a website, using the expert
membership of spinal surgeons to provide opinions and interpret the current evidence in explaining common conditions including the natural history, the need for further investigations, surgical and non-surgical options and expectations and complications following surgery, from a surgeon’s perspective. Such an information resource would serve patients by providing them with important information in order to educate them about their condition, facilitate the shared decision making process with their physicians, and better align their expectations towards treatment. FPs could use this resource to allow patients to obtain their own information once a provisional diagnosis is made, and therefore take some of the anxiety of explaining the natural history and the (lack of) need for any further investigation or consultations. Furthermore, FPs may even educate themselves directly through this potential resource or indirectly through discussions with their patients who have done their own research. For surgeons, the patients that come to see them have already been given accurate information to begin with and therefore, the conversation between patient and surgeon could take place at a deeper level of understanding. Furthermore, the surgeon may even change their practice to conform to the information explained on the website, thus reducing regional variation in surgery.

Ultimately, a decision support based on clinical symptoms interpretable by patients and endorsed by the Canadian community of spinal surgeons (i.e. through the Canadian Spine Society) could increase patient and FP knowledge, improve the efficiency of referrals, enhance the shared decision making process for patients and physician, (and possibly surgical outcomes), and even reduce variation in surgical rates.

4.3.2 Communication Feedback

At a higher level of intervention (secondary), between FPs and surgeons, another strategy primarily aimed at informing practicing physicians of individual surgeon preferences is consultant information feedback. Facilitating the dialogue for an individual consultant’s referral base has been identified as an important area for improving the understanding between generalist and specialist physicians. Postcards, for example, whereby referring physicians check-off on a list of the most important reasons for requesting referral and, in return, surgeons similarly identify reasons for recommending surgery for a patient, could be exchanged with the referral request or consultation note from FPs and surgeons, respectively. A sample of a potential postcard from FP and surgeon is shown in Figure 5-1.
By facilitating this dialogue, FPs would better appreciate a surgeon’s criteria for appropriateness of referral or surgery and, over time, the FP’s recommendations might better align with the surgeon. Conversely, the surgeon may better understand the patient population encountered by that referring physician and if little changes are made in referral practices, he/she may choose to deliver a local seminar on dealing with DDLS in primary care for those FPs.

Communication feedback strategies, such as postcards and local seminars, may possess the potential to increase FP knowledge, influence referral rates, and even enhance the shared decision making process but further work is needed to determine the feasibility and costs of this strategy.

### 4.3.3 Centralized Referral Centres

At the highest level of intervention (tertiary), centralized referral intake centres could provide the most efficient way to bring about change in referral practices. In this model, intake centres, in which specially trained FPs, physiotherapists, or nurse practitioners accept and screen a wide variety of referrals from primary care and facilitate referrals to sub-specialists with either specific expertise (deformity correction, minimally invasive techniques) or with shorter waiting lists, thus improving the efficiency of the referral process. Through this rapid review and triage, patients who are deemed to be potential surgical candidates or who would require a surgical consultation could then obtain appropriate diagnostic investigations and referral to surgeons with a sub-specialty interest in spinal disorders. The Bone and Joint Health Network’s Model of Care (www.boneandjointhealthnetwork.ca) for hip and knee replacement, already underway in Ontario, provides an excellent model (Figure 5-2). In circumstances where FPs are unsure about the need for surgical referral or have difficulty in accessing a consultant, referral centres could serve a useful purpose. These intake centres would accept all referrals, prior to advanced imaging, and therefore take the burden off of FPs who might otherwise have challenges managing these patients. Once deemed appropriate for surgical referral, these patients could then be seen by surgeons with a special interest in spinal surgery, rather than the local general-practice or non-spinal surgeon in the community who may not possess the most up-to-date knowledge or skill to manage these patients. The main disadvantages to this model is the costs and manpower needed to develop this, however, it is possible that costs could be offset by a reduction in advanced imaging (i.e. CT, MRI) or surgical consultation for patients who only
require reassurance and not further investigation. Another limitation is that this modality will not change the knowledge or comfort of managing these patients in primary care and would therefore require a sustained commitment. Ultimately, however, centralized intake centres would make the referral process more efficient, potentially cut down on costs due to advanced imaging, and even possibly reduce variation in surgical rates by reducing access disparities. It remains unclear if the costs would be offset by the potential benefits.

4.3.4 Reimbursement Incentives

Pay for performance (P4P) aims to link physician reimbursement to the provision of high quality care. In the US and the UK, government and third-party payers are increasingly relying on the justification of treatments with proven clinical and cost-effective benefits. In this model, physician reimbursement incentives (or disincentives) may affect physician behaviour. Systematic reviews have found that P4P programs can improve the quality of care provided, particularly in the primary care setting. Referral practices from FPs may be modified by implementing financial incentives. For surgeons, however, P4P strategies are far more challenging. Simple performance measures must include methods that assess patients’ overall improvement in quality of life and not simply measures of utilization. Further work in this controversial field will help define the role and establish if potential benefits for improving the quality of care exist.

5 Concluding Remarks

In summary, we found that wide variation exists in the preferences of patients, FPs, and surgeons for the initiation of surgical referral or the consideration of surgery for degenerative conditions of the lumbar spine. Clinical factors of patients were viewed very differently between surgeons, FPs, and patients concerning the appropriateness of surgical referral or surgery. Although preference for referral and surgery was highest for FPs, their referral practices were not well explained by patient clinical factors alone either through individual FP preferences or by recommendations from clinical guidelines highlighting the influence of other non-clinical factors. Variation in surgical rates for treatment of the degenerative lumbar spine is significant and may be driven primarily by the enthusiasm of surgeons rather than referring FPs or patients. Local health care resources and personnel as well as variation in disease prevalence play little role.
By appreciating the variation in preferences between surgeons, FPs, and patients, FPs may be able to improve the efficiency of referrals and all physicians may enhance the shared decision making process thus improving the delivery of care for patients. With an understanding of the influence that surgeons have in driving variation in surgical rates, further research into why surgeons may be less enthusiastic for surgery may allow us to direct strategies at surgeons in low-rate regions thereby improving access and allowing for a more equitable delivery of care for patients suffering from degenerative disease of the lumbar spine.
### Figure 5-1. Correspondence Postcards

<table>
<thead>
<tr>
<th>Referral Request</th>
<th>Consultation Response</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Most Important Clinical Factors</strong>&lt;br&gt; (check all that apply)&lt;br&gt; - Leg-Dominant Pain&lt;br&gt; - Prolonged Duration of Pain&lt;br&gt; - Severity of Pain&lt;br&gt; - Frequent/constant Pain&lt;br&gt; - Neurological Symptoms&lt;br&gt; - Functional Limitation</td>
<td><strong>Surgical Treatment</strong>&lt;br&gt; - Recommended&lt;br&gt; - Possible in Future&lt;br&gt; - None</td>
</tr>
<tr>
<td><strong>Other Main Reasons for Referral</strong>&lt;br&gt; - Strong Patient Preference&lt;br&gt; - Patient Education/Reassurance&lt;br&gt; - Other _________________&lt;br&gt; - Other _________________</td>
<td><strong>If Surgery Recommended/Possible:</strong>&lt;br&gt; <strong>Most Important Clinical Factors</strong>&lt;br&gt; (check all that apply)&lt;br&gt; - Leg-Dominant Pain&lt;br&gt; - Prolonged Duration of Pain&lt;br&gt; - Severity of Pain&lt;br&gt; - Frequent/constant Pain&lt;br&gt; - Neurological Symptoms&lt;br&gt; - Functional Limitation&lt;br&gt; - Other _________________</td>
</tr>
</tbody>
</table>
Figure 5-2. The Bone and Joint Health Network’s Model of Care

