IMAGING RATES AND APPROPRIATENESS FOR PATIENTS WITH LOW

BACK PAIN AS ORDERED BY PRIMARY CARE PHYSICIANS

By © Gabrielle S Logan

A thesis submitted to the School of Graduate Studies

in partial fulfillment of the requirements for the degree of Master of Science

Clinical Epidemiology, Faculty of Medicine

Memorial University of Newfoundland

October 2019

St. John's, Newfoundland and Labrador

ABSTRACT

Choosing Wisely recommends reducing unnecessary radiological imaging for low back pain. This thesis explored imaging appropriateness for low back pain compared to these recommendations in Newfoundland and Labrador and globally. A minor part included a descriptive study to provide age-sex standardised rates of lumbar spine Computed Tomography in the Eastern Health Region. The main thesis portion focused on imaging appropriateness and included two main studies; a systematic review and meta-analysis of lumbar spine x-ray and CT appropriateness, and a medical record review of lumbar spine CT referrals. The systematic review and meta-analysis, conducted using the PRISMA statement, found that 44% of x-rays and 54% of CTs were appropriate. The medical record review conducted in 2016 in Eastern Health included 3,595 lumbar spine CTs referrals. It found 5.5% were appropriate, 75.8% were potentially appropriate, and 16.8% were inappropriate. Unnecessary imaging occurs despite guidelines advising against them. Future research to understand why physicians order imaging for back pain patients and effective reduction interventions are necessary.

ACKNOWLEDGEMENTS

This thesis was made possible by a whole army of support, and without that support, I would not be here writing this acknowledgement!

My sincerest thanks to my supervisor, Dr Amanda Hall, who accepted me as a student with very little background information to go on. She provided me with more support then I could have possibly asked for and has given me a solid foundation to which I will be able to build a career. I am deeply grateful and indebted to her for generosity and kindness.

Many thanks are also due to Dr Holly Etchegary, my committee member, whose energy and passion were an inspiration and to Dr Patrick Parfrey, my other committee member, for providing me with this opportunity. I also wish to acknowledge Yvonne Thibault, Clin Epi Secretary and wizard, for treating me like family when I needed it the most and your behind-the-scenes efforts that keep Clin Epi together. To all the other faculty in Clin Epi, thank you for your teaching, mentorship, and care. I am a better person for having been around you. Another thank you is owed to the NL Support Team who helped me with logistical issues and data collection for this thesis! A specific "thank you" goes to Tammy Benteau and Amy Roebothan for helping me transcribe referrals.

I have made some amazing friends while at Memorial University of Newfoundland, both at the Faculty of Medicine, and Bitters. Thank you to you all who made time in your lives for this CFA!

I also want to thank my parents, Murray and Dolores Logan, for raising me, loving me, and for providing the framework from which I was able to grow from. Thank you to my siblings, Faith, Heinrich, Alex, Simon, and Andrew, you all believed in me. I love you so much! And love to my niece, Eshné, and nephew, Finntan, who remind me to experience joy and for whom I wish to build a better future.

Finally, this thesis is dedicated to my heart, my life, and now husband (!!), Nicholas McGoldrick, the one who coached me through all the struggles and self-doubt and frustrations. You showered me with love and buoyed me up when I did not think I could do it anymore. Your patience, intelligence, fairness, and sincerity are an inspiration. My world revolves around you and always will. I am grateful for you every day.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND SYMBOLS

ALBP- Acute low back pain

BMI- Body mass index

CADTH- Canadian Agency for Drugs and Technologies in Health

CDS- Clinical Decision Support

CIHI- Canadian Institute of Health Information

CIHR-SPOR- Canadian Institutes of Health Research Strategy for Patient-

Oriented Research

CLBP- Chronic low back pain

CPG- Clinical Practice Guidelines

CT- Computed Tomography

CWC- Choosing Wisely Canada

DI- Diagnostic imaging

DALY- Disease Adjusted Life Years

ED- Emergency department

EHR- Electronic health records

EH- Eastern Health

FBSS- Failed Back Surgery Syndrome

FP- Family Physician

GP- General Practitioner

LBP- Low back pain

LS- Lumbar Spine

MRI- Magnetic Resonance Imaging

MSK- Musculoskeletal

NSAID- Non-steroid anti-inflammatory drug

NSLBP- Non-Specific Low back pain

PCORI- Patient-Centered Outcomes Research Institute

POR- Patient-Oriented Research

PRISMA- Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QALY- Quality-adjusted life years

RECORD- REporting of studies Conducted using Observational Routinely-
collected Data

RHA- Regional Health Authority

SD- Standard deviation

SR- Systematic Review

SR&MA- Systematic review and meta-analysis

STROBE- STrengthening the Reporting of OBservational studies in Epidemiology

TENS- transcutaneous electrical nerve stimulation

YLD- Years lived with disability

DECLARATION OF PUBLICATION INTENT

These chapters have not been published in a scientific journal; therefore, none of the figures or tables are copyrighted. However, all three main chapters will be submitted to journals for publication, and two are currently under review. All journals that we are seeking publication from are open access journals. The submission details are as follows:

Title: Lumbar Spine CT Referral Rates by Family Physicians using routinely collected data in One Health Region in Newfoundland and Labrador, Canada: A Brief Report
Authors: Gabrielle S Logan, Bethan Copsey, Holly Etchegary, Patrick Parfrey, Krista Mahoney, & Amanda Hall
Journal: Submitted to CMAJ Open on May 22, 2019
Manuscript ID: CMAJOpen-2019-0076

Title: What Do We Really Know About the Appropriateness of Imaging for Low Back Pain in Primary Care? A Systematic Review and Meta-Analysis of Medical Record Reviews
Authors: Gabrielle S Logan, Andrea Pike, Bethan Copsey, Patrick Parfrey, Holly Etchegary, Amanda Hall
Journal: Submitted to Plos One on June 11, 2019
Manuscript Number: bmjoq-2019-000661

Title: What are the Reasons Family Physicians Refer Patients with Low Back Pain for CTs and are They Appropriate According to the Guidelines? A Retrospective Review of 4,435 Medical Records in Newfoundland using Routinely Collected Data from Linked Databases
Authors: Gabrielle S Logan, Russell Dawe, Kris Aubrey-Bassler, Patrick Parfrey, Holly Etchegary, Amanda Hall
Journal: Formatted for publication in CMAJ and will be submitted to the journal

CHAPTER 1: INTRODUCTION AND BACKGROUND

1.1 Introduction to Low Back Pain

Often thought of as a symptom of another disease, low back pain (LBP) is an incredibly common condition that affects nearly all humans, regardless of country and income (1). Though there are many definitions of LBP, some researchers have proposed using one unified definition to allow for collaboration among researchers (2). This is still not used in all LBP research, but it is noted as pain in the lumbar region of the spine that starts at the end of the ribcage, also known as the costal margin of the back at the 12th rib, and ends at the top of the buttock at the gluteal fold (2-4). Leg pain also commonly occurs in conjunction with back pain.

The anatomical features of this area include muscle, circulatory vessels, lumbar vertebrae, intervertebral discs, the spinal cord, and nerves that stem from the spinal cord. Thus, LBP can be pain from any of these sources in the lumbar spine (LS). Often it is difficult for physicians to locate the source of pain to any one of these anatomical features. If it can be determined, it is often very serious in pathology. The main serious pathologies are cancer/tumour, fractures, neurological deficits, and infection, and can also include inflammatory diseases such as arthritis or ankylosing spondylitis (5,6).

LBP has many different classifications based on how long the person has been experiencing pain, and whether the cause of the pain can be diagnosed (6).

Back pain can be classified based on the length of time a patient has it, with a pain duration for less than three months being considered acute LBP (ALBP). Pain for greater than three months is called Chronic LBP (CLBP). LBP can also be classified into one of three categories based on the cause of the pain. If no specific pathology caused by the aforementioned causes can be found, the pain is often classified as non-specific low back pain (NSLBP), and if there are neurological symptoms that affect the legs, such as numbness or weakness, a patient is said to have radicular syndrome (6,7). The third category is called the serious spinal pathology category, including pain that is caused by cauda equina, infection, fracture, or cancer/tumour. Further information is provided in the section on usual care for LBP.

1.2 Epidemiology and Etiology

Low back pain is very rarely caused by a serious condition, though 80% of the population globally will experience it. Prevalence of serious pathologies such as fractures, infection, inflammatory arthritis, and cancer have been estimated to be present in less than 1% of all cases of LBP (6). Prevalence of NSLBP is much more difficult to estimate, but it is thought that the one-year prevalence of any type of LBP ranges from 0.8% to 82.5%, with a point prevalence ranging from 1% to 58% (3). A systematic review of the incidence of any episode of LBP has been estimated to be between 1.5% to 36% (3). This wide range of estimates shows that the estimates of LBP are inconsistent and still unclear, which could be due to the available data and different definitions used in different countries.

There are many pathoanatomical findings that are thought to cause back pain, such as spinal stenosis, radiculopathy, degenerative disc disease, disc herniation, and more (8,9). How pain is triggered varies from one another due to the different anatomical pathways. For example, radiculopathy is not the same as radiating pain. Radiculopathy usually is the involvement of just one of the nerve roots in the spine and is associated with numbness and weakness, while radicular pain is not caused by just one nerve root and typically pain is worse in the legs than the back pain itself (1,6,9). Conditions that involve radiculopathy, spinal stenosis, and radicular pain/sciatica are classified together into a subset of LBP called radicular syndrome and are thought to be present in 5 to 10% of all cases (6).

Other than its ubiquity, one of the main reasons LBP is so heavily studied is because once someone experiences back pain, they tend to have recurrent back pain long term(1). In fact, a history of LBP is one of the strongest predictors of chronic back pain (10). One long term study on LBP found that within five years of an episode of back pain, 70% of the participants had another recurrence of pain (11). Psychosocial factors, often referred to as 'yellow flags', can indicate to care providers if the patient will develop chronic or recurrent pain (12,13). These yellow flags include attitudes and beliefs about back pain (e.g., maladaptive coping strategies), other psychiatric comorbidities, and external factors such as social interactions and whether the patient is off word due to LBP (12,13). The intermittent, unpredictable nature of the LBP can strongly affect a patient's

quality of life (1). Recurrent chronicity has been associated with patients taking early retirement, sick leave, and frequent utilization of health services (1).

1.2.1 Risk factors of LBP

Factors that can put a patient at risk of developing LBP are extensive and not fully understood. The most strongly linked risk factors are age and physically demanding jobs (1). Other risk factors have been noted, such as sex, obesity, smoking status, and other physical and mental comorbidities (1). However, the majority of this research has occurred in high-income countries. Further research into these risk factors in middle- to low-income countries is necessary to understand what factors are predictors of LBP risk.

Sex is one of the more complicated risk factors because, in high-income countries, women are more susceptible to LBP than men. For the purpose of this thesis, only sex will be considered as a risk factor, as most research in this area uses the term sex interchangeably with gender, and no studies for LBP have investigated the influence of both sex and gender on LBP risk. In females, menopause marks a decrease in hormones that may result in rapid disc and spinal degeneration compared to males (14). However, this correlation is not the only predictor, because in middle to low-income countries, sex no longer is a predictor (1). In Africa, more males than females have LBP, but in Latin America, this is not the case. There may be other cultural factors that may influence this discrepancy in risk due to sex. In Western countries, females may be more likely

to seek care than males, and inequalities in other countries may influence the opposite (1).

Obesity has not been globally found to be a universal risk factor for LBP either. In Russian and Finland, obesity was strongly related, but other countries such as South Africa, Spain, and China had either a weak association or no association at all (15). A meta-analysis on obesity as a risk factor for LBP showed that there was a slightly increased risk for LBP if the person was overweight (OR 1.15, 95% CI 1.08–1.21) and a more increased risk if a person was obese (OR 1.36, 95% CI 1.18–1.57) (16); however, this increased risk is minimal. A twin study that controlled for genetic factors found that obesity was not significant as a risk factor for LBP after genetics was considered (15). This was in contradiction to a systematic review that found that obesity had an OR of 1.9 after controlling for genetics (17).

Genetic predisposition to LBP has been a popular area of research within the last few years. The systematic review by Ferreira et al. (2013) found that genetic factors were most strongly associated with chronic LBP versus acute LBP (17). In a large-scale twin study that occurred in Denmark, 38% of LS pain most likely had a genetic component that meant the patient was more susceptible to LBP because of their genetics (18). This same study also found that women were more genetically susceptible to LBP then men, and this genetic component increased their susceptibility to LBP as they aged. Though there is likely a genetic factor that predisposes a person to LBP, the exact genes have not been identified.

1.2.2 The Burden of Disease for Low Back Pain

Closely tied to the prevalence, incidence, and aetiology of any given disease is the personal effect the disease causes to sufferers, called the Burden of Disease, or disease burden (19). This has to do with the effect that a disease has on the global population and is measured in outcomes such as disability-adjusted life years (DALY), premature death, or quality-adjusted life years (QALY) (19,20). These outcomes are important because they show that poor disease management can have an effect on more than quantifiable measures such as blood pressure or BMI. QALY and DALY allow researchers to quantify previously qualitative measures such as the quality of time spent in a person's life. There are also newer measures that have calculated the economic burden that a disease condition might have, as well as environmental factors that can influence the burden of disease.

The Global Burden of Disease is a massive study that aims to describe and quantify QALY, DALY, and other similar outcomes for various diseases in the world's population. Many different diseases are focused on in this study, but one of the major diseases is LBP, as it is the disease associated with the most years lived with disability (YLD) and is the sixth-largest contributor to overall DALY (18). In 2010, when the most recent Global Burden of Disease reports was published, the study found that 83 million DALYs were lived with low back pain (18). This is a massive increase from the 58.2 million DALY when the first Global Burden of Disease study was conducted in 1990. Though the methods used for assessing disease burden are imperfect, due to aspects of the burden of disease

such as the impact on family members being hard to quantify, they do give an idea into the extent, and negative impact something as ubiquitous as LBP can have on the world's population.

In Canada, a recent study of the Canadian population's burden of disease has shown that Musculoskeletal (MSK) disorders, including LBP, are now the third-highest cause of DALYs (all-ages, both sexes combined), behind cancer, and cardiovascular disease (21). This is in comparison to data from 2006, where MSK disorders were only the fourth highest cause of DALY, behind cancer, cardiovascular disease, and mental and substance abuse DALYs. There was a 3.1% increase in the age-standardised DALYs from 2006 to 2016. In 2016, there were also 1, 035, 204 YLD (all-age) from MSK disorders, and this makes MSK disorders the highest cause of YLD in Canada. MSK YLD showed a 22% increase from 2006. This trend of MSK disorders increasing in disease burden is something that has been found in other countries. Though it is not known how much LBP contributed to the MSK disorder category, the prevalence of LBP suggests that it is a large portion of the Canadian population.

1.3 Guidelines

Clinical Practice Guidelines (CPGs) are tools that have been developed to help physicians and other healthcare providers deliver the best treatment to their patients according to the best evidence and highest quality research available (22). For the most part, most disease conditions have their own set of guidelines that help with the identification, diagnosis, testing, and treatment of these

conditions. Guidelines are based on synthesised research from systematic reviews and meta-analyses of the literature that summarise effective tests and treatments for the condition of interest. They are usually created and endorsed by interested national organizations in different countries globally, and there have been multiple pushes to create internationally recognised guidelines to ensure consistent messaging and patient care (22).

While some guidelines are internationally known, they are not widely accepted and implemented into practice. A study of Canadian physicians' attitudes towards guidelines found that most viewed guidelines as a positive tool (23). The same study also found that 52% of surveyed physicians only used the guidelines monthly, indicating that though there is awareness of the guidelines, they are often not applied. The physicians generally were confident that the guidelines were reliable, but they were concerned that guidelines represented decreased autonomy in clinical skills (23).