* Identified non surgical patients managed through arthritis pathway
** Total LOS including acute care and in pt rehab beds

Reproduced with permission by The Bone and Joint Health Network (www.boneandjointhealthnetwork.ca).
References


Appendices

Appendix A. Questionnaires for Surgeons, Family Physicians, and Patients
Survey of Patient and Physician Attitudes Towards Lumbar Spinal Stenosis

For Surgeons

Principal Investigators:  James G. Wright, MD MPH FRCSC
Professor, Department of Surgery
Faculty of Medicine, University of Toronto

S. Samuel Bederman, MD MSc FRCSC
Research Fellow, Department of Surgery,
Faculty of Medicine, University of Toronto

Other Investigators:
Warren J. McIsaac, MD MSc CCFP FCFP
Department of Family Medicine,
Mount Sinai Hospital
Faculty of Medicine, University of Toronto

Peter C. Coyte, PhD
Department of Health Policy, Management & Evaluation
Faculty of Medicine, University of Toronto

Hans J. Kreder, MD MPH FRCSC
Division of Orthopaedic Surgery,
Sunnybrook Health Sciences Centre
Faculty of Medicine, University of Toronto

Nizar N. Mahomed, MD ScD FRCSC
Division of Orthopaedic Surgery,
University Health Network
Faculty of Medicine, University of Toronto

Institute for Social Research
York University
2007

SIGNATURE FOR CONSENT

I acknowledge that I have been informed about the questionnaire that I will complete and the research that is being conducted. I have been assured that any information relating to me will be kept confidential and that no information will be released or printed that would disclose my identity without my permission. I hereby consent to participate.

_____________________________________  ____________________________
Signature       Date
Section A. Information About Your Practice

1. Do you currently see patients in consultation for lumbar spinal stenosis?
   
   Yes 1
   No 5 → You do not need to complete the rest of the questionnaire. Please return your questionnaire in the postage-paid envelope provided.

2. In a typical month of regular office practice, about how many patients with lumbar spinal stenosis would you see?
   
   __________ patients

3. Do you currently perform spinal surgery for patients with lumbar spinal stenosis?
   
   Yes 1 → In a typical month of regular office practice, about how many spinal procedures for patients with lumbar spinal stenosis would you perform?
   No 5
   
   __________ procedures

4. Selecting appropriate patients with lumbar spinal stenosis for:
   
   Strongly disagree 1 2 3 4 5 6
   Strongly agree
   
   a. surgical referral is straightforward?
   b. surgery is straightforward?
The next section lists a series of 16 hypothetical clinical scenarios (A to P). For each scenario, please consider only the information provided and give your best opinion regarding appropriateness of surgical referral (inappropriate or appropriate) and likelihood of recommending surgery (unlikely or likely). Assume that the physical examination and imaging studies support your clinical suspicion. We encourage you to return to any situation and change your response if you wish.

To help you distinguish between scenarios, these scenarios vary on the following six factors:

[1] Duration of symptoms (less than four months OR more than eight months)
[2] Amount of Pain (mild to moderate OR severe)
[3] Dominant location of pain (back OR legs)
[4] Onset of pain (only with walking/activities OR even at rest or at night)
[5] Neurologic symptoms (occasional numbness and tingling OR moderate weakness)
[6] Walking tolerance (more than six blocks OR less than two blocks)

A. A patient has more than eight months of severe pain mostly in the back. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

very inappropriate 1
very unlikely 2
very likely 6

2. Based on this information, how likely are you to recommend surgery for this patient?

very inappropriate 1
very unlikely 2
very likely 6

B. A patient has less than four months of mild to moderate pain mostly in the back. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

very inappropriate 1
very unlikely 2
very likely 6

2. Based on this information, how likely are you to recommend surgery for this patient?

very inappropriate 1
very unlikely 2
very likely 6

C. A patient has less than four months of mild to moderate pain mostly in the back. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

very inappropriate 1
very unlikely 2
very likely 6

2. Based on this information, how likely are you to recommend surgery for this patient?

very inappropriate 1
very unlikely 2
very likely 6
D. A patient has less than four months of severe pain mostly in the legs. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very inappropriate</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very appropriate</td>
<td>6</td>
</tr>
</tbody>
</table>

2. Based on this information, how likely are you to recommend surgery for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

E. A patient has more than eight months of mild to moderate pain mostly in the legs. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very inappropriate</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very appropriate</td>
<td>6</td>
</tr>
</tbody>
</table>

2. Based on this information, how likely are you to recommend surgery for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

F. A patient has less than four months of mild to moderate pain mostly in the legs. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very inappropriate</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very appropriate</td>
<td>6</td>
</tr>
</tbody>
</table>

2. Based on this information, how likely are you to recommend surgery for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

G. A patient has more than eight months of severe pain mostly in the back. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very inappropriate</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very appropriate</td>
<td>6</td>
</tr>
</tbody>
</table>

2. Based on this information, how likely are you to recommend surgery for this patient?

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>very likely</td>
<td>6</td>
</tr>
</tbody>
</table>
H. A patient has less than four months of severe pain mostly in the back. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   very inappropriate 1
   2
   3
   4
   5
   very appropriate 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   very unlikely 1
   2
   3
   4
   5
   very likely 6

I. A patient has more than eight months of severe pain mostly in the legs. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   very inappropriate 1
   2
   3
   4
   5
   very appropriate 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   very unlikely 1
   2
   3
   4
   5
   very likely 6

J. A patient has less than four months of severe pain mostly in the legs. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   very inappropriate 1
   2
   3
   4
   5
   very appropriate 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   very unlikely 1
   2
   3
   4
   5
   very likely 6

K. A patient has less than four months of severe pain mostly in the back. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   very inappropriate 1
   2
   3
   4
   5
   very appropriate 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   very unlikely 1
   2
   3
   4
   5
   very likely 6
L. A patient has less than four months of mild to moderate pain mostly in the legs. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   - very inappropriate
   - 1
   - 2
   - 3
   - 4
   - 5
   - very appropriate
   - 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   - very unlikely
   - 1
   - 2
   - 3
   - 4
   - 5
   - very likely
   - 6

M. A patient has more than eight months of mild to moderate pain mostly in the legs. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   - very inappropriate
   - 1
   - 2
   - 3
   - 4
   - 5
   - very appropriate
   - 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   - very unlikely
   - 1
   - 2
   - 3
   - 4
   - 5
   - very likely
   - 6

N. A patient has more than eight months of mild to moderate pain mostly in the back. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   - very inappropriate
   - 1
   - 2
   - 3
   - 4
   - 5
   - very appropriate
   - 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   - very unlikely
   - 1
   - 2
   - 3
   - 4
   - 5
   - very likely
   - 6

O. A patient has more than eight months of mild to moderate pain mostly in the back. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   - very inappropriate
   - 1
   - 2
   - 3
   - 4
   - 5
   - very appropriate
   - 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   - very unlikely
   - 1
   - 2
   - 3
   - 4
   - 5
   - very likely
   - 6
P. A patient has more than eight months of severe pain mostly in the legs. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how appropriate is surgical referral for this patient?

   very inappropriate 1
   2
   3
   4
   5
   very appropriate 6

2. Based on this information, how likely are you to recommend surgery for this patient?

   very unlikely 1
   2
   3
   4
   5
   very likely 6

Q. Are there any other important factors not included in these vignettes, apart from further imaging, that you consider equally or more important in your surgical decision making for patients with lumbar spinal stenosis?