Since the '90s, some form of LBP CPGs have been compiled and distributed so that the best patient care could be provided to patients with LBP (22,24). These guidelines have been reviewed and updated multiple times since their inception. They provide advice on diagnosing LBP, and though they do provide various treatments that have evidence to show they are effective, no one single treatment method is recommended. This is because medicine is not a one-size-fits-all practice, and many treatments may be effective for one person, but not for others. Guidelines are also not mandatory to follow, and because of this, there may be a tendency to think that they are not important or necessary.

1.3.1 LBP Guidelines

Low back pain guidelines have evolved many times since the 90s, as the latest research results are published to support or oppose the recommendations listed (21,23). Different national guidelines have been dedicated to updating their guidelines as soon as the latest evidence is published, while others are not. These varying national guidelines have been compared to one another several times to determine what key differences are found between them. A systematic review found that most national guidelines differed in the population the guidelines were targeting, with some focusing solely on acute LBP and others providing recommendations for both acute and chronic (22). Other than this key difference, most guidelines are similar in recommending what tests are performed by the physician, assessing for yellow flags, advising when diagnostic imaging is required, and what types of advice, medications, and referrals should be provided to the patient (22). Though all the guidelines are mostly consistent in their messaging, there are still discrepancies, which makes caring for patients with LBP challenging.

1.3.2 DI guidelines for LBP

The majority of guidelines have reached consensus on the recommendation that diagnostic imaging (DI) should not be used for routine use in the case of NSLBP (22). This recommendation came about for three main reasons: first, because diagnostic imaging such as CT and MRI can be expensive to use; second, because CT and x-ray imaging emit radiation, and third, because research has shown that important outcomes are not improved by the routine use

of DI for LBP (25,26). Radiation is a mutagen, meaning that it can damage cellular DNA, and exposure to radiation increases a patient's risk of cancer (27). Radiologists operate under the recommendation of optimising imaging to provide the lowest-effective dose, meaning that they only expose patients to the least amount of radiation necessary to provide an accurate image (28,29). Evidence from many studies, including clinical trials, have shown that routine imaging does not improve recovery or change LBP treatment (30). Knowing this, it is often advised in reputable guidelines that imaging should be avoided unless needed to confirm red flag pathology (31,32). Nevertheless, there are inconsistencies in different national guidelines for NSLBP, as some guidelines recommend that imaging should be used for NSLBP after a 6-week trial of conservative therapy, and other guidelines recommend never performing DI for these cases (22,24).

DI is, however, recommended in the case of serious spinal pathologies, for which there are red flag symptoms and injury mechanisms that indicate a patient may have the suspected condition (Table 1) (6). The guidelines are also very specific about which modality type (e.g., x-ray, CT, MRI) is useful for which suspected pathology (33). Only in cases of a suspected fracture are CT imaging recommended, though an x-ray is the preferred modality in this situation due to the lower radiation risk (33). Otherwise, MRI is the best method for investigating suspected cancer, cauda equina, and infection (34). In the case where a patient has radicular syndrome (radiculopathy, spinal stenosis, sciatica), imaging will not help patient recovery unless there is an indication that the patient may be a

surgical candidate. Progressive, degenerative neurological findings and intractable pain are primary indicators for possible surgical intervention (6). If a patient is not a surgical candidate, research has shown that imaging does not change the course of pain management and treatment (35). However, imaging is still routinely ordered, as 12% to 32.2% of all commercial insurance patients in the United States with LBP received an x-ray, 16% to 21% received an MRI, 1.4% to 3% received a CT, and 10.9% to 16.1% received an MRI and/or CT (36).

Table 1. LBP symptoms and situations, recommended imaging strategy, and timing of imaging. Adapted from The American College of Physicians Clinical Guideline for Diagnostic Imaging for low back pain (33).

Clinical Situation	Description	Imaging Strategy	Timing
Major risk factors for cancer	new onset of low back pain with history of cancer, multiple risk factors for cancer, or strong clinical suspicion for cancer	Radiography plus erythrocyte sedimentation rate	Immediate imaging
Risk factors for spinal infection	new onset of low back pain with fever and history of intravenous drug use or recent infection	Magnetic resonance imaging	Immediate imaging
Risk factors for or signs of the cauda equina syndrome	new urine retention, faecal incontinence, or saddle anaesthesia	Magnetic resonance imaging	Immediate imaging
Severe neurologic deficits	progressive motor weakness or motor deficits at multiple neurologic levels	Magnetic resonance imaging	Immediate imaging
Weak risk factors for cancer	unexplained weight loss or age >50 years	Radiography with or without erythrocyte sedimentation rate	Defer imaging after a trial of therapy
Risk factors for or signs of ankylosing spondylitis	morning stiffness that improves with exercise, alternating buttock pain, awakening because of back pain during the second part of the night, or younger age [20 to 40 y]	Radiography with or without erythrocyte sedimentation rate	Defer imaging after a trial of therapy
Risk factors for vertebral compression fracture	history of osteoporosis, use of corticosteroids, significant trauma, or older age [>65 y for women or >75 y for men]	Radiography with or without erythrocyte sedimentation rate	Defer imaging after a trial of therapy
Signs and symptoms of radiculopathy	back pain with leg pain in an L4, L5, or S1 nerve root distribution or positive result on straight leg raise or crossed straight leg raise test in patients who are candidates for surgery or epidural steroid injection	Magnetic resonance imaging	Defer imaging after a trial of therapy
Risk factors for or symptoms of spinal stenosis	radiating leg pain, older age, or pseudoclaudication in patients who are candidates for surgery or epidural steroid injection	Magnetic resonance imaging	Defer imaging after a trial of therapy
No criteria for immediate imaging and back pain improved or resolved after a 1-month trial of therapy			No Imaging
Previous spinal imaging with no change in clinical status			No imaging

1.3.3 LBP Guideline implementation

For over a decade, research into the implementation of back pain guidelines into practice has been conducted. The theory behind guideline implementation studies is that once the guidelines are delivered to physicians, the physicians will practice medicine accordingly in order to improve their practice and provide evidence-based patient care. The different categories of outcomes for guideline implementation research are ones related to the patient such as pain ratings and quality of life, outcomes related to physicians' learning and behaviour change, and outcomes related to the cost-effectiveness of such a strategy (37,38). There are many guideline implementation strategies that occur, some passively (e.g., posting guidelines or mailing them out), some actively (e.g., in-person education sessions, Clinical Decision Support), often with multiple interventions, or, more commonly, with a simple one-time intervention (38). Recently a systematic review (SR) synthesized these types of guideline implementation interventions, and though the SR was unable to conduct a meta-analysis due to the large heterogeneity between included studies, the SR found that simple, one-time intervention was not effective at changing practice (38). Another SR looking at multifaceted guideline implementation for both neck and back pain found that a multifaceted implementation strategy was also not effective in promoting behaviour that was adherent to the guidelines (37). Thus, the current challenge for guideline implementation is to determine if the guidelines are actually used in practice and if they are not, how to best promote their use by physicians.

Promising strategies to promote the use of guidelines do appear to be more active strategies as opposed to passive strategies (39).

1.3.4 LBP Imaging Guideline Implementation and Appropriateness

Many studies have researched various DI guideline implementation strategies, measuring the outcome of decreased diagnostic imaging use for LBP. An SR focussing on the implementation strategies for imaging found that Clinical Decision Support (CDS) was the most effective strategy for decreasing rates of imaging, while audit and feedback had mixed results; education with guideline dissemination was not effective (40). Recently, a published SR about radiography guideline implementation for low back pain in the emergency department (ED) setting found that, of the five interrupted time-series analysis studies on guideline implementation, there was limited evidence to truly recommend any specific implementation strategy (41). Though the setting limits the generalizability of the SR findings, it does indicate that this is a complicated task that requires more evidence to prove which strategy is best to promote adherence to the guidelines. This type of research has also looked into patient health outcomes to ensure that the patient does not suffer or experience worse health as a result of following guidelines. Generally, patient care does not improve if the patient is given imaging, and patient recovery is also not affected by withholding imaging (35).

Another way to determine whether imaging guidelines are being used in practice is to compare the imaging referrals to the guidelines published for imaging (25,35). This outcome is typically called appropriateness, justification, indicated, concordance, or adherence to the guidelines (42). It is a useful measure

because it provides stakeholders such as policymakers, researchers, and clinicians, with a simple, understandable outcome to demonstrate guideline use. The downside of using this measurement is that it only shows how often a guideline is adhered to in practice but does not provide any information regarding the knowledge that a provider may have on the guidelines, or if there were any factors that lead to the decision to not follow the guideline in that case. Thus, it does not help to determine if guidelines are truly helpful in clinical practice. Guidelines are only guides, not a rule. Not every healthcare centre has equal access to the options necessary to follow them (e.g., lack of access to ideal modality types) or is obliged to follow them.

In Canada, very few studies have investigated the appropriateness of imaging according to guidelines on DI for the LS (39). A Canadian government commissioned SR found that the rates of appropriateness varied globally, but as this study was not peer-reviewed, the methodology section was missing critical information for replicability (36). Busse et al. also included all providers who referred for imaging, including Chiropractors (39). In some public healthcare systems, patients do not typically seek treatment from a chiropractor, which is a private service not always covered by the public healthcare system. Instead, it is most common to receive imaging referrals from family physicians (FP) or general practitioners (GP). Thus, focusing in on these physician groups is important to understand the size of the issue for imaging. If practice is to change, understanding what the common reasons are that a physician orders imaging, and why they are doing so is crucial to understanding DI appropriateness.

1.4 Usual Care for LBP

One of the first points of contact with the healthcare system for most patients is an appointment with their FP (43). Typically, at this appointment, the physician's goal is to collect information on the patient that would aid in making a diagnosis. This informative data normally comes from patients' medical history and from the physical examination (44,45). From this, a physician should be able to form an idea regarding what is causing the patient pain and recommend conservative evidence-based treatments.

A physician is looking for red flags that, with the combination of LBP, indicate if there may be a serious pathology causing the patient's pain (see Table 1) (46). These serious spinal pathology red flag indicators include but are not limited to: Fever which can suggest infection; sudden unexplained weight loss which suggests cancer; severe neurological deficits like incontinence which can suggest cauda equina; and a physical exam with trauma that suggests a fracture (46,47). Other red flags are a history of cancer, pain at night, intravenous drug usage, and steroid use. Weak risk factors for serious pathology include age greater than 50 years old (33). These indicators direct the physician to refer the patient for further testing (e.g., DI or bloodwork) and to secondary care specialists such as a neurologist, a neurosurgeon, an oncologist, or an orthopaedic surgeon. If there are no red flags present, the physician must rely on other aspects of the patient's examination to determine if there is a specific cause of the pain. However, even if there are red flags, they are not often the best indicators that there is, in fact, a serious underlying pathology (43). In fact, a study found that

out of 1,172 patients presenting to primary care with LBP, 80.4% of these patients had a red flag indication, but only 0.9% had a serious underlying pathology (1,48). While red flags are important to note for physicians and help them triage patients more effectively, the poor specificity of these red flags shows how physicians may rely on further unnecessary imaging to provide reassurance that there is no serious spinal pathology.

Common low back disorders that are thought to cause pain in the low back and legs are spinal stenosis, degenerative disc disease, radiating pain/radiculopathy, and disc herniation (6). These are conditions that can affect the nerve roots that branch off from the spinal cord, the spinal cord itself, or the discs between the vertebrae. The complicated nature of LBP is such that even if a patient has signs of any of these diseases, these abnormalities may not, in fact, be the source of the pain at all (35). A study showed that radiologists who imaged patients who did not report LBP frequently revealed spinal abnormalities (40). Since the people included in the study were asymptomatic, these findings show that the so-called "defects" or findings do not always correlate to pain and do not always indicate where the problem is (43). Thus, it is difficult to state with 100% certainty that a patient's spinal abnormalities are causing LBP, and may explain why many patients develop chronic LBP, because the pain may not be attributed to the correct cause. This also explains why imaging is not always the best tool to aid in the diagnosis of the source of low back pain.

1.4.1 Usual Care for LBP- Guideline Concordant Treatment

Most LBP will resolve on its own with or without treatment intervention from a healthcare professional (7). When a patient does seek help from a healthcare professional, the guidelines recommend treatments with good quality evidence as support that a physician should provide to the patient (46,49). These involve treatments such as giving advice on staying active, advising against bed rest, and referring to allied health professionals such as physiotherapists (7,36,46). Unfortunately, a systematic review and meta-analysis (SR & MA) have shown that no treatment to date has a high magnitude of effect on patient pain and function when compared to placebo (50).

Various pharmacological treatments have been prescribed in the past that are recommended to help patients cope with pain, with NSAIDs being the most recommended treatment. However, an SR & MA has shown that acetaminophen (also called paracetamol) is not effective in reducing low back pain in the short term (51). Opioids were a common medication in the past, but since the US Food and Drug Administration and Health Canada have recognised that overprescribing of opioids led to an opioid crisis, this medication group has become more restricted in what it is prescribed for (49,52). Muscle relaxants are also prescribed at times, but the pain relief they offer comes with the downside of the sedative effects of the medication and is most effective for only ALBP (49). Though there are various medications that a patient can take, most guidelines recommend a trial of simple analgesics before trying stronger ones (43).

Surgical interventions are not the first line of care for patients with NSLBP due to the various risks that come with surgery (43). Most guidelines only recommend surgery when every other therapy option has failed, and when the patient has radicular symptoms in their legs with neurological deficits (43,53). Cases that may require surgery are those that have severe radiculopathy with intractable pain caused by compressed nerve roots or spinal stenosis. Compressed nerve roots can be caused by a herniated disc and can be treated by a discectomy, where the intervertebral disc that is causing the issues is partly or fully removed. Spinal stenosis is treated by removing the tissue that is compressing the spinal cord. However, 10% to 40% of surgical patients experiences pain even after surgery, resulting in what is called Failed Back Surgery Syndrome (FBSS) (54), suggesting that surgery may not be the best treatment for certain causes of LBP. With varying rates of spinal surgeries in different countries and within regions of one country, it is likely that there are different standards as to when surgery is considered (7). In Canada, patients seeking these kinds of surgeries often have to wait upwards of two years before they receive them, often due to referrals where a patient is determined not to be a surgical candidate (55). It is possible that patients who are waiting for surgery would be better served by providing them with standardised, evidence-based conservative care.

Other common types of LBP treatments that patients either seek on their own or are referred to are typically manual therapies or alternative therapies from allied health professionals or traditional Chinese medicine practitioners (50). There are many types of treatments that exist, such as spinal manipulation,

massage, acupuncture, TENS, and Kinesio-tape (50). However, the evidence for the effectiveness of these treatments are minimal, and most were found to have very little effect on a patient's pain (56). Most of these treatments have minimal, short term effects on the patient's pain and are not a solution or cure (7,49).

1.4.2 Usual Care for LBP Imaging

Though the guidelines recommend only ordering imaging in a very narrow amount of cases, one might ask, how often are patients actually referred for LS imaging? What is the usual practice for ordering imaging when a patient with a low back pain complaint visits a primary care physician? Physicians will typically report their adherence to the guidelines, but that may not truly reflect what is done in practice (57). Many individual studies in both the emergency department and in family practice have described such practices without any conclusive evidence. In 2000, a survey of United States family physicians was conducted, with a reported 40% stating that they ordered routine imaging for acute LBP (57). This study was further corroborated by an audit of Medicare records in the US that found that about 30% of patients received imaging for LBP that did not have any serious coding associated with it in less than 28 days of diagnosis (58). Recently, a systematic review complied all known studies that described ordering patterns for low back pain imaging (59). The results of this SR show that simple imaging is occurring for 16.3% of patients presenting with LBP to primary care, and complex imaging is occurring for 9.2% of LBP patients in primary care; however, this study included all primary care providers who ordered imaging and

not just on FPs. The self-reported survey's discrepancy with the SR is likely due to the different providers targeted with each.

1.5 Impacts of Unnecessary Imaging

There are many downsides to unnecessary imaging, both to the patient and to the healthcare system. Certain modalities such as x-ray and CT imaging expose patients to radiation, which is a mutagen and carcinogen (25). Since CT exposes patients to the highest amount of radiation, ensuring that patients are not getting unnecessary exposure to these dangerous emissions is of great importance (8).

The downside of unnecessary imaging to the healthcare system is that there is a high monetary cost to imaging, especially if the image does not help to confirm a diagnosis that results in a change in treatment. By receiving imaging, a patient may end up using more healthcare dollars in the long run than a patient who didn't receive imaging due to unforeseeable downstream costs. In a public healthcare system that is mostly single-payer, like the one in Canada, any extra costs to the healthcare system put a strain on the quality and access that patients have. With the increasing costs of healthcare, being able to manage healthcare budgets properly is in the best interest of all.

Receiving imaging is not necessarily correlated with quicker recovery (8). In fact, because there is increased use of healthcare system resources, unnecessary imaging can increase wait times for *all* patients requiring imaging, which delays timely treatment. Wait times for imaging can also increase the risk of chronicity because a patient may delay conservative treatment while waiting

for the test itself (8). It is also possible that the patient who receives imaging experiences further psychological stress from being diagnosed with a disease (8). Imaging can also reveal incidental findings unrelated to their pain, which can send the patient down a path that involves more tests. Often there may be delays, but the test does not result in the proper diagnosis of pain.

1.6 Medical Record Review

Medical record reviews, also known as chart audits, are an ever-increasing trend in research for many reasons, but mostly because this is data that is routinely collected by physicians on any individual who uses the healthcare system. This type of methodology now makes up a reported 25% of all published health studies (60). The benefits of a chart audit methodology are numerous: The data is already conveniently collected, it can be easy to access, it is cost-effective, it can provide a large sample size, and can allow for a long-term follow-up. However, there are also negatives associated with this type of data and method as this data is not collected for research purposes. There are potential data quality issues (e.g., illegible handwriting, transcription errors, etc.), missing data, varying follow-ups, and standard/usual care guideline adherence issues. Another limitation of observational data is that it cannot be randomised. Randomisation is a common method for dealing with both known and unknown biases, but when randomization either does not, or cannot occur, that data can be biased due to factors such as seasonal effects, provider preferences, patient requests, or other unknown sources (60).

Health information can be found in different databases, depending on the so-called "custodians" of the data. In Canada, with a public healthcare system, the custodian of health information tends to vary province to province but is typically the regulating government organization. However, there may be different databases that hold different types of information, such as medication records, radiological imaging records, insurance claims, workers' compensation records, etc. Paper records used to be the norm, but with the popularity and convenience of computers, there has been a shift to digital records.

This type of method is excellent to answer research questions related to practice patterns and utilization, quality of patient care, resource allocation, and guideline adherence. Choosing Wisely is an organization that was formed to encourage a more thoughtful approach to caring for patients in light of evidence showing over-use of certain healthcare tests and treatments. The use of chart audits is a common way to shows patterns of overuse and lack of guideline adherence.

1.7 The Local Context of Newfoundland and Labrador

The local context is often an important factor to consider as it can be an influential force in research. Local trends can also confound results and findings and limit the generalisability of research. Important details about the local context such as imaging rates, population, number of physicians, and number of imaging units in the Eastern Health (EH) Region are useful to report for this thesis.

The Canadian Institute of Health Information (CIHI) reported that NL ordered CT imaging at a rate much greater than most other provinces in Canada at 169.5 CTs/1,000 people (compared to Ontario at 116.9/1,000, British Columbia at 115.9/1,000) in 2012 (61). The only other province that ordered more CT imaging was New Brunswick. In 2017, the Canadian Agency for Drugs and Technologies in Health (CADTH) reported that there were 14 sites in NL that housed the 16 CT units in the province (62). This means that there are 30.26 CT units/1,000,000 people in the province. That is the largest number of CT units per population than anywhere else in Canada, as the next highest number of CT units are found in the Yukon, Nunavut, and the Northwest Territories and Ontario has 13.02 CT units/1,000,000 (62). In 2016, there were approximately 522,537 people who resided in Newfoundland and Labrador (63,64). Out of the 1,298 licenced and practising physicians in NL, 880 total physicians practice in the EH Region alone (65). Family physicians in NL are unable to order MRIs for their patients.

1.8 Patient-Oriented Research and Patient Engagement

The healthcare system was created with the purpose of attending to those who are sick, making patients the user for which the system is providing a service. However, the system is not always designed and researched with the end-user (patients) in mind, and often research is conducted on patients without considering how patients will be impacted by the findings(66). Many organizations and initiatives such as the Canadian Institutes of Health Research's Strategy for Patient-oriented Research (CIHR-SPOR), Patient-centred Outcomes

Research Institute (PCORI), and the National Health Service's INVOLVE have been created to promote, fund, and advise researchers who conduct Patient-Oriented Research (POR) (66-68) There are various domains that need to be considered when engaging patients in the research that a healthcare system funds, such as time for the project, funding available, research questions, and patient interest. Patients can be engaged in a basic level, such as informing them that this research is occurring and getting their feedback, to a more advanced level with patients joining research teams to assist in every aspect of research design, data collection, and knowledge translation (68).

This study was made possible through funds awarded by the Newfoundland and Labrador Support for People and Patient-Oriented Research and Trials Unit (NL SUPPORT). The funding provided the opportunity to engage patients in the formation of the project aims and objectives, as well as with Knowledge Translation. Engagement sessions with stakeholders, including patients, occurred as a part of this project's priority-setting and was used to set objectives for the fourth chapter of this thesis.

1.9 Thesis Objectives

This manuscript formatted Masters in Medicine thesis was focussed on DI ordered by primary care physicians for patients with LBP, specifically focusing on x-ray and CT imaging appropriateness. MRI modalities were not included in this thesis due to the restricted access for MRI that precludes family physicians from ordering them. The thesis objective was to describe CT utilization patterns locally,

and global and local DI appropriateness ordered in primary care settings. The thesis produced three manuscripts that focused on different aspects of the primary research objectives:

I.   The first manuscript determined the rates of imaging ordered here in the local context by FPs. This involved calculating age-sex standardised CT imaging utilization patterns for four recent years in Newfoundland and Labrador's Eastern Health Region. These rates were statistically compared to one another and contextualised using other rates published in other provinces and countries to frame CT imaging ordering habits for LBP in the local context.

II.   The second manuscript of this research thesis involved a systematic review and meta-analysis following PRISMA guidelines on what is already known about global appropriateness of x-ray and CT imaging in primary care settings and the different guidelines that have been used to determine appropriate proportions of imaging.

III.   The final phase in this thesis was an audit of CT referrals from FPs for patients with low back pain to establish the proportion of CTs that are ordered appropriately according to serious pathology red-flag guidelines. Those referrals that do not have a suspicion of a serious pathology were described according to the suspected cause of the pain. The audit was for the year 2016 and included LS CTs that were performed in the Eastern Health Region of Newfoundland and Labrador.

The three manuscripts of this thesis provided insight into the issue of imaging both globally and in the local context. It was hypothesized that there would be high rates of CT ordering in our local context. It was also hypothesized that there would be inconsistency as to what is known globally about CT and x-ray imaging appropriateness. Finally, it is hypothesized that there will be a small proportion of images that are ordered for a suspected serious spinal pathology and thus considered appropriate. The following chapters were prepared for publication in various journals and are thus formatted according to those journals' specifications.

1.10 References

(1) Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. The Lancet 2018 9–15 June 2018;391(10137):2356-2367.

(2) Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine 2008;33(1):95-103.

(3) Hoy D, Brooks P, Blyth F, Buchbinder R. The Epidemiology of low back pain. Best Practice & Research Clinical Rheumatology 2010 December 2010;24(6):769-781.

(4) Schnurrer-Luke Vrbanić T. Low back pain-from definition to diagnosis. Reumatizam 2011;58(2):105-107.

(5) Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet 2016 10/08; 2018/08;388(10053):1545-1602.

(6) Bardin LD, King P, Maher CG. Diagnostic triage for low back pain: a practical approach for primary care. Med J Aust 2017 Apr 3;206(6):268-273.

(7) Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. The lancet 2012;379(9814):482-491.

(8) Flynn TW, Smith B, Chou R. Appropriate use of diagnostic imaging in low back pain: a reminder that unnecessary imaging may do as much harm as good. J Orthop Sports Phys Ther 2011 Nov;41(11):838-846.

(9) Allegri M, Montella S, Salici F, Valente A, Marchesini M, Compagnone C, et al. Mechanisms of low back pain: a guide for diagnosis and therapy. F1000Res 2016 Jun 28;5:10.12688/f1000research.8105.2. eCollection 2016.

(10) Stanton TR, Henschke N, Maher CG, Refshauge KM, Latimer J, McAuley JH. After an episode of acute low back pain, recurrence is unpredictable and not as common as previously thought. Spine 2008;33(26):2923-2928.

(11) Hestbaek L, Leboeuf-Yde C, Manniche C. Low back pain: what is the long-term course? A review of studies of general patient populations. European Spine Journal 2003;12(2):149-165.

(12) New Zealand Guidelines Group. New Zealand acute low back pain guide: incorporating the guide to assessing psychosocial yellow flags in acute low back . 2004; Available at: https://www.healthnavigator.org.nz/media/1006/nz-acute-low-back-pain-guide-acc.pdf. Accessed 07/23, 2019.

(13) Chou R, Shekelle P. Will This Patient Develop Persistent Disabling Low Back Pain? JAMA 2010;303(13):1295-1302.

(14) Wang YX, Wang JQ, Kaplar Z. Increased low back pain prevalence in females than in males after menopause age: evidences based on synthetic literature review. Quant Imaging Med Surg 2016 Apr;6(2):199-206.

(15) Dario AB, Loureiro Ferreira M, Refshauge K, Luque-Suarez A, Ordonana JR, Ferreira PH. Obesity does not increase the risk of chronic low back pain when genetics are considered. A prospective study of Spanish adult twins. Spine J 2017 Feb;17(2):282-290.

(16) Zhang TT, Liu Z, Liu YL, Zhao JJ, Liu DW, Tian QB. Obesity as a Risk Factor for Low Back Pain: A Meta-Analysis. Clin Spine Surg 2018 Feb;31(1):22-27.

(17) Ferreira PH, Beckenkamp P, Maher CG, Hopper JL, Ferreira ML. Nature or nurture in low back pain? Results of a systematic review of studies based on twin samples. European Journal of Pain 2013;17(7):957-971.

(18) Hoy D, March L, Woolf A, Blyth F, Brooks P, Smith E, et al. The global burden of neck pain: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014 Jul;73(7):1309-1315.

(19) Lopez AD, Murray CC. The global burden of disease, 1990–2020. Nat Med 1998;4(11):1241.

(20) Prüss-Üstün A, Mathers C, Corvalán C, Woodward A. Introduction and methods: assessing the environmental burden of disease at national and local levels. 2003.

(21) Lang JJ, Alam S, Cahill LE, Drucker AM, Gotay C, Kayibanda JF, et al. Global Burden of Disease Study trends for Canada from 1990 to 2016. CMAJ 2018 Nov 5;190(44):E1296-E1304.

(22) Koes BW, van Tulder MW, Ostelo R, Burton AK, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. Spine 2001;26(22):2504-2513.

(23) Hayward RS, Guyatt GH, Moore KA, McKibbon KA, Carter AO. Canadian physicians' attitudes about and preferences regarding clinical practice guidelines. CMAJ 1997 Jun 15;156(12):1715-1723.

(24) Koes BW, van Tulder M, Lin CC, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. European Spine Journal 2010 12/01;19(12):2075-2094.

(25) Choosing Wisely Canada. Imaging Tests for Lower Back Pain: When you need them and when you don't. Available at: https://choosingwiselycanada.org/imaging-tests-low-back-pain/. Accessed January 9, 2019.

(26) Lemmers GPG, van Lankveld W, Westert GP, van dW, Staal JB. Imaging versus no imaging for low back pain: a systematic review, measuring costs, healthcare utilization and absence from work. European Spine Journal 2019 05/01;28(5):937-950.

(27) Smith-Bindman R, Lipson J, Marcus R, Kim K, Mahesh M, Gould R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Arch Intern Med 2009;169(22):2078-2086.

(28) Brink JA, Amis Jr ES. Image Wisely: a campaign to increase awareness about adult radiation protection. 2010.

(29) Amis Jr ES, Butler PF, Applegate KE, Birnbaum SB, Brateman LF, Hevezi JM, et al. American College of Radiology white paper on radiation dose in medicine. Journal of the american college of radiology 2007;4(5):272-284.

(30) Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. Radiol Clin North Am 2012 Jul;50(4):569-585.

(31) Davis PC, Wippold FJ,2nd, Brunberg JA, Cornelius RS, De La Paz RL, Dormont PD, et al. ACR Appropriateness Criteria on low back pain. J Am Coll Radiol 2009 Jun;6(6):401-407.

(32) Patel ND, Broderick DF, Burns J, Deshmukh TK, Fries IB, Harvey HB, et al. ACR appropriateness criteria low back pain. Journal of the American College of Radiology 2016;13(9):1069-1078.

(33) - Chou R, - Qaseem A, - Owens DK, - Shekelle P. - Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. - Annals of internal medicine (- 3):- 181.

(34) Clinically Organized Relevant Exam (CORE) Back Tool . 2013; Available at: http://www.health.gov.on.ca/en/pro/programs/ecfa/docs/lb_tk_core_tool_bw.pdf. Accessed May 13, 2019.

(35) Chou R, Fu R, Carrino JA, Deyo RA. Imaging strategies for low-back pain: systematic review and meta-analysis. Lancet 2009 /;373(9662):463-472.

(36) Dagenais S, Galloway EK, Roffey DM. A systematic review of diagnostic imaging use for low back pain in the United States. The Spine Journal 2014;14(6):1036-1048.

(37) Suman A, Schaafsma FG, Elders PJ, van Tulder MW, Anema JR. Cost-effectiveness of a multifaceted implementation strategy for the Dutch multidisciplinary guideline for nonspecific low back pain: design of a stepped-wedge cluster randomised controlled trial. BMC Public Health 2015 May 31;15:522-015-1876-1.

(38) Mesner SA, Foster NE, French SD. Implementation interventions to improve the management of non-specific low back pain: a systematic review. BMC musculoskeletal disorders 2016;17(1):258.

(39) Busse J, Alexander P, Abdul-Razzak A, Riva J, Alabousi M, Dufton J. Appropriateness of spinal imaging use in Canada. Hamilton, ON: McMaster University 2013.

(40) Jenkins HJ, Hancock MJ, French SD, Maher CG, Engel RM, Magnussen JS. Effectiveness of interventions designed to reduce the use of imaging for low-back pain: a systematic review. CMAJ 2015 Apr 7;187(6):401-408.

(41) Liu C, Desai S, Krebs LD, Kirkland SW, Keto-Lambert D, Rowe BH, et al. Effectiveness of interventions to decrease image ordering for low back pain presentations in the emergency department: a systematic review. Acad Emerg Med 2018;25(6):614-626.