_____________________________________________________________________________
_____________________________________________________________________________

Section C. Background Information

1. Year of postgraduate medical training completed: ___ ___ ___ ___

2. Medical School of Graduation:

   North-American 1
   Foreign 5

3. Subspecialty Fellowship training in Spinal Surgery

   Yes 1
   No 5

4. Age, in years: ___ ___

5. Gender:

   Female 1
   Male 5

6. Are you affiliated with a university?

   Yes 1
   No 5

Thank you for your participation!

Please SIGN and return your completed questionnaire to the Institute for Social Research in the postage-paid envelope provided.
Survey of Patient and Physician Attitudes Towards Lumbar Spinal Stenosis

For Family Physicians

Principal Investigators:  
James G. Wright, MD MPH FRCSC  
Professor, Department of Surgery  
Faculty of Medicine, University of Toronto  

S. Samuel Bederman, MD MSc FRCSC  
Research Fellow, Department of Surgery,  
Faculty of Medicine, University of Toronto

Other Investigators:

Warren J. McIsaac, MD MSc CCFP FCFP  
Department of Family Medicine,  
Mount Sinai Hospital  
Faculty of Medicine, University of Toronto

Peter C. Coyte, PhD  
Department of Health Policy, Management & Evaluation  
University of Toronto

Hans J. Kreder, MD MPH FRCSC  
Division of Orthopaedic Surgery,  
Sunnybrook Health Sciences Centre  
Faculty of Medicine, University of Toronto

Nizar N. Mahomed, MD ScD FRCSC  
Division of Orthopaedic Surgery,  
University Health Network  
Faculty of Medicine, University of Toronto

Institute for Social Research  
York University  
2007

SIGNATURE FOR CONSENT

I acknowledge that I have been informed about the questionnaire that I will complete and the research that is being conducted. I have been assured that any information relating to me will be kept confidential and that no information will be released or printed that would disclose my identity without my permission.

I hereby consent to participate.

_____________________________________  ____________________________
Signature       Date
Section A. Information About Your Practice

1. Do you currently see patients in your practice with lumbar spinal stenosis?
   Yes 1
   No 5 → You do not need to complete the rest of the questionnaire.
   Please return your questionnaire in the postage-paid envelope provided.

2. In a typical month of regular office practice, about how many patients with existing or new symptoms of lumbar spinal stenosis did you see?
   __________ patients

3. Selecting patients with lumbar spinal stenosis for:
   Strongly agree | Strongly disagree
   a. surgical referral is straightforward?  1 2 3 4 5 6
   b. surgery is straightforward?  1 2 3 4 5 6

4. Do you currently refer patients with lumbar spinal stenosis to surgeons?
   Yes 1 → Which surgical specialty do you refer to most often?
   Orthopaedics 1
   Neurosurgery 2
   No 5

5. Which other specialists do you commonly refer patients with lumbar spinal stenosis to?
   (circle all that apply)
   Neurologist 1
   Rheumatologist 1
   Physiatrist 1
   General Internist 1
   Other Family Physician 1
   Chiropractor 1
   Pain Specialist 1
   Other (please specify) ____________ 1
Section B. Patient Vignettes

The next section lists a series of 16 hypothetical clinical scenarios (A to P). For each scenario, please consider only the information provided and give your best opinion regarding likelihood of initiating surgical referral (unlikely or likely) and appropriateness for surgery (inappropriate or appropriate). Assume that the physical examination and imaging studies support your clinical suspicion. We encourage you to return to any situation and change your response if you wish.

To help you distinguish between scenarios, these scenarios vary on the following six factors:

1. Duration of symptoms (less than four months OR more than eight months)
2. Amount of Pain (mild to moderate OR severe)
3. Dominant location of pain (back OR legs)
4. Onset of pain (only with walking/activities OR even at rest or at night)
5. Neurologic symptoms (occasional numbness and tingling OR moderate weakness)
6. Walking tolerance (more than six blocks OR less than two blocks)

A. A patient has more than eight months of severe pain mostly in the back. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

- very unlikely
- 1
- 2
- 3
- 4
- 5
- very likely
- 6

2. Based on this information, how appropriate do you think surgery is for this patient?

- very inappropriate
- 1
- 2
- 3
- 4
- 5
- very appropriate
- 6

B. A patient has less than four months of mild to moderate pain mostly in the back. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

- very unlikely
- 1
- 2
- 3
- 4
- 5
- very likely
- 6

2. Based on this information, how appropriate do you think surgery is for this patient?

- very inappropriate
- 1
- 2
- 3
- 4
- 5
- very appropriate
- 6

C. A patient has less than four months of mild to moderate pain mostly in the back. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

- very unlikely
- 1
- 2
- 3
- 4
- 5
- very likely
- 6

2. Based on this information, how appropriate do you think surgery is for this patient?

- very inappropriate
- 1
- 2
- 3
- 4
- 5
- very appropriate
- 6
D. A patient has less than four months of severe pain mostly in the legs. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

- very unlikely
- 1
- 2
- 3
- 4
- 5
- very likely
- 6

2. Based on this information, how appropriate do you think surgery is for this patient?

- very inappropriate
- 1
- 2
- 3
- 4
- 5
- very appropriate
- 6

E. A patient has more than eight months of mild to moderate pain mostly in the legs. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

- very unlikely
- 1
- 2
- 3
- 4
- 5
- very likely
- 6

2. Based on this information, how appropriate do you think surgery is for this patient?

- very inappropriate
- 1
- 2
- 3
- 4
- 5
- very appropriate
- 6

F. A patient has less than four months of mild to moderate pain mostly in the legs. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

- very unlikely
- 1
- 2
- 3
- 4
- 5
- very likely
- 6

2. Based on this information, how appropriate do you think surgery is for this patient?

- very inappropriate
- 1
- 2
- 3
- 4
- 5
- very appropriate
- 6

G. A patient has more than eight months of severe pain mostly in the back. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

- very unlikely
- 1
- 2
- 3
- 4
- 5
- very likely
- 6

2. Based on this information, how appropriate do you think surgery is for this patient?

- very inappropriate
- 1
- 2
- 3
- 4
- 5
- very appropriate
- 6
H. A patient has less than four months of severe pain mostly in the back. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

I. A patient has more than eight months of severe pain mostly in the legs. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

J. A patient has less than four months of severe pain mostly in the legs. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

K. A patient has less than four months of severe pain mostly in the back. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6
L. A patient has less than four months of mild to moderate pain mostly in the legs. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

very unlikely 1
2
3
4
5
very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?

very unlikely 1
2
3
4
5
very likely 6

M. A patient has more than eight months of mild to moderate pain mostly in the legs. They have pain even at rest or at night. There is some moderate weakness in the legs. They can walk more than six blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

very unlikely 1
2
3
4
5
very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?

very unlikely 1
2
3
4
5
very likely 6

N. A patient has more than eight months of mild to moderate pain mostly in the back. They have pain even at rest or at night. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

very unlikely 1
2
3
4
5
very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?

very unlikely 1
2
3
4
5
very likely 6

O. A patient has more than eight months of mild to moderate pain mostly in the back. They have pain only with walking or activities. There is some moderate weakness in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?

very unlikely 1
2
3
4
5
very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?

very unlikely 1
2
3
4
5
very likely 6
P. A patient has more than eight months of severe pain mostly in the legs. They have pain only with walking or activities. There is occasional numbness and tingling in the legs. They can walk less than two blocks.