(42) Jenkins HJ, Downie AS, Maher CG, Moloney NA, Magnussen JS, Hancock MJ. Imaging for low back pain: is clinical use consistent with guidelines? A systematic review and meta-analysis. Spine J 2018 May 3.

(43) Maher CG, Williams C, Lin C, Latimer J. Managing low back pain in primary care. Aust Prescr 2011 /;34(5):128-132.

(44) Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CC, Chenot J, et al. Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. European Spine Journal 2018;27(11):2791-2803.

(45) NSW Agency for Clinical Innovation. Management of people with acute low back pain: model of care. 2016; Available at: https://www.aci.health.nsw.gov.au/__data/assets/pdf_file/0007/336688/acute-low-back-pain-moc.pdf. Accessed 05/15, 2019.

(46) Chou R, Qaseem A, Snow V, Casey D, Cross JT, Shekelle P, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 2007;147(7):478-491.

(47) Raison NT, Alwan W, Abbot A, Farook M, Khaleel A. The reliability of red flags in spinal cord compression. Arch Trauma Res 2014 Mar 30;3(1):e17850.

(48) Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, et al. Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. Arthritis Rheum 2009 Oct;60(10):3072-3080.

(49) Chou R, Deyo R, Friedly J, Skelly A, Hashimoto R, Weimer M, et al. Nonpharmacologic therapies for low back pain: a systematic review for an American College of Physicians Clinical Practice Guideline. Ann Intern Med 2017;166(7):493-505.

(50) Machado L, Kamper S, Herbert R, Maher C, McAuley J. Analgesic effects of treatments for non-specific low back pain: a meta-analysis of placebo-controlled randomized trials. Rheumatology 2008;48(5):520-527.

(51) Machado GC, Maher CG, Ferreira PH, Pinheiro MB, Lin C-C, Day RO, et al. Efficacy and safety of paracetamol for spinal pain and osteoarthritis: Systematic review and meta-analysis of randomised placebo controlled trials. BMJ (Online) 2015 2015/03;350.

(52) Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. JAMA 2016;315(15):1624-1645.

(53) Savigny P, Kuntze S, Watson P, Underwood M, Ritchie G, Cotterell M, et al. Low back pain: early management of persistent non-specific low back pain. London: National Collaborating Centre for Primary Care and Royal College of General Practitioners 2009;14.

(54) Chan C, Peng P. Failed back surgery syndrome. Pain medicine 2011;12(4):577-606.

(55) Coyle MJ, Roffey DM, Phan P, Kingwell SP, Wai EK. The Use of a Self-Administered Questionnaire to Reduce Consultation Wait Times for Potential Elective Lumbar Spinal Surgical Candidates: A Prospective, Pragmatic, Blinded, Randomized Controlled Quality Improvement Study. JBJS 2018;100(24):2125-2131.

(56) Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. The Lancet 2018;391(10137):2368-2383.

(57) Di Iorio D, Henley E, Doughty A. A survey of primary care physician practice patterns and adherence to acute low back problem guidelines. Arch Fam Med 2000 Nov-Dec;9(10):1015-1021.

(58) Pham HH, Landon BE, Reschovsky JD, Wu B, Schrag D. Rapidity and modality of imaging for acute low back pain in elderly patients. Arch Intern Med 2009 May 25;169(10):972-981.

(59) Downie A, Hancock M, Jenkins H, Buchbinder R, Harris I, Underwood M, et al. How common is imaging for low back pain in primary and emergency care? Systematic review and meta-analysis of over 4 million imaging requests across 21 years. Br J Sports Med 2019 Feb 13.

(60) Worster A, Haines T. Advanced statistics: understanding medical record review (MRR) studies. Acad Emerg Med 2004;11(2):187-192.

(61) Canadian Institute of Health Information. Medical Imaging in Canada 2012. 2013; Available at: https://www.cihi.ca/en/mit_summary_2012_en.pdf. Accessed January 9, 2019.

(62) Canadian Agency for Drugs and Technologies in Health. The Canadian Medical Imaging Inventory, 2017. 2018; Available at: https://cadth.ca/canadian-medical-imaging-inventory-2017. Accessed 05/21, 2019.

(63) Statistics Canada. **Population and Dwelling Count Highlight Tables, 2016 Census**. 2019; Available at: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/Table.cfm?Lang=Eng&T=101&S=50&O=A. Accessed 07/24, 2019.

(64) Newfoundland and Labrador Statistics Agency. **Quick Facts**. 2019; Available at: https://www.stats.gov.nl.ca/Default.aspx. Accessed 05/17, 2019.

(65) Newfoundland and Labrador Medical Association. NLMA Membership Statistics. 2019; Available at: http://nlma.nl.ca/FileManager/About-NLMA/docs/2019.07.22_NLMA_Membership_Statistics.pdf. Accessed 07/24, 2019.

(66) Canadian Institutes of Health Research. **Patient engagement**. 2018; Available at: http://www.cihr-irsc.gc.ca/e/45851.html. Accessed 05/14, 2019.

(67) Forsythe LP, Ellis LE, Edmundson L, Sabharwal R, Rein A, Konopka K, et al. Patient and stakeholder engagement in the PCORI pilot projects: description and lessons learned. Journal of general internal medicine 2016;31(1):13-21.

(68) Involve N. Briefing notes for researchers: involving the public in NHS, public health and social care research. INVOLVE Eastleigh, UK 2012.

CO-AUTHORSHIP STATEMENT

This thesis was made possible by a team of co-authors who generally contributed to the research question, protocol, analysis, and manuscript preparation. My committee was formed of Drs. Amanda Hall, Holly Etchegary, and Patrick Parfrey. Dr Amanda Hall was my primary supervisor and responsible for assisting with the research question, project designs, and forming of protocols. Dr Holly Etchegary helped coordinate patient engagement sessions and provided manuscript preparation assistance. Dr Patrick Parfrey provided further resources and advice on research designs. Further statistical analysis training was provided by Bethan Copsey (PhD candidate). Drs Kris Aubrey-Bassler and Russell Dawe provided content expertise to the medical record review. Andrea Pike provided insights for manuscript preparation for the systematic review and meta-analysis. Dr Krista Mahoney provided assistance and expertise to the CT rates manuscript.

I certify that the majority of this thesis was performed by me, from protocol to data extraction to analysis and manuscript drafting. To the best of my knowledge, I have acknowledged all co-authors and contributors properly. All co-authors have received and approved of the manuscripts. I have produced this thesis as the primary author.

CHAPTER 2: LUMBAR SPINE CT REFERRAL RATES BY FAMILY PHYSICIANS USING ROUTINELY COLLECTED DATA IN ONE HEALTH REGION IN NEWFOUNDLAND AND LABRADOR, CANADA: A BRIEF REPORT

Authors: Gabrielle S Logan MSc(c)*[1], Bethan Copsey PhD(c)[2], Holly Etchegary PhD[1], Patrick Parfrey MD[1], Krista Mahoney PhD[1], Amanda Hall PhD[1,3]

Affiliations: (1) Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada; (2) Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK (3) Primary Health Research Unit, Memorial University of Newfoundland, St. John's, NL, Canada.

*Corresponding author: Gabrielle S Logan; Address: 300 Prince Philip Dr, St. John's, Newfoundland and Labrador, Canada, A1B 3V6; Email: glogan@mun.ca (Orcid ID: https://orcid.org/0000-0001-7007-9373)

2.1 Introduction

Choosing Wisely recommends reducing unnecessary lumbar spine (LS) Computed Tomography (CT) imaging for low back pain, primarily to improve patient safety by avoiding unnecessary exposure to carcinogenic ionising radiation and secondarily to reduce healthcare spending associated with over-testing (1,2). However, only a handful of studies have examined population-based utilisation of lumbar spine (LS) CTs, with most studies reporting the proportion of patients with low back pain who receive CT imaging compared to those who do not (3,4). Australia and the US provide population-level data on LS CT utilisation for their countries, which is helpful for comparisons of usage internationally (5,6). These estimates range from 209/100,000 to 2,464/100,000 individuals (5,6). To our knowledge, there are no peer-reviewed publications of Canadian LS CT utilisation rates. The only data available were provided in a government-commissioned report on appropriate imaging, and it focused on the LS CT rates in just two Canadian provinces, Manitoba and Ontario, and found different estimates in both provinces (7).

The objective of this study is to determine the yearly age-sex standardised rates of LS CT imaging for adults (≥19 years old) by family physicians in the Eastern Health (EH) Region of Newfoundland and Labrador (NL), Canada. This study adds to the body of work in this area by presenting LS CT rates from a third province in Canada. It has been estimated that NL has a higher use of CTs (any procedure) than any other Canadian province (8). While we could hypothesize that the rate of LS CTs may also be higher in NL than other provinces,

comparisons between provinces are beyond the scope of this study due to lack of data access required for this analysis.

2.2 Methods

Data Source. The third-party data custodian identified the dataset from the administrative code for LS CTs from Meditech, an electronic medical records database in the EH Region of NL, and provided it to the researchers. Records from 2013 to 2016 were accessed, and the following variables were collected: number of LS CTs with or without contrast, age, sex, ordering physician speciality, and imaging service date.

Data Cleaning. The dataset contained all LS CT scans conducted between January 1st, 2013 and December 31st, 2016. The inclusion criteria were adults (>19 years old) who received a CT scan, and referrals that were ordered from an FP (any speciality other than family medicine or general practitioner was excluded).

Data that did not fit the inclusion criteria were removed. The pediatric population (<18) was removed because different diagnostic imaging guidelines apply to children. Patients aged 19 were removed because when standardising a population on age and sex, those aged 19 are in the 15 to 19 years old category, of which the majority is a pediatric classification. The physician that orders a CT image for a patient with LBP typically is the patient's family physician (FP). As such, we focused on this group of providers. Finally, yearly totals of LS CT imaging were obtained.

Data Analysis. Crude rates of LS CT referrals were calculated by dividing the total number of CTs performed in EH Region in each year of interest (numerator) by the total population of EH Region in that same year of interest (denominator) and multiplying that proportion by 100,000 people. The Newfoundland and Labrador Centre of Health Information (NLCHI) provided population estimates. The rate from 2016 used 2015 EH Region population estimates, as the population estimates for 2016 were not available.

Age-sex standardised rates of LS CTs were calculated by categorising all records of CT referrals into appropriate age groups and sex of the patient for each year of interest. Each age group contained 5 different ages (e.g., 20 to 24). Each year of interest's CT rate for the applicable age-sex categories was determined by dividing the CT count for an age-sex category by the population estimate for that same age-sex category and multiplying the proportion by 100,000.

CT counts for 2014, 2015, and 2016 were estimated using 2013 population age-sex estimates from the EH region in NL. For example, this was calculated for 2014 by taking the 2014 rate for each age-sex category and dividing it by 100,000 to get the proportion, and multiplying the proportion by the same 2013 age-sex categorised population estimate. Then we summed the estimated CT counts for each year of interest. The total estimate of CT counts for each year was used to calculate the age-sex adjusted rate by taking the CT count estimate for each year of interest, dividing it by the 2013 population estimate and multiplying it by 100,000.

Rate ratios and 95% confidence intervals (CIs) were calculated to compare whether or not rates of CT referrals per 100,000 people in a 1-year period were increasing over time. Each year's age-sex standardised LS CT rate was compared to the previous year's rate to see if there was a statistically significant change. CIs were calculated, and if either the upper or lower CI crossed one, this indicated that the rates were not significantly different.

2.3 Results

There was a total of 18,358 LS CTs performed in the EH Region between 2013 and 2016. 3987 records were excluded due to provider (n= 2831), patient age (n=98), or insufficient information (n=1058) resulting in 14,371 included records.

The age-sex standardised rates were similar to the crude rate and are as follows: 1,225/100,000 (95% CI 1,223.5, 1,226.2) in 2013, 1,393/100,000 in 2014 (95% CI 1,405.3, 1,408.2), 1,556/100,000 in 2015 (95% CI 1,566.3, 1,569.3) and 1,395/100,000 in 2016 (95% CI 1,406.2, 1,409.1). Age-sex standardised rates are presented in Figure 1. Crude rates of CT referrals per 100,000 were as follows: 1,225/100,000 in 2013, 1,399/100,000 in 2014, 1,568/100,000 in 2015, and 1,408/100,000 in 2016. The rate ratios comparing a year to an adjacent year are presented in Table 1. The greatest increase in rates was between 2014 and 2013, and there was a decrease in rates between 2016 and 2015.
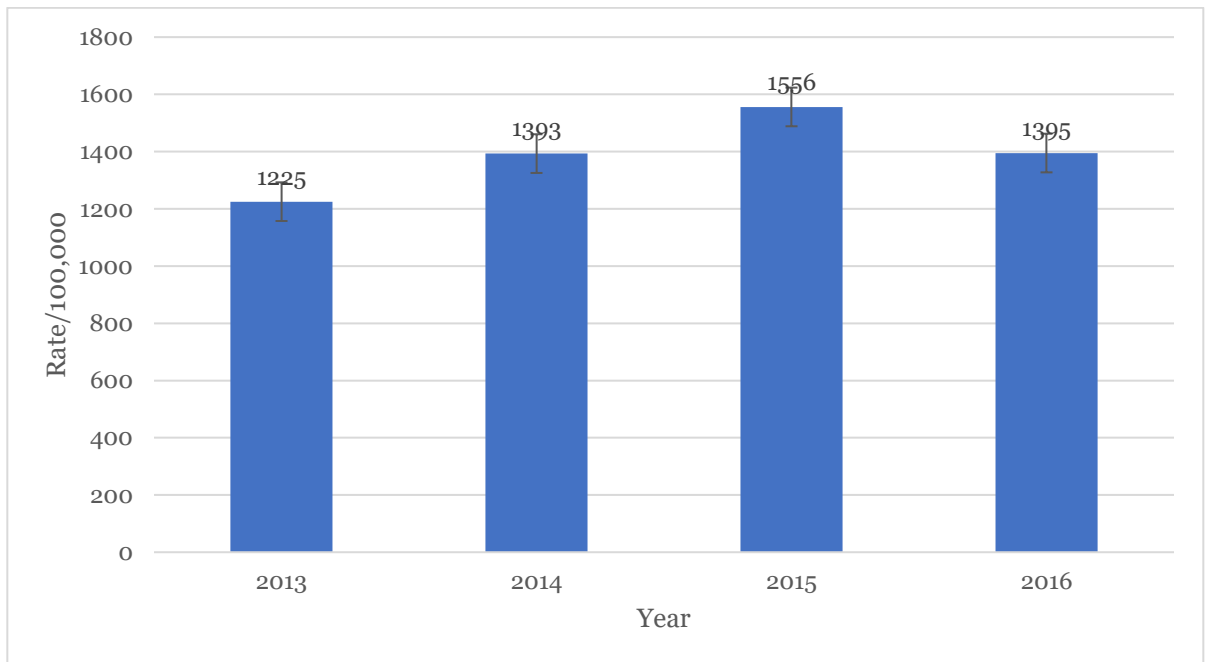
Figure 2.1. Age-sex standardised rate of adult CT referrals per 100,000 people for the lumbar spine from GPs in the Eastern Health Region from 2013 to 2016.

Table 2.1. Rate ratios comparing age-sex standardised rate estimates in adjacent years

Year comparison	Rate Ratio* (95% confidence interval)
2014 to 2013	1.137[#] (95% CI 1.084, 1.194)
2015 to 2014	1.117[#] (95% CI 1.067, 1.169)
2016 to 2015	0.896[#] (95% CI 0.857, 0.938)
*Calculated by dividing the more recent year by the year previous. [#] Statistically significant due to the large sample size	

2.4 Discussion

The age-sex standardised LS CT rate ranged from 1253 to 1556/100,000 individuals over four years. While our rate ratio analysis identified that the observed differences in rates were statistically different, the magnitude of these differences was so small they are likely clinically irrelevant. Thus, the LS CT rate in NL has remained fairly steady from 2013-2016. Diagnostic imaging data from a larger timeframe would allow for accurate trend analysis.

To put our findings in context with other populations, we found data from Canada, Australia, and the USA (5-7). Busse et al. published grey-literature rates and found that in Manitoba, Canada, in 2010/11, the age-sex standardised rate of LS CTs ordered was 1000 LS CTs per 100,000 individuals, and in Ontario, Canada, the age-sex standardised rate was approximately 660 LS CTs per 100,000 persons (7). However, direct comparisons are difficult, as the reference population in our NL context used NL specific age-sex standardised population estimates in the analysis techniques and Busse et al. did not use the same reference populations. It is also noteworthy that family physicians were the target provider for the NL age-sex standardised rates; thus, all other providers were excluded. This may not have been the case for Busse et al.

In Australia, we found age-standardised rates only, which varied from 209/100,000 to 2,464/100,000 individuals (6). In the USA, there were also different rates of spinal imaging from different hospital referral regions, which ranged from 320/100,000 to 2,370/100,000 individuals (age, sex, and race

standardised) (5). Caution needs to be taken when comparing NL CT utilisation rates to other countries. While numerically our rates are within the range of these other countries, differing population estimates for the reference populations of these other countries limits the direct comparison of these rates.

It is important to note the limitations in our dataset and findings. First, there were no 2016 age and sex population estimates available from NLCHI for analysis in time for this publication; thus age-sex standardised rate for 2016 was based on population estimates from 2015. Given that the number of people in the EH Region may have changed from 2015 to 2016, the accuracy of the 2016 estimate may not be as comparable to estimates with accurate population estimates. Second, the data used for analysis were routinely collected health data not collected with research purposes in mind. We cannot know if there was misclassified or missing data, if the quality and accuracy of the data were considered, or if there were other unforeseen confounders (9). Finally, FPs in NL are not authorised to order MRIs; this imaging modality is limited to secondary and tertiary physicians only. It is possible that the inability for FPs to order MRI skews the CT rates, but further research is needed to address this limitation.

In conclusion, there appears to be a high rate of LS CTs ordered in the EH Region of NL and this seems similar or higher compared to other countries or larger Canadian provinces. While direct or indirect comparisons were beyond the scope of this paper, future research could look more closely at comparisons of LS CT utilization rates amongst provinces, especially given the high prevalence of

this condition and lack of clinical utility of CTs for providing conservative care to patients with LBP. Similarly, further research is needed to better understand how many CTs were necessary for the management of a patient's condition. It is important to focus research on health system targeted interventions to improve the appropriateness of CT referrals, which would ensure patient safety is prioritized, and healthcare funding is spent appropriately.

2.5 References

(1) Choosing Wisely Canada. Imaging Tests for Lower Back Pain: When you need them and when you don't. Available at: https://choosingwiselycanada.org/imaging-tests-low-back-pain/. Accessed January 9, 2019.

(2) Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. Radiol Clin North Am 2012 Jul;50(4):569-585.

(3) Mafi JN, McCarthy EP, Davis RB, Landon BE. Worsening trends in the management and treatment of back pain. JAMA Intern Med 2013 Sep 23;173(17):1573-1581.

(4) Eccles M, Steen N, Grimshaw J, Thomas L, McNamee P, Soutter J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. The Lancet 2001 5 May 2001;357(9266):1406-1409.

(5) Lurie JD, Birkmeyer NJ, Weinstein JN. Rates of advanced spinal imaging and spine surgery. Spine 2003;28(6):616-620.

(6) Australian Commission on Safety and Quality in Health Care. Computed tomography of the lumbar spine. 2015; Available at: https://www.safetyandquality.gov.au/wp-content/uploads/2015/11/SAQ201_03_Chapter2_v9_FILM_tagged_merged_2-3.pdf. Accessed May 1, 2019.

(7) Busse J, Alexander P, Abdul-Razzak A, Riva J, Alabousi M, Dufton J. Appropriateness of spinal imaging use in Canada. Hamilton, ON: McMaster University 2013.

(8) Canadian Institute of Health Information. Medical Imaging in Canada 2012. 2013; Available at: https://www.cihi.ca/en/mit_summary_2012_en.pdf. Accessed January 9, 2019.

(9) Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Peteresen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLoS Med 2015 2015/;12(10).

CHAPTER 3: WHAT DO WE REALLY KNOW ABOUT THE APPROPRIATENESS OF IMAGING FOR LOW BACK PAIN IN PRIMARY CARE? A SYSTEMATIC REVIEW AND META-ANALYSIS OF MEDICAL RECORD REVIEWS

Authors: Gabrielle Logan BSc Hons[1], Andrea Pike MSc[2], Bethan Copsey[3], Patrick Parfrey MD[1], Holly Etchegary PhD[1], Amanda Hall PhD[1,2]

Affiliations: (1) Faculty of Medicine, Memorial University, St. John's, NL, Canada; (2) Primary Healthcare Research Unit, Memorial University, St. John's, NL, Canada; (3) Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

 *Corresponding author: Gabrielle Logan; Address: 300 Prince Philip Dr, St. John's, NL, Canada, A1B 3V6; Email: glogan@mun.ca (Orcid ID: https://orcid.org/0000-0001-7007-9373)

3.1 Introduction

Guidelines for the assessment and treatment of low back pain (LBP) have been in circulation since the 1980s with more than 11 countries publishing their own LBP clinical guidelines in the last two decades.[1] While most early versions of LBP guidelines did not recommend routine use of radiographic imaging for assessment of LBP, there were discrepancies about when to image (e.g., some guidelines provided specific criteria or timeframes for imaging, and others did not). In the 1980s and 1990s, x-ray imaging was commonly recommended in the assessment of LBP persisting longer than four weeks[1], and Computed Tomography (CT) was often recommended in patients experiencing neurological deficits, including radicular symptoms.[2,3] For the last 25 years, there has been increased congruence among LBP guidelines regarding when and under what circumstances to use diagnostic imaging (DI). Since 2000, the recommendations typically state that DI is warranted only when patients with LBP present with red flag symptoms that suggest the presence of one of four known specific spinal pathologies (severe cauda equina, infection, fracture, and cancer).[4,5] Guidelines have also been updated with respect to the potential direct and indirect patient harms of DI, particularly x-ray and CT, as well as their lack of clinical utility for non-specific LBP.

Harms of over-testing

Overuse of x-ray and CT imaging for LBP is not a benign issue. It can result in concerning direct and indirect harms to the patient. Over-testing also results in significant economic burden to health systems.

Patient harms

Both x-ray and CT imaging expose patients to radiation, a known mutagen that can increase the risk of cancer.[6] The human body can tolerate some radiation, but the more exposure that a patient has to radiation, the greater their cancer risk. This risk of radiation is even greater to young patients as radiation can affect both male and female fertility.[7] Thus, radiologists typically recommend using x-ray and CT only when medically necessary and clinically justified to patient care.[8,9]

In addition to the harms from radiation, imaging can reveal anatomical abnormalities or incidental findings.[10] These abnormalities are also extremely common in asymptomatic patients, are only weakly correlated with patient symptoms, and are not always the cause of a patient's pain.[10] For example, a systematic review in 2014 found that disc degeneration was present in 96% of asymptomatic adults aged 80 and up, and disc bulges found in 80%.[11] The harm of incidental findings is that patients may have to be sent for further tests to confirm that the finding is, in fact, benign, which may delay the patient receiving the appropriate treatment.

Moreover, patients who receive DI do not have better patient outcomes compared to those treated without imaging.[5,10] Chou et al. performed a

systematic review and meta-analysis to compare physical outcomes of patients

with LBP who received imaging to those who did not.[12] They found that

patients who received immediate imaging for non-serious LBP had similar pain

and function outcomes both in the short and long term compared to patients who

received usual care without imaging.[12] This systematic review called into

question the value and safety of DI, given that imaging does not appear to help

patients' physical outcomes and may also have negative effects on a patients'

psychological well-being.

Health system burden

In addition to patient harms, over-testing results in a substantial economic

burden to healthcare systems. In the United States, the dollar value spent on all

CTs in 2000 was $975 million, and by 2006, the amount increased to $2.17

billion.[13,14] In countries with a public healthcare system, it is difficult to

quantify in dollars the cost of unnecessary imaging, but in Canada, the rate of CT

imaging has almost doubled since 2003,[15] suggesting that the cost of imaging

has also drastically increased. This financial increase also is associated with

trickle-down effects such as the increased need for follow-up, further

investigations of incidental findings, referrals to specialists, and even

surgery.[10,16]

Importance of assessing appropriateness

Given the potential patient harms and added health care costs of using DI,

it is essential to understand if these tests are being used appropriately according

to the current guidelines. This information allows us to understand whether and to what degree patient safety and quality of care are compromised with the use of unnecessary testing. A recent systematic review of DI appropriateness for LBP found that approximately one third of imaging referrals were not appropriate; however, this review included imaging referrals from any healthcare provider for any imaging modality (including MRIs).[17] X-ray and CT pose the most direct harm to patients; thus we intend to provide a focused estimate of appropriateness for these tests only. Additionally, since physicians in family practice or emergency department settings are the most common setting for imaging referrals for patients with LBP and follow the same guidelines for imaging ordering, we will focus our question to this provider population. This will also allow us to reduce any heterogeneity in our estimate due to potentially different ordering practices or guidelines amongst different providers.

Aim

We aim to synthesize the evidence from all studies investigating the appropriateness of physician-made referrals for CTs and x-rays for LBP. Our review adds to the literature by providing clinicians, implementation researchers and policy makers with an estimate of imaging appropriateness that is specific to physicians working in family practice and emergency department settings.

3.2 Methods

Search Strategy

Four databases, PubMed, CINAHL, EMBASE and The Cochrane Database of Systematic Reviews, were searched for terms related to the PICO keywords of low back pain, guidelines, and adherence. The search string was developed with a research librarian. Databases were searched from inception to May 2018 (see Supplementary file). Titles and abstracts from each database search were imported to Endnote (version 10), and duplicates were removed before screening. Forward and backward citation tracking as well as reference lists of relevant systematic reviews and policy documents was done on all included papers in order to ensure our database search captured all applicable published research articles.

Inclusion Criteria

Studies were included if (i) the design was a retrospective or prospective review/audit of medical records, (ii) the data item was data on lumbar CT and x-ray images, (ii) the imaging referrals were made by a physician in either general practice or emergency department settings, (iii) the analysis compared the reason for imaging referral to a guideline source, and (iii) the outcome was the proportion of appropriate or inappropriate referrals based on adherence to the guidelines. All LBP types were eligible for inclusion. Studies that looked at appropriateness of images referred by other providers such as chiropractors, physiotherapists, nurse practitioners, or pharmacists were excluded. Only studies that reported individual or aggregate data from chart reviews for CT and x-ray imaging were included. If other tests or imaging modalities (e.g., MRI) were

combined with x-rays or CTs, the study was excluded, the study authors were contacted to confirm if x-ray and CT data could be reported separately, if not the study would be excluded. Other study designs, such as self-reported surveys or simulated patient visits were excluded. Since there was potential for variation in imaging recommendations found in guidelines published prior to the year 2000 which could impact in the definition of appropriateness, we excluded all studies in which the data were obtained prior to the year 2000.

Two reviewers (GL, AH) screened titles and abstracts and created a shortlist of full texts to be screened. Full texts were scrutinized by two reviewers (GL, AH) to assess eligibility against the inclusion/exclusion criteria. Any discrepancy was resolved upon discussion of the difference and consensus of the categorization for inclusion. Authors of studies that did not have a full text available (abstract or conference proceedings only) were contacted to determine if there was a published full text. Authors of studies that did not report imaging modalities included were contacted to determine if MRI was included in the aggregate data.

Data extraction

An electronic data collection form was developed to extract information from all included studies on study characteristics and outcome data. For each study the healthcare setting, LBP type, sample size, and outcome data were extracted. Outcomes included both the proportion of appropriate and inappropriate images. Additional outcome information included: the guidelines source used for comparison, the definition used to assess appropriateness (or inappropriateness),

the outcome denominator (if outcome reported the number of patients, images, visits), and measurement error (if reported) was extracted.

Quality of Reporting and Risk of Bias Assessment

Quality of reporting was assessed for each study according to the "Reporting of studies Conducted using Observational Routinely-collected health data" (RECORD) Statement checklist, which is an expansion of the "Strengthening the Reporting of Observational Studies in Epidemiology" STROBE Statement checklist.[18-21] Every included study was compared to the RECORD Statement's 35-item checklist to determine if the study reported pertinent information.

No widely accepted tool exists for assessing Risk of Bias (RoB) for this type of observational study. Guidance was provided by a review authored by Sanderson et al. which provides a list of specific domains to be considered.[22] RoB for these observational, non-randomised studies was determined by using items that related to the following 4 domains: Representativeness of patients, misclassification of patients, misclassification of outcome measurement, and inconsistent data. Overall study RoB was judged to be low if 4 out of the 4 domains judged low risk, moderate if 3 domains were considered low risk or high if two or less domain items were low risk.

Data synthesis and analysis

Data were summarized separately for appropriateness of x-rays and appropriateness of CTs. We extracted estimates of the proportion of appropriate x-rays or CTs (and 95% confidence intervals) from each included study. In one case, the study only included an estimate of inappropriateness.[48] In this case the authors were contacted and confirmed that we could accurately use the inverse of their estimate as the proportion of appropriate x-rays. When studies did not provide CIs for their appropriate percentage, we calculated the 95% CI using the formula for calculating confidence intervals for a single proportion in STATA (v 15). Meta-analysis for a single proportion using a random effects model was completed on studies that were determined to be clinically homogenous.[23] The effect size was calculated with STATA (v 15).

We applied the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) approach to assess certainty of the estimates of appropriateness.[24] Certainty was downgraded based on 4 factors:

- Risk of Bias: Twenty-five percent or more of the participants were from studies rated as having a high RoB.

- Inconsistency in results: Determined by examining whether the estimates were similar in magnitude (overlapping confidence intervals).

- Indirectness of evidence: More than 50% of the participants were outside the target group.

- Imprecision of evidence: Determined based on the width of the confidence interval (CI) associated with the proportion of appropriateness (+/- 3%) and the overall sample size (at least 2000 participants).