1. Based on this information, how likely are you to initiate surgical referral for this patient?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

2. Based on this information, how appropriate do you think surgery is for this patient?
   - very inappropriate 1
   - 2
   - 3
   - 4
   - 5
   - very appropriate 6

Q. Are there any other important factors not included in these vignettes, apart from further imaging, that you consider equally or more important in deciding about surgical referral for patients with lumbar spinal stenosis?

_____________________________________________________________________________
_____________________________________________________________________________
_____________________________________________________________________________

Section C. Background Information

1. Year of postgraduate medical training completed: __ __ __ __

2. Medical School of Graduation:
   - North-American 1
   - Foreign 5

3. Age, in years: ___ ___

4. Gender:
   - Female 1
   - Male 5

5. Do you work in a university setting or a teaching hospital?
   - Yes 1
   - No 5

Thank you for your participation!

Please SIGN and return your completed questionnaire to the Institute for Social Research in the postage-paid envelope provided.
Survey of Patient and Physician Attitudes Towards Lumbar Spinal Stenosis

For Patients

Principal Investigators:

James G. Wright, MD MPH FRCSC
Professor, Department of Surgery
Faculty of Medicine, University of Toronto

S. Samuel Bederman, MD MSc FRCSC
Research Fellow, Department of Surgery,
Faculty of Medicine, University of Toronto

Other Investigators:

Warren J. McIsaac, MD MSc CCFP FCFP
Department of Family Medicine,
Mount Sinai Hospital
Faculty of Medicine, University of Toronto

Hans J. Kreder, MD MPH FRCSC
Division of Orthopaedic Surgery,
Sunnybrook Health Sciences Centre
Faculty of Medicine, University of Toronto

Peter C. Coyte, PhD
Department of Health Policy, Management & Evaluation
Faculty of Medicine, University of Toronto

Nizar N. Mahomed, MD ScD FRCSC
Division of Orthopaedic Surgery,
University Health Network
Faculty of Medicine, University of Toronto

Institute for Social Research
York University
2007
Section A. Information About Your Condition

1. Have you ever suffered from daily back or leg pain lasting more than two months?
   - Yes 1
   - No 5   → You do not need to complete the rest of the questionnaire.
     Please return your questionnaire in the postage-paid envelope provided.

2. Regarding your chronic back or leg pain, deciding with your family doctor about:

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  2  3  4  5  6</td>
<td></td>
</tr>
</tbody>
</table>

   a. surgical referral is straightforward?
   b. undergoing surgery is straightforward?

3. Which specialist would you most prefer to see for your condition?

   - Neurologist (doctor specialized in diseases of nervous system) 1
   - Neurosurgeon (surgeon specialized in surgery on nerves, spinal cord, and brain) 1
   - Rheumatologist (doctor specialized in rheumatic diseases, including arthritis, spinal and soft tissue disorders, and chronic musculoskeletal pain syndromes) 1
   - Orthopaedic Surgeon (surgeon specialized in surgery on bones, joints, and ligaments) 1
   - Physiatrist (doctor specialized in physical medicine) 1
   - General Internist (doctor specialized in diseases affecting adults) 1
   - Pain Specialist 1
   - Other Family Physician 1
   - Chiropractor 1
   - Other (please specify) ____________ 1

4. Have you ever seen a surgeon for your condition?
   - No 1
   - Yes 5   → Which surgical specialty did you see?
     - Orthopaedic Surgeon 1
     - Neurosurgeon 2
   - Have you ever had spinal surgery?
     - No 1
     - Yes 2
5. Please describe which option most closely describes your previous or current situation related to your back or leg pain for each of the following six factors.

A. How long you have had your pain?
   - Less than four months 1
   - More than eight months 2

B. How bad is your pain?
   - Mild to moderate 1
   - Severe 2

C. Where is your pain worst?
   - Back 1
   - Legs 2

D. When do you have your pain?
   - Only with walking/activity 1
   - Even at rest or at night 2

E. Nerve problems in your legs:
   - Occasional numbness/tingling 1
   - Moderate leg weakness 2

F. How far can you walk?
   - More than six blocks 1
   - Less than two blocks 2

Section B. Pretend Situations

The next section lists a series of 16 pretend situations (A to P). For each situation, please imagine that it describes your current condition and give your best opinion about your likelihood of wanting to be referred to a surgeon (unlikely or likely) and your likelihood of considering surgery (unlikely or likely). Imagine that other non-surgical methods (pain-killers, physiotherapy, chiropractic, etc.) are no longer helping. The situations are not given in any specific order. We encourage you to return to any situation and change your response if you wish.

A. You have more than eight months of severe pain mostly in the back. The pain is there even at rest or at night. You have occasional numbness and tingling in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

B. You have less than four months of mild to moderate pain mostly in the back. The pain is there only with walking or activities. You have some moderate weakness in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?
   - very unlikely 1
   - 2
   - 3
   - 4
   - 5
   - very likely 6
C. You have less than four months of mild to moderate pain mostly in the back. The pain is there only with walking or activities. You have occasional numbness and tingling in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likely Level</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

D. You have less than four months of severe pain mostly in the legs. The pain is there only with walking or activities. You have occasional numbness and tingling in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likely Level</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

E. You have more than eight months of mild to moderate pain mostly in the legs. The pain is there only with walking or activities. You have occasional numbness and tingling in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likely Level</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

F. You have less than four months of mild to moderate pain mostly in the legs. The pain is there even at rest or at night. You have occasional numbness and tingling in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likely Level</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>
G. You have more than eight months of severe pain mostly in the back. The pain is there only with walking or activities. You have some moderate weakness in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?
2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

H. You have less than four months of severe pain mostly in the back. The pain is there even at rest or at night. You have some moderate weakness in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?
2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

I. You have more than eight months of severe pain mostly in the legs. The pain is there even at rest or at night. You have some moderate weakness in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?
2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>

J. You have less than four months of severe pain mostly in the legs. The pain is there only with walking or activities. You have some moderate weakness in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?
2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very unlikely</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Very likely</td>
<td>6</td>
</tr>
</tbody>
</table>
163

K. You have less than four months of severe pain mostly in the back. The pain is there even at rest or at night. You have occasional numbness and tingling in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

L. You have less than four months of mild to moderate pain mostly in the legs. The pain is there even at rest or at night. You have some moderate weakness in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

M. You have more than eight months of mild to moderate pain mostly in the legs. The pain is there even at rest or at night. You have some moderate weakness in the legs. You can walk more than six blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

N. You have more than eight months of mild to moderate pain mostly in the back. The pain is there even at rest or at night. You have occasional numbness and tingling in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

   very unlikely  1  very unlikely  1
   2
   3
   4
   5
   very likely  6

163
O. You have more than eight months of mild to moderate pain mostly in the back. The pain is there only with walking or activities. You have some moderate weakness in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

P. You have more than eight months of severe pain mostly in the legs. The pain is there only with walking or activities. You have occasional numbness and tingling in the legs. You can walk less than two blocks.

1. Based on this information, how likely would you want to be referred to a surgeon for an opinion about surgery?

2. Based on this information, how likely are you to consider surgery if the risks of surgery are low?

Q. What other factors, not included in these situations, do you consider to be equally or more important in your decision for surgical referral or to consider surgery for your condition?

_____________________________________________________________________________

Section C. Background Information

1. Age, in years: _____ _____

2. Gender: Female 1 Male 5

3. What is the highest level of education you have completed?

- Did not complete high school 1
- High School 2
- Post-secondary, non-University 3
- University (undergraduate) 4
- University (professional or graduate) 5

Thank you for your participation!
Please return your completed questionnaire to the Institute for Social Research in the postage-paid envelope provided.
Appendix B. Diagnostic and Procedural Inclusion and Exclusion Codes for Degenerative Lumbar Spinal Surgery Cohort Identification

(adapted from Deyo et al., 2005\textsuperscript{13})

CCI – Canadian Classification of Interventions

ICD-10 – International Classification of Diseases (10\textsuperscript{th} revision)

OHIP – Ontario Health Insurance Plan Procedural Billing Codes

Table A 1. Identification of Surgical Cases from CCI Code Discharges

<table>
<thead>
<tr>
<th>CCI Procedure Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.AW.72</td>
<td>Release spinal cord (includes decompression, spinal cord or nerve root)</td>
</tr>
<tr>
<td>1.SC.74</td>
<td>Fixation, spinal vertebrae</td>
</tr>
<tr>
<td>1.SC.75</td>
<td>Fusion, spinal vertebrae</td>
</tr>
<tr>
<td>1.SC.80</td>
<td>Repair, spinal vertebrae (includes laminectomy, laminoplasty, etc.)</td>
</tr>
<tr>
<td>1.SC.87</td>
<td>Excision partial, spinal vertebrae</td>
</tr>
<tr>
<td>1.SC.89</td>
<td>Excision total, spinal vertebrae</td>
</tr>
<tr>
<td>1.SE.53</td>
<td>Implantation of internal device, intervertebral disc</td>
</tr>
<tr>
<td>1.SE.87</td>
<td>Excision partial, intervertebral disc</td>
</tr>
<tr>
<td>1.SE.89</td>
<td>Excision total, intervertebral disc</td>
</tr>
</tbody>
</table>

Table A 2. ICD-10 Diagnostic Codes for Exclusions in Surgical Case Finding

<table>
<thead>
<tr>
<th>ICD-10 Diagnosis Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C00-D48</td>
<td>All neoplasms</td>
</tr>
<tr>
<td>G00-G09</td>
<td>Intraspinal abscess</td>
</tr>
<tr>
<td>O00-O99</td>
<td>Pregnancy-related diagnoses</td>
</tr>
<tr>
<td>M45-M46</td>
<td>Inflammatory spondyloarthropathies</td>
</tr>
<tr>
<td>M86</td>
<td>Osteomyelitis</td>
</tr>
<tr>
<td>M80</td>
<td>Pathologic fracture</td>
</tr>
<tr>
<td>M84</td>
<td>Non-union/mal-union of fracture</td>
</tr>
<tr>
<td>S12, S22, S32</td>
<td>Fractures of spinal column</td>
</tr>
<tr>
<td>S13.1, S23.1, S33.1</td>
<td>All vertebral dislocations</td>
</tr>
<tr>
<td>V01-X59</td>
<td>Accidents</td>
</tr>
</tbody>
</table>

A case with any of these codes is excluded, no matter what other code may be present.

Exclude all patients <50 years of age as of the listed discharge date. Then exclude if listed in any diagnosis field:

<table>
<thead>
<tr>
<th>ICD-10 Diagnosis Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>G54.2, G54.3</td>
<td>Cervical/thoracic root lesion</td>
</tr>
<tr>
<td>M47.81-M47.84</td>
<td>Cervical/thoracic spondylosis without myelopathy</td>
</tr>
<tr>
<td>M47.11-M47.14</td>
<td>Cervical/thoracic spondylosis with myelopathy</td>
</tr>
<tr>
<td>M47.21-M47.24</td>
<td>Cervical/thoracic spondylosis with radiculopathy</td>
</tr>
<tr>
<td>M50, M51.21-M51.24</td>
<td>Displacement cervical disc</td>
</tr>
<tr>
<td>M51.31-M51.34</td>
<td>Degeneration of cervical/thoracic disc</td>
</tr>
<tr>
<td>M51.01-M51.04</td>
<td>Cervical/thoracic disc disease with myelopathy</td>
</tr>
<tr>
<td>M51.11-M51.14</td>
<td>Cervical/thoracic disc disease with radiculopathy</td>
</tr>
<tr>
<td>M96.11-M96.14</td>
<td>Cervical/thoracic disc post laminectomy syndrome</td>
</tr>
<tr>
<td>M51.91-M51.94</td>
<td>Unspecified disc disorder, cervical/thoracic</td>
</tr>
<tr>
<td>M51.81-M51.84</td>
<td>Other specified intervertebral disc disorder, cervical/thoracic</td>
</tr>
<tr>
<td>M48.01-M48.04, M54.2</td>
<td>Cervical/Thoracic spinal stenosis</td>
</tr>
<tr>
<td>M53.1</td>
<td>Brachial neuritis</td>
</tr>
</tbody>
</table>

Exclude if listed as main diagnosis; do not exclude if it is a secondary diagnosis.
Table A 3. OHIP Procedural Codes for Exclusions in Surgical Case Finding

A case with any of these codes is excluded, no matter what other code may be present.

Exclude if listed in *any* OHIP procedure field:

<table>
<thead>
<tr>
<th>OHIP Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fractures</strong></td>
<td></td>
</tr>
<tr>
<td>E562</td>
<td>Reduction - Fractures or Fracture Dislocations - Reapplication of Halo traction (IOP) - counter traction pins or vest</td>
</tr>
<tr>
<td>F103</td>
<td>Reduction - Fractures or Fracture Dislocations - Closed reduction</td>
</tr>
<tr>
<td>F105</td>
<td>Reduction - Fractures or Fracture Dislocations - Open reduction - posterior approach</td>
</tr>
<tr>
<td>F107</td>
<td>Reduction - Fractures or Fracture Dislocations - Open reduction - anterior approach</td>
</tr>
<tr>
<td>E913</td>
<td>Posterior, Anterior, Anterolateral or Posterolateral Spinal Decompressive Procedures - With spinal cord injury</td>
</tr>
<tr>
<td>R494</td>
<td>Arthrodesis - Fusion by different surgeon - two or more levels</td>
</tr>
<tr>
<td>R493</td>
<td>Arthrodesis - Fusion with other procedure(s) by different surgeon - one level</td>
</tr>
<tr>
<td>N126</td>
<td>Intra-oral approach to lesions of the skull base and upper cervical spine</td>
</tr>
<tr>
<td>F103</td>
<td>Reduction - Fractures or Fracture Dislocations - Closed reduction</td>
</tr>
<tr>
<td>N572</td>
<td>Open reduction, any single level, spine fracture/dislocation,</td>
</tr>
<tr>
<td>N573</td>
<td>Anterior odontoid screw fixation</td>
</tr>
<tr>
<td>N570</td>
<td>Vertebroplasty (injection of bone cement) as sole procedure,</td>
</tr>
<tr>
<td>N583</td>
<td>Kyphoplasty (balloon tamp and injection of bone cement)</td>
</tr>
<tr>
<td><strong>Cervical/Thoracic</strong></td>
<td></td>
</tr>
<tr>
<td>E924</td>
<td>Arthrodesis - Fusion by different surgeon – anterior cervical interbody fusion, per level</td>
</tr>
<tr>
<td>E929</td>
<td>Arthrodesis, Fusions - Anterior spinal cervical interbody fusion - per level</td>
</tr>
<tr>
<td></td>
<td>Decompression - Anterior, Anterolateral or Posterolateral – Anterior cervical spinal cord or nerve root decompression, including removal of disc or vertebral body, single disc level</td>
</tr>
<tr>
<td>N182</td>
<td>Decompression - Anterior, Anterolateral or Posterolateral – Simple anterior cervical discectomy</td>
</tr>
<tr>
<td>R419</td>
<td>Decompression - Posterior - Cervical hemilaminectomy for disc disease, with or without foraminotomy</td>
</tr>
<tr>
<td>R447</td>
<td>Disc excision (one level)</td>
</tr>
<tr>
<td>N500</td>
<td>Disc excision (one level)</td>
</tr>
<tr>
<td>N501</td>
<td>Vertebrectomy (removal of vertebral body and excision of adjacent discs)</td>
</tr>
<tr>
<td>N502</td>
<td>Disc excision (one level)</td>
</tr>
<tr>
<td>N503</td>
<td>Vertebrectomy (removal of vertebral body and excision of adjacent discs)</td>
</tr>
<tr>
<td>N504</td>
<td>Disc excision (one level)</td>
</tr>
<tr>
<td>N505</td>
<td>Vertebrectomy (removal of vertebral body and excision of adjacent discs)</td>
</tr>
<tr>
<td>N569</td>
<td>Anterior cervical decompression by intra-oral approach</td>
</tr>
<tr>
<td>N516</td>
<td>One disc level</td>
</tr>
<tr>
<td>N517</td>
<td>One disc level</td>
</tr>
<tr>
<td>N518</td>
<td>One disc level</td>
</tr>
<tr>
<td>N509</td>
<td>One level - unilateral</td>
</tr>
<tr>
<td>N510</td>
<td>One level - bilateral</td>
</tr>
<tr>
<td>N520</td>
<td>One level - laminoplasty (includes fixation of lamina)</td>
</tr>
<tr>
<td>N528</td>
<td>C1/C2 screw fixation (transarticular, pedicle, lateral mass)</td>
</tr>
<tr>
<td>N513</td>
<td>One disc level - below C2</td>
</tr>
<tr>
<td>N519</td>
<td>C1/C2 fusion using graft/posterior wires</td>
</tr>
<tr>
<td>N514</td>
<td>One disc level - below C2</td>
</tr>
<tr>
<td>N532</td>
<td>C1/C2 screw fixation (transarticular, pedicle, lateral mass)</td>
</tr>
<tr>
<td>N515</td>
<td>One disc level - below C2</td>
</tr>
</tbody>
</table>

Table continued on next page …
Exclude if listed in *any* OHIP procedure field:

<table>
<thead>
<tr>
<th>OHIP Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tumours</strong></td>
<td></td>
</tr>
<tr>
<td>N313</td>
<td>Removal of spinal tumour by anterior or anterolateral cervical or anterior,</td>
</tr>
<tr>
<td></td>
<td>anterolateral or posterolateral thoracic or lumbar approach - two surgeons - excision</td>
</tr>
<tr>
<td>N314</td>
<td>Removal of spinal tumour by anterior or anterolateral cervical or anterior,</td>
</tr>
<tr>
<td></td>
<td>anterolateral or posterolateral thoracic or lumbar approach - one surgeon</td>
</tr>
<tr>
<td>N317</td>
<td>Extradural partial or total removal by laminectomy (including posterolateral cervical approach)</td>
</tr>
<tr>
<td>N318</td>
<td>Intradural (extramedullary) partial or total removal of spinal tumour</td>
</tr>
<tr>
<td>N319</td>
<td>Intramedullary - biopsy and/or decompression</td>
</tr>
<tr>
<td>N320</td>
<td>Intramedullary - removal</td>
</tr>
<tr>
<td>N321</td>
<td>A.V. malformation of cord - Excision or operative obliteration; with or without evacuation of haematoma</td>
</tr>
<tr>
<td>R634</td>
<td>Excision - Muscle/Soft Tissue - Tumours - simple</td>
</tr>
<tr>
<td>R635</td>
<td>Excision - Muscle/Soft Tissue - Tumours - radical resection</td>
</tr>
<tr>
<td>R993</td>
<td>Exc. of tumours not specifically listed</td>
</tr>
<tr>
<td>E978</td>
<td>A.V. malformation of cord - Excision or operative obliteration; with or without evacuation of haematoma - per segment after three segments</td>
</tr>
<tr>
<td>E900</td>
<td>Removal of spinal tumour by anterior or anterolateral cervical or anterior,</td>
</tr>
<tr>
<td></td>
<td>anterolateral or posterolateral thoracic or lumbar approach - repeat</td>
</tr>
<tr>
<td>N553</td>
<td>Simple soft tissue tumour excision under 5cm</td>
</tr>
<tr>
<td>N554</td>
<td>Radical soft tissue tumour excision 5cm and greater.</td>
</tr>
<tr>
<td>N559</td>
<td>Intradural extramedullary spinal tumour - partial or total removal.</td>
</tr>
<tr>
<td>N560</td>
<td>Intradural intramedullary spinal tumour - partial or total removal.</td>
</tr>
<tr>
<td><strong>Infections</strong></td>
<td></td>
</tr>
<tr>
<td>R251</td>
<td>Incision and Drainage (Osteomyelitis) - Bone - incision and drainage only</td>
</tr>
<tr>
<td>R252</td>
<td>Incision and Drainage (Osteomyelitis) - Saucerization with bone grafting - posterior</td>
</tr>
<tr>
<td>R270</td>
<td>Incision and Drainage (Osteomyelitis) - Saucerization with bone grafting – anterior</td>
</tr>
<tr>
<td>R254</td>
<td>Incision and Drainage (Osteomyelitis) - Sequestrectomy - posterior</td>
</tr>
<tr>
<td>R234</td>
<td>Incision and Drainage (Osteomyelitis) - Sequestrectomy - anterior</td>
</tr>
<tr>
<td>N549</td>
<td>Incision and drainage including sequestrectomy, anterior approach.</td>
</tr>
<tr>
<td>N548</td>
<td>Incision and drainage only, posterior approach.</td>
</tr>
<tr>
<td>N550</td>
<td>Sequestrectomy, posterior approach.</td>
</tr>
</tbody>
</table>
### Table A 4. Diagnosis Codes for “Definite” Lumbar Surgery

A case with any ‘definite’ diagnosis was used to generate OHIP procedure lists for exploration and subsequent categorization into Fusion (incl. deformity correction) and Decompression (incl. discectomy).