3.3 Results

We identified a total of 919 publications from database searching (n=918) and reference lists (n=1), which was reduced to 696 studies after deduplication (Figure 3.1). We reviewed 185 full texts, of which 22 were excluded for specific reasons (see Supplementary file). [25-46] Of the six final included studies,[47-52] one study was published primarily in Spanish,[52] and two studies were abstracts only for which there was no full publication.[47,48]
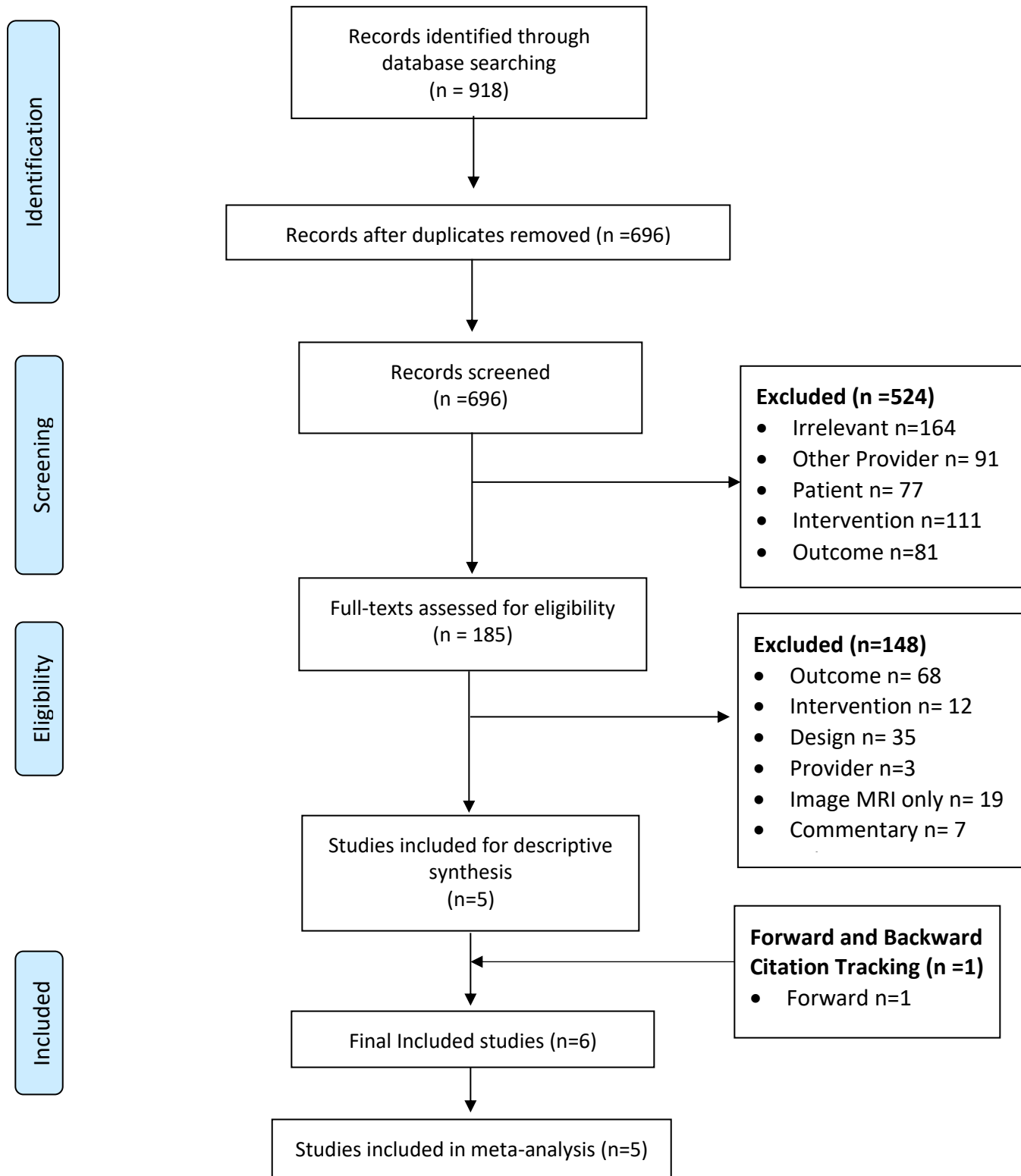
Figure 3.1. PRISMA flow diagram of the search strategy

Study characteristics

The studies were conducted in Finland, Ireland, Spain, & the United States (Table 3.1). Study settings were primary care settings that were often a mix of both ED and GPs, such as in a hospital or clinic. Sample sizes ranged from 30 to 3908. The duration of LBP in the different studies was undefined. Five of 6 studies assessed the appropriateness of x-rays; two of the six studies assessed the appropriateness of CTs. The studies used a range of different guidelines to select the criteria for determining appropriateness. Of the six studies, nine different guidelines were used; some studies were guided by more than one guideline source.

Table 3.1. Study characteristics and reported outcomes of appropriateness organised by image type

Study / Country	Setting[1] Patient age	Database / Data source	Guideline Source	Definition of Appropriateness	Denominator sample size[2]	% Appropriate (95%CI)	Risk of Bias
x-ray							
Baez 2011 USA	Mixed 18-40years	EMR Imaging referral[3]	ACR, ACP, APS	Adherence to EBGs	Per image (100)	34% (25, 43%)	High
Culleton 2013 Ireland	Mixed ≥65years NR	EMR Radiology findings	RCR	Adherence to RCR guidelines	Per image (414)	18% (14, 22%)	High
Muntion-Alfaro 2006, Spain	Mixed NR	Medical Records Unclear	RCGP, AHCPR, ICSI	No red flags	Per patient (538)	47% (43, 51%)	Moderate
Schlemmer 2015 USA	ED NR	Insurance Claims Imaging referral[3]	ACR, NCQA	Red flag indicators, >6-weeks of LBP	Per claim (3908)	56% (55, 58%)	Low
Tahvonen 2016 Finland	Mixed NR	Medical Records Imaging referral	EC	Unclear	Per image (50)	32% (19, 45%)	High
CTs							

Oikarinen 2009 Finland	Mixed <35years	Medical Records Imaging referral[3]	EC	Situations of trauma	Per patient (30)	23% (8, 39%)	High
Schlemmer 2015 USA	ED NR	Insurance Claims Imaging referral[3]	ACR, NCQA	Red flag indicators, >6-weeks of LBP	Per claim (648)	56% (52, 60%)	Low

[1] A mixed setting refers to studies that used a data source of imaging referrals in which the referring physician could be practicing in a family practice, in-hospital or emergency department setting.

[2] The number of lumbar spine imaging referrals reviewed.

[3] In addition to the referral, patient charts may have been accessed to determine patient information for determining appropriateness

NR: not reported

EBG: Evidence Based Guidelines

Guideline Abbreviations: NCQA: National Committee for Quality Assurance; RCGP: Royal College of General Practitioners; AHCPR: Agency for Health Care Policy and Research; ICSI: Institute for Clinical Systems Improvement; RCR: Royal College of Radiologists; ACR: American College of Radiologists; ACP: American College of Physicians; APS: American Pain Society; EC: European Commission

Note: the type of low back pain (e.g. acute, chronic) was not specified in any of the studies.

Reporting quality using the RECORD Checklist

Study Design

The included studies were all retrospective chart reviews/audits (see Supplementary file), though not all used common terms to indicate that.[47] The majority of studies were a general chart audit/review done specifically to quantify appropriate imaging for LBP. However, one study's objective was to quantify appropriateness of CT imaging in young patients and included more than CT imaging of the lumbar spine.[49]

Setting

All included studies were a general chart review of medical records and were conducted in a primary care provider setting and reported adequate information for the settings according to the RECORD checklist. The settings were identified as a hospital or health centre, with only one study mentioning data coming from the ED settings solely.[51]

Participants and Study Size

Participants were largely identified either by patient records or records of images. Coding used to identify the included records was clearly described in only two studies.[51,52] These two studies were the only studies to justify the study's sample size.

Data Sources/Variables

Most studies took information from the patients' hospital or clinic charts directly. If there was a specific database or computer program that was accessed, it was not communicated in the published paper. Electronic medical records were specified in three studies, but the applications were not identified by name.[48,51,52] One study utilized an insurance claims database.[51]

Data Access, Cleaning, Linkage, and Supplementary Information

These reporting criteria were poorly or not at all discussed in the studies. If there was linkage involved, it was not clarified, and if the data cleaning occurred, the details were not explained sufficiently. No study mentioned the level of database accessed by researchers. Only Schlemmer et al. provided supplementary data that was available for access online.[51]

Risk of Bias

The four domains that were assessed for RoB were representativeness of patients, misclassification of patients, misclassification of outcome measurement, and inconsistency in data reporting (Figure 3.2). Four studies were judged to have a high risk of bias, one to have a moderate risk of bias[52] and one to have a low RoB.[51]

Author, year	Representativeness	Mis-classification Patient	Mis-classification Outcome	Inconsistent data	Final Judgement
Baez 2011	✓	✗	!	✓	High
Culleton 2013	✓	✗	!	✓	High
Muntion-Alfaro	✓	✓	!	✓	Moderate
Oikarinen 2009	✗	✗	✓	✓	High
Schlemmer 2015	✓	✓	✓	✓	Low
Tahvonen 2016	✓	✗	!	✓	High

Figure 3.2 Risk of Bias of Included studies as determined by the the representativeness of patients, risk of misclassification of patients, misclassification of outcome of interest, and inconsistent data.

Estimates of Appropriateness

X-rays

We found five studies that reported the appropriateness of x-rays, with four studies that used the reason for referral to determine appropriateness (Table 1).[47,50-52] One study, by Culleton et al., used the radiology findings report interpreting the image to determine appropriateness.[48] It was excluded from the meta-analysis due to the heterogeneity of outcome assessment and data source. From the four studies with 4184 participants, we found low-quality evidence that 44% (95% CI: 34%, 54%) of x-rays were appropriate (Figure 3.3). The quality of evidence was downgraded for two reasons; inconsistency and indirectness (Table 3.2). The estimate was determined to be inconsistent based on non-overlapping confidence intervals of individual estimates across studies. As well, the estimate was downgraded due to indirectness as one of the studies was conducted solely in an ED setting. Assuming the ED is functioning as intended, people are presenting with trauma, which is often an indicator for imaging; thus, we anticipate higher rates of appropriateness in the ED compared to general practice settings.

CTs

We found two studies with 678 participants that reported the appropriateness of CTs (Table 3.1). Both studies used the reason for referral to determine appropriateness but used different criteria to define the outcome. Schlemmer et

al.[51] defined appropriateness as any red flag condition or pain that has persisted greater than 6 weeks and Oikarinen et al.[49] restricted the definition to only situations of trauma. Using both studies, we found very low-quality evidence that 54% (95% CI: 51%, 58%) of CTs for LBP were appropriate (Figure 3.3). Similar to the outcome of x-ray appropriateness, the certainty of the estimate for CT appropriateness was downgraded due to inconsistency because of non-overlapping confidence intervals and indirectness because there were differences in the setting that would influence the outcome. Additionally, the estimate was downgraded due to imprecision, although the confidence intervals were somewhat narrow, the estimate is based on a sample size that is less than 2000 participants which challenge the certainty of the estimate (Table 3.2).

Table 3.2. GRADE Summary of Findings for the outcome of appropriateness of x-ray and CT imaging for patients with low back pain.

Appropriateness of x-ray and CT imaging in patients with LBP ordered by primary care physicians			
Population: Patients with any type of low back pain			
Setting: Emergency department, General Practice, Hospital			
Comparison: Back pain guidelines for imaging, assumed to focus on red flag indicators			
Outcome	**Effect**	**Number of participants in Studies**	**Certainty**
Appropriateness of x-ray	44% (34 to 54%)	n=5010; five studies	Low[2,4] ⊕⊕OO
Appropriateness of CTs	54% (51 to 58%)	n=678; two studies	Very low[2,3,4] ⊕OOO
* GRADE Working Group grades of evidence **High quality:** Further research is very unlikely to change our confidence in the estimate of effect. **Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. **Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. **Very low quality:** We are very uncertain about the estimate. [1] Downgraded due to Risk of Bias [2] Downgraded on Inconsistency [3] Downgraded imprecision [4] Downgraded on indirectness			
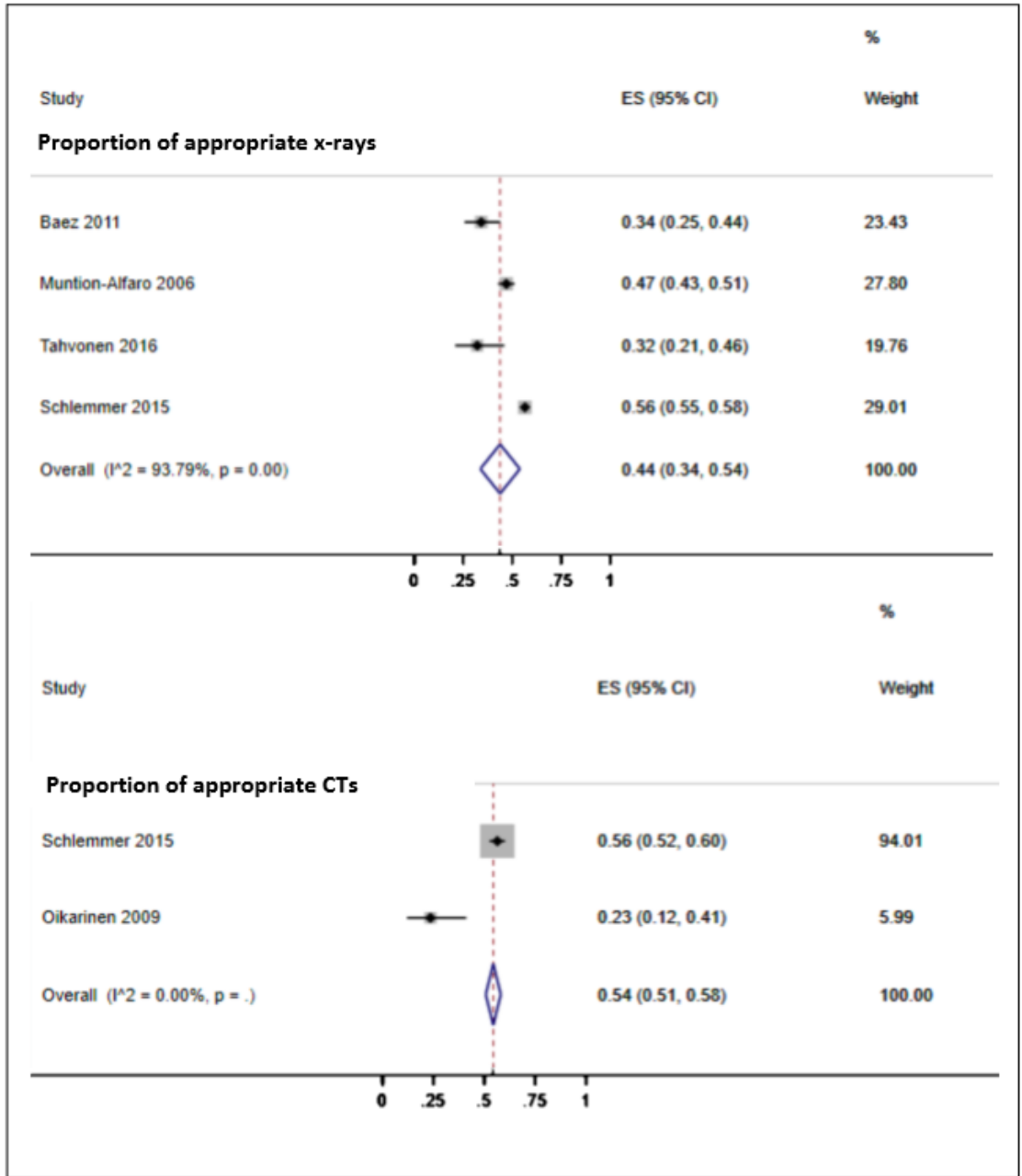
Figure 3.3. The proportion of appropriate x-rays and CT scans for low back pain.

3.4 Discussion

Few studies have been published reporting on the appropriateness of x-ray and CT scans ordered by primary care physicians (in general practice or emergency medicine) individually for patients with LBP. Among the studies we identified, most were conducted in European countries. No audit was conducted in countries such as Canada and Australia despite these countries having ongoing national campaigns to reduce unnecessary imaging for LBP (e.g., Choosing Wisely Canada).[7] From the available evidence, we found that only half of x-rays and CTs are being ordered according to guidelines. However, due to several factors related to inconsistency and indirectness, we have low certainty in this estimate. Our lack of certainty stems largely from the variation or lack of reporting on how appropriateness had been defined in these studies. Moreover, the majority of the studies we identified were conducted with very small sample sizes (and were thus underpowered to provide reliable estimates) and were of low methodological and reporting quality. In order to advance the science in this area, better quality studies that are adequately powered and adhere to guidelines for conducting and reporting clinical audits using routinely collected data are required.

Prior to our review, it was difficult to say anything regarding the appropriateness of imaging for LBP according to the guidelines. While another systematic review has investigated imaging appropriateness, it had significant heterogeneity by including multiple providers who may be following different guidelines and included multiple imaging modality types, including MRI.[17] Our review adds to the current knowledge base in this area by answering a specific

67

question regarding the appropriateness of x-ray and CT for patients with LBP. Given that there have been several recent (past 5 years) international campaigns targeting physicians in general practice and emergency departments to reduce x-ray and CT imaging, providing a robust assessment of the appropriateness specific to this recommendation is necessary to help clarify the issue and set targets for change.[7]

With respect to the estimate of imaging appropriateness, it is important to discuss that we found a wide variation in the methods and reporting of the included studies. The six included studies cited 9 different guideline sources which were not always internationally recognized. In addition, although the names and sometimes references of guidelines were mentioned as the source for determining appropriateness, it was not clear which criteria were used to define the outcome. For example, many guidelines recommended imaging only when red flags were present, and others provided additional criteria, which recommended imaging after a certain duration of LBP and non-response to treatment. It was unclear how these criteria were operationalized to code the reasons for referral as appropriate or not. This could lead to misclassification of the outcome or low reliability of the results. Better reporting of criteria for defining appropriateness and examples of operationalizing the coding protocol would improve our understanding of possible heterogeneity in the outcomes across studies.

Other sources of potential heterogeneity included the differences in inclusion criteria regarding patient population, the setting in which imaging referrals were made, and the medical record data sources. For example, two studies looked at patients that were under the age of 40, while one study looked only at patients older than 65 years. While most studies included a mixture of settings with referrals made from hospital-based or general practice-based physicians, one study focused solely on referrals made within an emergency department setting. Lastly, one study collected data from an insurance database, while two looked at EMR, and three did not describe the database other than to mention medical records. These potential sources of clinical heterogeneity may explain some of the inconsistency in the estimates across studies.

Strengths

This review has several strengths. We adhered to the PRISMA guidance for conducting and reporting systematic reviews and meta-analysis using observational data.[53,54] This included a) having two reviewers screen studies and extract data and b) providing an assessment of methodological quality and heterogeneity among the included studies. For the meta-analysis, we used a random-effects model to account for differences in the study conduct. Finally, we used the "RECORD checklist" to provide a robust assessment of the quality of reporting which allowed us to make sound recommendations for advancing the quality and replicability of the science in these types of study designs.

Limitations

Despite its strengths, the study is limited by two key factors. First, due to personnel resource constraints, we chose to use a more specific search strategy, meaning that it may not have been sufficiently sensitive to identify an exhaustive list of all potentially relevant studies. However, after consultation with a librarian about this decision, we included forward and backward citation tracking to enhance our specific search of electronic databases. While additional citation tracking did identify several potentially relevant studies all but one[51] were later excluded for various reasons (see Supplementary file).

Another limitation of this systematic review involves the quality and risk of bias assessments of the included studies. Many of the studies were not described in sufficient detail to assess the quality for replicability. Since a tool does not already exist to help grade the studies that are reporting routinely collected health data, the domains for potential introduction of bias were selected based on expert opinion. This makes it difficult to compare to other systematic reviews.

Future research

Based on this review's findings, we identified several areas for future research that would improve our knowledge about the appropriateness of LBP imaging. First, only 2 studies assessed the appropriateness of CT images for LBP that were ordered by physicians. One of these studies had a very small sample size and a high risk of bias, and the other was methodologically sound but was conducted in an ED setting. Future studies in other countries using similar methods to Schlemmer et al. in both general practice and emergency settings

would be helpful to confirm appropriateness of CTs for LBP. This would involve adhering to the RECORD statement for improved reporting quality. Additionally, for both outcomes of x-rays and CTs, we found that the definition of appropriateness varied among studies and in many cases the definition was often unclear or too vague to allow meaningful interpretation or replication. Thus, as a first essential step, we recommend future research clearly report the definition of appropriateness they are using and the operationalization of the definition for coding purposes. Second, and possibly most important, this field of research would benefit from a standardized definition of appropriateness for x-rays and CTs. This could be based on a spectrum to reflect some variation in the guidelines, ranging from a very strict cut-off (e.g., appropriate if only trauma-indicated used in the Oikarinen et al. study) to more inclusive definitions (e.g., any red-flag indicated and/or having pain greater than 6 weeks as was used in Schlemmer et al.).[49,51]

Implications for practice

The results of this systematic review show that in several countries, about half of the referrals for LBP imaging (x-rays and CTs) are not appropriate according to the guidelines. Due to the associated patient harms of x-ray and CTs scans, including radiation exposure, high rates of incidental findings and risk of delayed recovery, non-adherence to the guidelines represents low-value care for patients.[27] Hence, it is important to better understand why these referrals are made through future research.

Conclusion

Recently there has been a push to reduce unnecessary and inappropriate imaging, not only to save costs but also to provide better patient care.[10] Before this review, it was difficult to say anything regarding how appropriate imaging for low back pain is according to the guidelines. We now have an estimate of appropriateness, which indicates that only about half of imaging is appropriate according to recent guidelines. However, due to lack of published research, this estimate was not informed by data from many of the countries promoting the reduction of inappropriate imaging such as Canada, Australia and the UK. Moving forward, what we need is for more countries to undertake high-quality studies with sufficiently large sample sizes using clear definitions of appropriateness.

3.5 References

1 Koes BW, van Tulder MW, Ostelo R, et al. Clinical guidelines for the management of low back pain in primary care: an international comparison. *Spine J* 2001;26(22):2504-13.

2 Deyo RA. Early diagnostic evaluation of low back pain. *J Gen Intern Med* 1986;1(5):328-38.

3 Schroth WS, Schectman JM, Elinsky EG, et al. Utilization of medical services for the treatment of acute low back pain. *J Gen Intern Med* 1992;7(5):486-91.

4 Davis PC, Wippold FJ, Brunberg JA, et al. ACR Appropriateness Criteria on low back pain. *J Am Coll Radiol* 2009;6(6):401-7.

5 Chou R, Qaseem A, Owens DK, et al. Clinical Guidelines Committee of the American College of Physicians. Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. *Ann Intern Med* 2011;154(3):181-9.

6 Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII, Phase I, Letter Report. Washington, DC: The National Academies Press; 1998.

7 Choosing Wisely Canada. Imaging Tests for Lower Back Pain: When you need them and when you don't. Available at: https://choosingwiselycanada.org/imaging-tests-low-back-pain/. Accessed January 9, 2019.

8 Bushberg JT. Eleventh annual Warren K. Sinclair keynote address—science, radiation protection and NCRP: building on the past, looking to the future. *Health Phys* 2015;108(2):115-23.

9 Answers to common questions about the use and safety of CT scans. Mayo Clinic Proceedings: Elsevier; 2015.

10 Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. *Radiol Clin North Am* 2012;50(4):569-85.

11 Brinjikji W, Luetmer PH, Comstock B, et al. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. *AJNR Am J Neuroradiol* 2015;36(4):811-6.

12 Chou R, Fu R, Carrino JA, et al. Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet* 2009;373(9662):463-72.

13 Medicare Part B. imaging services: rapid spending growth and shift to physician offices indicate need for CMS to consider additional management practices. Washington, DC: Government Accountability Office 2008.

14 Lehnert BE, Bree RL. Analysis of appropriateness of outpatient CT and MRI referred from primary care clinics at an academic medical center: how critical is the need for improved decision support? *J Am Coll Radiol* 2010;7(3):192-7.

15 Canadian Institute of Health Information. Medical Imaging in Canada 2012. 2013; Available at: https://www.cihi.ca/en/mit_summary_2012_en.pdf. Accessed January 9, 2019.

16 Deyo RA. Diagnostic evaluation of LBP: reaching a specific diagnosis is often impossible. *Arch Intern Med* 2002;162(13):1444-7.

17 Jenkins HJ, Downie AS, Maher CG, et al. Imaging for low back pain: is clinical use consistent with guidelines? A systematic review and meta-analysis. *Spine J* Published Online First: 3 May 2018. doi:10.1016/j.spinee.2018.05.004

18 Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Med* 2015;12(10):e1001885.

19 Nicholls SG, Quach P, Von Elm E, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement: Methods for arriving at consensus and developing reporting guidelines. *PLoS ONE* 2015; 10(5):e0125620.

20 von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Int J Surg* 2014;12(12):1495-99.

21 Noah N. The STROBE Initiative STrengthening the Reporting of OBservational studies in Epidemiology (STROBE). *Epidemiol Infect* 2008;136(7):865.

22 Sanderson S, Tatt ID, Higgins J. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36(3):666-76.

23 Nyaga VN, Arbyn M, Aerts M. Metaprop: a Stata command to perform meta-analysis of binomial data. *Archives of Public Health* 2014;72(1):39.

24 Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64(4):383-94.

25 Fullen BM, Maher T, Bury G, et al. Adherence of Irish general practitioners to European guidelines for acute low back pain: A prospective pilot study. *Eur J Pain* 2007;11(6):614-23.

26 Raja AS, Ip IK, Cochon L, et al. Will publishing evidence-based guidelines for low back pain imaging decrease imaging use? *Am J Emerg Med* Published Online First: 2018 July 18. doi:10.1016/j.ajem.2018.07.039

27 Charlesworth CJ, Meath THA, Schwartz AL, et al. Comparison of low-value care in medicaid vs commercially insured populations. *JAMA Intern Med* 2016;176(7):998-1004.

28 Foo C, Pearson K. Are we choosing wisely at northern health?[abstract] *Intern Med J* 2017;47:12.

29 Kost A, Genao I, Lee JW, et al. Clinical decisions made in primary care clinics before and after Choosing Wisely™. *J Am Board Fam Med* 2015;28(4):471-4.

30 Lin IB, Coffin J, O'Sullivan PB. Using theory to improve low back pain care in Australian Aboriginal primary care: a mixed method single cohort pilot study. *BMC Fam Pract* 2016;17:44.

31 Rao S, Rao S, Harvey HB, et al. Low back pain in the emergency department - Are the ACR Appropriateness Criteria being followed? *J Am Coll Radiol* 2015;12(4):364-9.

32 Rego MH, Nagiah S. Over-imaging in uncomplicated low back pain: a 12-month audit of a general medical unit. *Intern Med J* 2016;46(12):1437-9.

33 Bishop PB, Wing PC. Compliance with clinical practice guidelines in family physicians managing worker's compensation board patients with acute lower back pain. *Spine J* 2003;3(6):442-50.

34 Buller-Close K, Schriger DL, Baraff LJ. Heterogeneous effect of an Emergency Department Expert Charting System. *Ann Emerg Med* 2003;41(5):644-52.

35 Day F, Hoang LP, Ouk S, et al. The impact of a guideline-driven computer charting system on the emergency care of patients with acute low back pain. *Proc Annu Symp Comput Appl Med Care* 1995:576-80.

36 Eccles M, Steen N, Grimshaw J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. *Lancet* 2001;357(9266):1406-9.

37 Espeland A, Albrektsen G, Larsen JL. Plain radiography of the lumbosacral spine. An audit of referrals from general practitioners. *Acta Radiol* 1999;40(1):52-9.

38 Halpin SF, Yeoman L, Dundas DD. Radiographic examination of the lumbar spine in a community hospital: an audit of current practice. *BMJ* 1991;303(6806):813-5.

39 Hourcade S, Treves R. Computed tomography in low back pain and sciatica. A retrospective study of 132 patients in the Haute-Vienne district of France. *Joint Bone Spine* 2002;69(6):589-96.

40 Gonzalez-Urzelai V, Palacio-Elua L, Lopez-de-Munain J. Routine primary care management of acute low back pain: adherence to clinical guidelines. *Eur Spine J* 2003;12(6):589-94.

41 Schectman JM, Schroth WS, Verme D, et al. Randomized controlled trial of education and feedback for implementation of guidelines for acute low back pain. *J Gen Intern Med* 2003;18(10):773-80.

42 Suarez-Almazor ME, Belseck E, Russell AS, et al. Use of lumbar radiographs for the early diagnosis of low back pain: Proposed guidelines would increase utilization. *JAMA* 1997;277(22):1782-6.

43 Tacci JA, Webster BS, Hashemi L, et al. Clinical practices in the management of new-onset, uncomplicated, low back workers' compensation disability claims. *J Occup Environ Med* 1999;41(5):397-404.

44 Deyo RA, Diehl AK. Lumbar spine films in primary care. *J Gen Intern Med* 1986;1(1):20-25.

45 Espeland A, Baerheim A, Albrektsen G, et al. Patients' views on importance and usefulness of plain radiography for low back pain. *Spine J* 2001;26(12):1356-63.

46 Richards PJ, Tins B, Cherian R, et al. The emergency department: an appropriate referral rate for radiography. *Clin Radiol* 2002;57(8):753-8.

47 Baez J, Khorasani R. Lumbar radiographs: Adherence to evidence based guidelines[abstract]. *Emerg Radiol* 2011;18(6):460.

48 Culleton S, Quinn C, O'Keeffe P. Imaging of low back pain in older people: An audit of current practice[abstract]. *Ir J Med Sci* 2013;182:S246.

49 Oikarinen H, Meriläinen S, Pääkkö E, et al. Unjustified CT examinations in young patients. *Eur Radiol* 2013;19(5):1161-5.

50 Tahvonen P, Oikarinen H, Niinimäki J, et al. Justification and active guideline implementation for spine radiography referrals in primary care. *Acta Radiol* 2017;58(5):586-92.

51 Schlemmer E, Mitchiner JC, Brown M, et al. Imaging during low back pain ED visits: A claims-based descriptive analysis. *Am J Emerg Med* 2015;33(3):414-8.

52 Muntión-Alfaro MT, Benítez-Camps M, Bordas-Julve JM, et al. Back pain: Do we follow the recommendations of the guidelines? *Aten Prim* 2006;37(4):215-20.

53 Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151(4):264-9.

54 Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA* 2000;283(15):2008-12.

CHAPTER 4: WHAT ARE THE REASONS FAMILY PHYSICIANS REFER PATIENTS WITH LOW BACK PAIN FOR CTS AND ARE THEY APPROPRIATE ACCORDING TO THE GUIDELINES? A RETROSPECTIVE REVIEW OF 4,435 MEDICAL RECORDS IN NEWFOUNDLAND USING ROUTINELY COLLECTED DATA FROM LINKED DATABASES

Authors: Gabrielle S Logan BSc Hons[1], Russell Eric Dawe MD[1,2], Kris Aubrey-Bassler MD[1,2], Patrick Parfrey MD[1], Holly Etchegary PhD[1], Amanda Hall PhD[1,2]

Affiliations: (1) Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada; (2) Primary Healthcare Research Unit, Memorial University of Newfoundland, St. John's, NL, Canada

*Corresponding author: Gabrielle S Logan; Address: 300 Prince Philip Dr, St. John's, NL, Canada, A1B 3V6; Email: glogan@mun.ca (Orcid ID: https://orcid.org/0000-0001-7007-9373)

4.1 Abstract

Background. CT Imaging referrals are often ordered for patients with low back pain by their family physicians. Though evidence-based guidelines have been created to help physicians manage low back pain patients effectively, it is not known if CT images are ordered according to guidelines.

Objectives. To evaluate the appropriateness of CT imaging referrals from family physicians for patients with low back pain.