<table>
<thead>
<tr>
<th>Diagnosis Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Herniated Disc</strong></td>
<td></td>
</tr>
<tr>
<td>M51.25-M51.28</td>
<td>Displacement of lumbar disc</td>
</tr>
<tr>
<td>M51.05-M51.08</td>
<td>Herniated lumbar disc with myelopathy</td>
</tr>
<tr>
<td>M51.15-M51.18</td>
<td>Herniated lumbar disc with radiculopathy</td>
</tr>
<tr>
<td><strong>Disc Degeneration</strong></td>
<td></td>
</tr>
<tr>
<td>M47.85-M47.88</td>
<td>Lumbosacral spondylosis, no myelopathy</td>
</tr>
<tr>
<td>M51.35-M51.38</td>
<td>Degeneration of lumbar disc</td>
</tr>
<tr>
<td>M51.95-M51.98</td>
<td>Lumbar disc calcification</td>
</tr>
<tr>
<td><strong>Spinal Stenosis</strong></td>
<td></td>
</tr>
<tr>
<td>M47.15-M47.18</td>
<td>Spondylogetic compression of lumbar spinal cord (with myelopathy)</td>
</tr>
<tr>
<td>M47.25-M47.28</td>
<td>Other lumbar spondylosis with radiculopathy</td>
</tr>
<tr>
<td>M48.05-M48.08</td>
<td>Lumbar stenosis</td>
</tr>
<tr>
<td>M47.95-M47.98</td>
<td>Spondylolisthesis, unspecified, lumbar</td>
</tr>
<tr>
<td><strong>Possible Instability</strong></td>
<td></td>
</tr>
<tr>
<td>M53.25-M53.28</td>
<td>Disorders of sacrum: includes instability of lumbosacral joint</td>
</tr>
<tr>
<td>M43.15-M43.18</td>
<td>Acquired spondylolisthesis (included because these are overwhelmingly lumbar)</td>
</tr>
<tr>
<td>M43.05-M43.08</td>
<td>Spondylolysis/Spondylolisthesis, lumbar</td>
</tr>
<tr>
<td><strong>Miscellaneous low back problems</strong></td>
<td></td>
</tr>
<tr>
<td>M51.45-M51.48</td>
<td>Schmorl's node, Lumbar region</td>
</tr>
<tr>
<td>M96.15-M96.18</td>
<td>Postlaminectomy syndrome, lumbar</td>
</tr>
<tr>
<td>M54.4,M54.5</td>
<td>Lumbago</td>
</tr>
<tr>
<td>M54.3</td>
<td>Sciatica</td>
</tr>
<tr>
<td>M99.(0-9)3</td>
<td>Non-allopathic lesions, lumbar spine</td>
</tr>
<tr>
<td>M99.(0-9)4</td>
<td>Non-allopathic lesions, sacral region</td>
</tr>
<tr>
<td>S33.7</td>
<td>Sprains and strains, lumbosacral and other sacral ligaments</td>
</tr>
<tr>
<td>S33.5</td>
<td>Sprains and strains, lumbar/sacral</td>
</tr>
<tr>
<td>M40.(1-5)(5-8)</td>
<td>Kyphosis and Lordosis, lumbar</td>
</tr>
<tr>
<td>M41.(5,8)(5-8)</td>
<td>Other secondary scoliosis, lumbar</td>
</tr>
<tr>
<td>M43.25-M43.28</td>
<td>Other fusion of spine, lumbar</td>
</tr>
<tr>
<td>M43.55-M43.58</td>
<td>Other recurrent vertebral subluxation, lumbar</td>
</tr>
<tr>
<td>M53.85-M53.88</td>
<td>Other specified dorsopathies, lumbar</td>
</tr>
<tr>
<td>M53.95-M53.98</td>
<td>Dorsopathy, unspecified, lumbar</td>
</tr>
<tr>
<td>M54.15-M54.18</td>
<td>Radiculopathy, lumbar</td>
</tr>
<tr>
<td>M54.85-M54.88</td>
<td>Other dorsalgia, lumbar</td>
</tr>
<tr>
<td>M54.95-M54.98</td>
<td>Dorsalgia, unspecified, lumbar</td>
</tr>
</tbody>
</table>
Table A 5. Diagnosis codes for “Possible” Lumbar Surgery

A case with a procedure code classified as either Fusion (Error! Reference source not found.) or Decompression (Error! Reference source not found.), and any diagnosis listed below is selected. All other cases without diagnoses in Error! Reference source not found. or Error! Reference source not found. are excluded.

<table>
<thead>
<tr>
<th>ICD-10 Diagnosis Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>M48.25-M48.28</td>
<td>Kissing spine</td>
</tr>
<tr>
<td>M47.89</td>
<td>Spondylosis, unspecified site</td>
</tr>
<tr>
<td>M51.29</td>
<td>Herniated disc, unspecified site</td>
</tr>
<tr>
<td>M51.49</td>
<td>Schmorl's node, site unspecified</td>
</tr>
<tr>
<td>M51.39</td>
<td>Degeneration intervertebral disc, unspecified site</td>
</tr>
<tr>
<td>M51.09</td>
<td>Intervertebral disc disorder with myelopathy, site unspecified</td>
</tr>
<tr>
<td>M51.19</td>
<td>Intervertebral disc disorder with radiculopathy, site unspecified</td>
</tr>
<tr>
<td>M96.19</td>
<td>Post-laminectomy syndrome, unspecified region</td>
</tr>
<tr>
<td>M48.09</td>
<td>Stenosis, unspecified site, not cervical</td>
</tr>
<tr>
<td>M47.89</td>
<td>Stenosis, other, not cervical</td>
</tr>
<tr>
<td>M79.25-M79.28</td>
<td>Thoracic or lumbosacral neuritis or radiculitis</td>
</tr>
<tr>
<td>M54.9</td>
<td>Backache, unspecified</td>
</tr>
<tr>
<td>M79.3</td>
<td>Other symptoms referable to back; Other unspecified back disorders; Sprain and strain, unspecified part of back</td>
</tr>
<tr>
<td>M51.(0-4,9),</td>
<td>All “Definite” Lumbar Surgery codes truncated to 3 numeric digits.</td>
</tr>
<tr>
<td>M47.(1,2,8,9),</td>
<td></td>
</tr>
<tr>
<td>M48.0, M53.(2,8,9),</td>
<td></td>
</tr>
<tr>
<td>M43.(0,1,2,5)</td>
<td></td>
</tr>
<tr>
<td>M96.1, M54.(1,8,9),</td>
<td></td>
</tr>
<tr>
<td>M99, M40.(1-5)</td>
<td></td>
</tr>
<tr>
<td>M41.(5,8)</td>
<td></td>
</tr>
</tbody>
</table>
Table A 6. “Fusion” OHIP Procedure Codes, identified by “Definite” Lumbar Surgery Cohort

A case with any ‘Fusion’ code is classified as ‘Fusion’ and included. Also includes deformity correction.

<table>
<thead>
<tr>
<th>OHIP Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E548</td>
<td>Fusion with other procedure(s), and/or by different surgeon - with instrumentation including first two levels of fixation</td>
</tr>
<tr>
<td>E567</td>
<td>Fusion with other procedure(s) - by same surgeon – one level</td>
</tr>
<tr>
<td>E568</td>
<td>Fusion with other procedure(s) - by same surgeon – two or more levels</td>
</tr>
<tr>
<td>E574</td>
<td>Fusion by different surgeon - repeat fusion, to any fusion,</td>
</tr>
<tr>
<td>R459</td>
<td>Anterior or posterior fusion of one level</td>
</tr>
<tr>
<td>R493</td>
<td>Fusion with other procedure(s) by different surgeon - one level</td>
</tr>
<tr>
<td>R494</td>
<td>Fusion by different surgeon - two or more levels</td>
</tr>
<tr>
<td>R336</td>
<td>Revision of entire instrumentation - without fusion</td>
</tr>
<tr>
<td>R346</td>
<td>Revision of entire instrumentation - with fusion</td>
</tr>
<tr>
<td>R362</td>
<td>Posterior (Harrington) - with or without fusion</td>
</tr>
<tr>
<td>R369</td>
<td>Readjustment of instrumentation</td>
</tr>
<tr>
<td>R371</td>
<td>Segmental procedure - with fusion</td>
</tr>
<tr>
<td>R271</td>
<td>Osteotomy (includes fixation/fusion) - Anterior – via chest and abdomen</td>
</tr>
<tr>
<td>R296</td>
<td>Osteotomy (includes fixation/fusion) – Posterior</td>
</tr>
<tr>
<td>R303</td>
<td>Osteotomy (includes fixation/fusion) - Anterior – via chest</td>
</tr>
<tr>
<td>R310</td>
<td>Osteotomy (includes fixation/fusion) – Circumferential</td>
</tr>
<tr>
<td>R366</td>
<td>Osteotomy (includes fixation/fusion) - Anterior – via abdomen</td>
</tr>
<tr>
<td>E367</td>
<td>one disc level (as secondary procedure)</td>
</tr>
<tr>
<td>N581</td>
<td>One disc level</td>
</tr>
<tr>
<td>E365</td>
<td>one disc level (as secondary procedure)</td>
</tr>
<tr>
<td>N526</td>
<td>Artificial disc insertion (includes approach)</td>
</tr>
<tr>
<td>N525</td>
<td>Artificial disc insertion (approach by separate surgeon)</td>
</tr>
<tr>
<td>N559</td>
<td>One disc level</td>
</tr>
<tr>
<td>N580</td>
<td>One disc level</td>
</tr>
<tr>
<td>E370</td>
<td>one disc level (as secondary procedure)</td>
</tr>
<tr>
<td>E387</td>
<td>fusion to sacrum</td>
</tr>
<tr>
<td>N513</td>
<td>One disc level</td>
</tr>
<tr>
<td>E372</td>
<td>one disc level (as secondary procedure)</td>
</tr>
<tr>
<td>N582</td>
<td>One disc level</td>
</tr>
<tr>
<td>N533</td>
<td>Pars reconstruction for spondylolysis.</td>
</tr>
<tr>
<td>N539</td>
<td>Anterior scoliosis correction - any number of levels</td>
</tr>
<tr>
<td>N540</td>
<td>Posterior scoliosis correction - up to six levels</td>
</tr>
<tr>
<td>N574</td>
<td>Osteotomy - Above cord and conus (includes partial rib resection)</td>
</tr>
<tr>
<td>N575</td>
<td>Osteotomy - Below conus - each level</td>
</tr>
<tr>
<td>N576</td>
<td>Smith Peterson Osteotomy - each level.</td>
</tr>
</tbody>
</table>
Table A 7. “Decompression” OHIP Procedure Codes, identified by “Definite” Lumbar Surgery Cohort

A case not classified as ‘Fusion’ (Error! Reference source not found.) with any ‘Decompression’ code is classified as ‘Decompression’ and included. If not classified as either ‘Fusion’ or ‘Decompression’ then excluded.

<table>
<thead>
<tr>
<th>OHIP Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N185</td>
<td>Posterior laminectomy one or two levels, cervical, thoracic, lumbar</td>
</tr>
<tr>
<td></td>
<td>Anterior, Anterolateral or Posterolateral - Anterolateral or posterolateral decompression, lumbar or thoracic spine, single disc level</td>
</tr>
<tr>
<td>N186</td>
<td>Posterior - Repeat posterior exploration or reopening of posterior exploration, more than six months after original procedure, includes foraminotomy, discectomy or neurolysis</td>
</tr>
<tr>
<td>N337</td>
<td>Lumbar hemilaminectomy for disc disease including removal of soft disc or osteophyte</td>
</tr>
<tr>
<td>R370</td>
<td>Percutaneous discectomy</td>
</tr>
<tr>
<td>R452</td>
<td>Anterior, Anterolateral or Posterolateral – Simple anterior lumbar discectomy</td>
</tr>
<tr>
<td>R457</td>
<td>Lumbar hemilaminectomy for disc disease including removal of soft disc or osteophyte</td>
</tr>
<tr>
<td>R373</td>
<td>Excision - Bone - Spinous process</td>
</tr>
<tr>
<td>R374</td>
<td>Excision - Bone - Lamina or transverse process</td>
</tr>
<tr>
<td>R450</td>
<td>Excision - Bone - Part of body or pedicle</td>
</tr>
<tr>
<td>N506</td>
<td>Disc excision (one level)</td>
</tr>
<tr>
<td>N507</td>
<td>Vertebrectomy (removal of vertebral body and excision of adjacent discs)</td>
</tr>
<tr>
<td>N508</td>
<td>Disc excision (one level)</td>
</tr>
<tr>
<td>N579</td>
<td>Vertebrectomy (removal of vertebral body and excision of adjacent discs)</td>
</tr>
<tr>
<td>N511</td>
<td>One level - unilateral</td>
</tr>
<tr>
<td>N512</td>
<td>One level - bilateral</td>
</tr>
<tr>
<td>N524</td>
<td>One level - bilateral canal enlargement - unilateral approach</td>
</tr>
<tr>
<td>N571</td>
<td>Percutaneous discotomy</td>
</tr>
</tbody>
</table>
Appendix C. Maps of Ontario

Figure A 1. Map of 49 Ontario Counties

Figure A 2. Map of 14 Ontario Local Health Integration Networks (LHINs)

Reproduced with permission by the Queen’s Printer for Ontario from Ontario’s Local Health Integration Networks (www.lhins.on.ca)
Copyright Acknowledgments

Figure 1-1. The Spectrum of Degenerative Disease of the Lumbar Spine (Reproduced with permission by the Publishing Division of the Massachusetts Medical Society from Deyo RA, Weinstein JN. Primary Care: Low Back Pain. *NEJM* 2001; 344:363-70).

Figure 1-2. Andersen's Behavioural Model of Health Service Use (Reproduced with allowance by the American Sociological Association from Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behavior 1995; 36:1-10).


Figure 5-2. The Bone and Joint Health Network's Model of Care (Reproduced with permission by The Bone and Joint Health Network).

Figure A 1. Map of 49 Ontario Counties (Reproduced with permission by the Ontario Ministry of Finance).

Figure A 2. Map of 14 Ontario Local Health Integration Networks (LHINs) (Reproduced with permission by the Queen’s Printer for Ontario from the Ontario Local Health Integration Networks).