Methods. A retrospective medical chart audit of administrative electronic health records was performed. All adult lumbar spine CT referrals ordered by all family doctor in one health region in Newfoundland and Labrador, Canada and performed between January 1st, 2016 and December 31st, 2016 were included. Each CT referral was identified and collected from two linked databases (Meditech and PACs). Data were manually extracted and categorised into three main groups: red flag indicated (determined to be an appropriate referral), radicular syndrome, or nonspecific low back pain (determined to be not appropriate).

Results. 3,596 lumbar spine CTs ordered by family doctors from 2016 were included. Demographic information indicates 54.5% of the records collected were female, with a mean age of 54.7 (SD 14 years). 5.5% (95% CI 4.3%-5.7%) of CTs referrals were for a suspected red flag condition, meaning only 5.5% were appropriate.

Interpretation. Guidelines recommend only ordering CTs for LBP when red flags are present. This audit found that the majority of CT referrals are being ordered for radicular syndrome, for which there are limited evidence-based guidelines for treatment.

4.2 Introduction

Low back pain (LBP) is a common health issue, identified as the leading cause of disability globally (1). Less than 1% of low back pain is due to a specific serious spinal condition (cancer, infection, cauda equina, or fracture) (2). If there is no indication of a serious pathology from the patient history or physical exam, LBP can be further classified as radicular syndrome (e.g., spinal stenosis, sciatica, radiculopathy or radicular pain) which occurs in 5 to 10% of cases or non-specific low back pain (NSLBP; defined as no cause that can be determined) which occurs in ~90-95% of all LBP cases (3). Patients with LBP often first seek treatment from their family physician (FP) (2).

Computed Tomography (CT), a form of diagnostic imaging (DI), is one modality used to investigate LBP. Guidelines from organisations such as Choosing Wisely Canada (CWC) recommend performing CTs of the lumbar spine (LS) to confirm the presence of a suspected serious pathology (4-6). In instances of NSLBP or radicular syndrome, imaging has limited use (4-6). In some cases of NSLBP or radicular syndrome where patients have not responded to conservative care and are considered potential candidates for surgery, guidelines state that DI would be recommended; however, CT may not be the best imaging modality (3,5,7). This is largely due to the safety risks posed to patients as one LS CT emits 6 mSV of radiation, compared to the 1.5 mSV of radiation from an x-ray (4,5,7-9). In fact, one lumbar spine CT emits 170 times the amount of radiation as a chest x-ray (9).

The prevalence of imaging for LBP has been reported to be high, given the small proportion of patients in which it will likely have a beneficial impact (10). Further investigation of the reasons for ordering imaging has indicated that 34.8% (95% CI: 27.1, 43.3) of all lumbar imaging were considered inappropriate when compared to red flag guidelines, and 31.6% (95% CI: 28.3, 35.1) were considered inappropriate when compared to guidelines regarding suspicion of clinical pathology (11). Few studies have quantified the appropriateness of CT, which is surprising given the safety risks to the patient and costs to the healthcare system (12). With a public healthcare system like Canada's, taking care to provide safe and affordable health services to the public should be a priority. However, there is variation in the quality of care provided in different provinces, with the province of Newfoundland and Labrador ordering more CTs than other Canadian provinces in 2012 and equal rates of all CTs as Ontario in 2017 (13,14).

Objective. To determine the proportion of LS CT referrals made by FPs that were to investigate symptoms of serious spinal pathology, radicular syndrome, and/or NSLBP. It is predicted that the proportion of red flag-indicated CT imaging for LBP will be very small, as the prevalence of serious spinal pathology is rare.

4.3 Methods

This study received ethical approval from the Newfoundland and Labrador Health Research Ethics Authority. The reporting of this study followed the

REporting of studies Conducted using Observational Routinely-collected Data (RECORD) checklist (15).

Study Design and Setting. We conducted a retrospective analysis of one year of CT imaging data using the administrative, electronic health records of the Eastern Health (EH) Regional Health Authority (RHA), in Newfoundland and Labrador (NL), Canada. Eastern Health is the largest of four RHAs that exist in NL, providing health services to over 300,000 individuals from approximately 13,000 health centres (16). There are seven hospitals with a radiology department within EH that perform CT imaging. Data were collected from January 1st, 2016 to December 31st, 2016

Study population: The reasons an FP ordered a CT image for adult patients with LBP were assessed. A family physician was defined as any physician who works in a family or general practice who had an ID code as a general practitioner or family physician. LBP due to spinal causes was the focus of our assessment. This included serious spinal pathologies which are cancer (including past history of cancer), infection, cauda equina, and fracture; radicular syndromes which include conditions like spinal stenosis, radiculopathy, radiating pain, sciatica; and non-specific causes which are defined as LBP from an unknown cause (3). LBP attributable to a non-spinal cause was excluded including, but not limited to, abdominal aortic aneurysm, pregnancy, or pancreatitis.

Data Eligibility, Data Sources, and Linkage. A third party retrieved a list of all patients that received a lumbar spine (LS) CT in 2016 using the billing codes for an LS CT with and without contrast from the Picture Archive and

Communication System (PACS) database. PACS is a medical, digital application that allows healthcare providers to store and view high-quality DI. Records were eligible for inclusion if an LS CT with or without contrast was performed between January 1st, 2016 and December 31st, 2016, the patient was older than 18 years, and the CT was ordered by an FP.

Patient CT referral forms were accessed from PACS, where the CT imaging, referral forms, and radiologist finding reports were also found. Demographic information was retrieved from the Meditech system, including age at the time of the scan, sex, and postal code. The EH RHA had already digitally linked these two databases.

Data Collection. Three research assistants collected referral form free-text data. The referral form from each patient record was retrieved, and the free text referral reason was transcribed into an Excel file word for word, using a codebook to ensure all physician shorthand was transcribed the same. Digital text from the radiology report in PACS was also collected and used only in instances where the referral form was illegible or missing. Due to time restrictions, it was not possible to perform a validation study of extracted data to ensure that each of the research assistants was being equally diligent at transcribing the physician handwriting.

Data Coding and Outcomes. The aim of this study was to determine the reason for referral and to categorise it into one of three categories: appropriate (defined as concordant with CWC recommendations regarding red flags), potentially appropriate (radicular syndromes) and not appropriate (non-specific routine back pain) (3,7). The definition for each category was determined by a

review of the evidence-based guidelines and CWC recommendations (3,7). Consultations with FPs occurred to ensure all referrals were captured accurately.

- Appropriate referral: For the serious spinal pathology category, keywords in the referral such as fracture, cancer, tumour, history of cancer, infection, and cauda equina, were clear and easily included into this category. Mentioning urinary retention and faecal incontinence, or any similar terms were determined to be indicators of suspected cauda equina, thus the referral was coded as red flag indicated. If there was a discussion of a recent injection at the LS, such as a lumbar puncture, the referral was assumed to be a suspected infection and was coded as a red flag pathology. Any referral with a history of cancer was also included; however, familial cancer history was not included, as a family history of cancer is not a strong predictor of spinal cancer (3,10). Any mention of a suspected fracture, including a compression fracture, was included in this group. If the patient had a remote history of fracture and the image was a follow up to this image, this was not considered a red flag pathology. As guidelines are not consistent regarding age >50 years as a red flag, it was not considered as an appropriate indicator for imaging.

- Potentially appropriate: For radicular syndrome, referrals with keywords of radiating pain, spinal stenosis, radiculopathy, or radicular symptoms were coded into a radicular syndrome category. If there was mention of pain below the hips, and specifically in the legs, without mentioning the

low back, it was assumed that the low back was affected and the symptoms indicated radicular syndrome.

- Not appropriate: Referrals for LBP, and any referral that did not contain keywords that described radiating pain, a specific congenital disease, or a red flag was placed in the NSLBP category. If the physician suspected an anatomical structure or a specific diagnosis in the patient that they suspected was causing pain but did not mention symptoms other than back pain, this was kept in the NSLBP category. This included conditions such as osteoarthritis, spondylosis, degenerative disc disease, etc. The reasoning behind this was that evidence shows that various groups of people have anatomical findings in their spine, but do not experience pain; thus, it is possible that findings in the spine such as disc bulges, and degenerative disc disease, might not be causing the patient's pain (3,4). Referrals that mentioned a history of surgery were ignored, and only the patient's symptoms were considered.

Process: Every referral form free text was coded according to the ordering physician's primary suspicion of the back pain source. This was done by searching for keywords in each referral text that indicated which of the three main categories the referral form best conformed to. If the ordering physician did not mention any key terms explicitly, the description of symptoms in the referral was reviewed with a research physician to determine which category was described in

the text. All referrals were then split into the three main categories above. Descriptions of these symptoms and examples follow (Table 1).

All data were coded by one researcher, GL, with approximately 10% of the data validated by a second person, AH, to minimise misclassification bias. See Table 1 for coding examples:

Table 4.1: Coding Terms with definitions and examples from referrals		
Category code	Definition	Examples of Referral form text
Appropriate		
Red Flag condition:	This refers to specific spinal pathologies requiring immediate imaging: cancer, fracture, cauda equina or infection symptoms.	"41 year old male multiple back surgeries now complains of increasing pain, difficulty urinating. He does say that he has had urinating difficulties more often and has been ongoing for several months. Diagnosis: Rule out cauda equina"
		"Back pain. Fall one week ago. ? Fracture L1. Pain out of proportion. Diagnosis: back pain"
Potentially appropriate		
Radicular Syndromes or Leg-dominant pain	This refers to the conditions of spinal stenosis radiculopathy or radicular pain (described as "radiation to legs", numbness, or shooting pain).	"Lower back with radiation to legs and numbness and tingling in her feet, shooting pain in toes. Diagnosis: Low back pain"
		"numbness left leg, mechanical low back pain"
		"Patient with radicular back into the gluteal region. Patient with x-ray L spine with OA.

		Diagnosis: Rule out nerve root compression."
Not-appropriate		
Non-specific Low Back Pain	This refers to any referral that did not describe symptoms that suggested a red flag or leg pain.	"Persistent low back pain. Degenerative disc disease with L3-4 narrowing. Diagnosis: ? Discogenic low back pain"
		"Increasing back pain. Diagnosis: OA"

Data Cleaning and Analysis. Statistical Package for the Social Sciences (SPSS, IMB®,**version** 25.0.0.0) was used to generate descriptive statistics for this dataset. Data cleaning was performed to ensure all pertinent information was available, that referrals were classified appropriately, and the dataset was deidentified. Percentages/proportions were generated for all referral codes, and confidence intervals (CI) for a single proportion were calculated (17).

4.4 Results

In 2016, there were 4,435 LS CTs ordered by any physician in the Eastern Health Region in Newfoundland and Labrador. 82% (n=3,655) were ordered by FPs and retained for analysis. Eleven records were excluded due to the patient's age (< 18) and forty-eight records were excluded for a suspected cause not related to the lumbar spine (e.g. post-partum pregnancy-related back pain, thoracic spine) leaving 3,596 records (Figure 4.1). The majority of data were obtained from the imaging referral; however, in 41 cases (1.13%), the referral was unavailable, and the physician's reason for referral was obtained from the

corresponding radiology report. There were an additional 69 cases (1.9%) where the referral form was missing, illegible, or did not provide enough information to code accurately. Participants who received CT imaging had a mean age of 54.7 (SD 14 years), of which 54.5% were women (Table 4.1). 5.3% of CT referrals mentioned a past history of surgery, and 6.1% of CT referrals mentioned a past history of trauma (e.g., fall or motor vehicle accident).
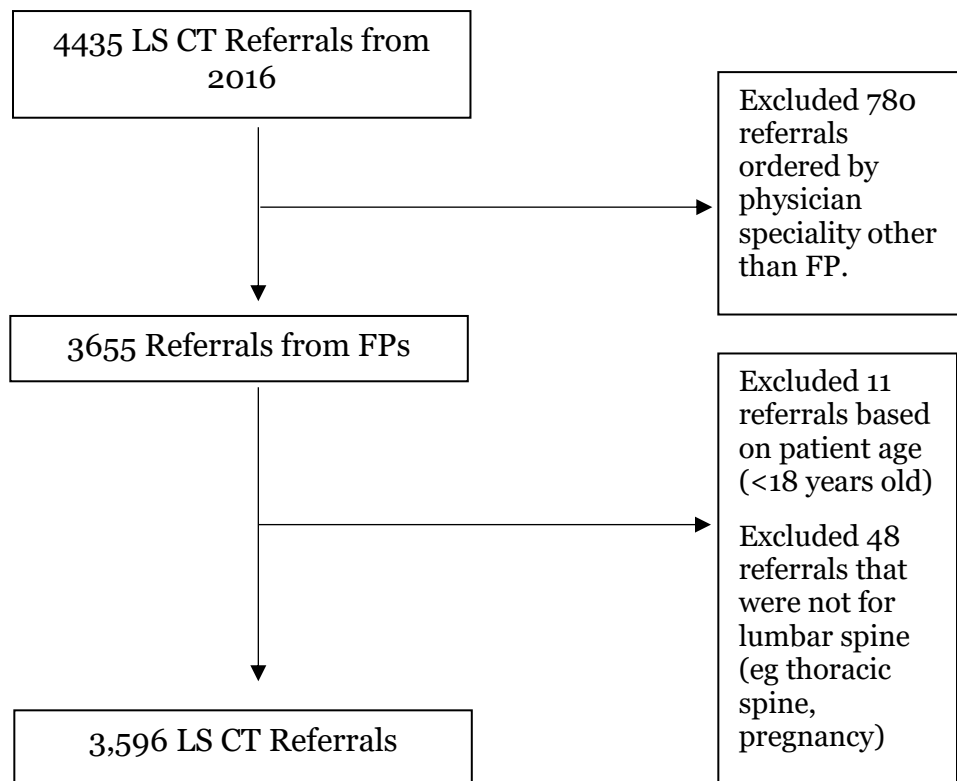
```
┌─────────────────────────────┐
│  4435 LS CT Referrals from  │
│             2016            │
└─────────────────────────────┘
              │
              │          ┌──────────────────────┐
              │─────────>│ Excluded 780         │
              │          │ referrals            │
              │          │ ordered by           │
              │          │ physician            │
              │          │ speciality other     │
              │          │ than FP.             │
              ↓          └──────────────────────┘
┌─────────────────────────────┐
│   3655 Referrals from FPs   │
└─────────────────────────────┘
              │          ┌──────────────────────┐
              │          │ Excluded 11          │
              │          │ referrals based      │
              │          │ on patient age       │
              │          │ (<18 years old)      │
              │─────────>│                      │
              │          │ Excluded 48          │
              │          │ referrals that       │
              │          │ were not for         │
              │          │ lumbar spine         │
              │          │ (eg thoracic         │
              │          │ spine,               │
              ↓          │ pregnancy)           │
┌─────────────────────────────┐ └──────────────────┘
│    3,596 LS CT Referrals    │
└─────────────────────────────┘
```

Figure 4.1. Flow diagram of included and excluded images from a medical record review of all LS CTs in 2016

Reasons for Referral

Red Flag indicated referrals: In 5.5% of referral forms, FPs indicated they suspected a red flag condition as the primary reason for ordering the CT image (Table 4.2). Of these red flag conditions, the most common red flag suspected condition was for cancer/tumour or history of cancer (2.4% of referrals). Fractures were suspected in 2.2% of referrals, cauda equina in 0.8% of referrals, and infection in 0.2% of referrals.

Radicular Syndrome: 75.8% of the referrals mentioned radicular syndrome symptoms. There were 233 referrals specifically for investigation of spinal stenosis (6.5% of the total referrals, 8.5% of the radicular syndrome category).

Non-specific LBP: 16.8% described symptoms that indicated NSLBP, meaning that there was no sign of radiating pain, the source of pain was unknown, and there were no indications of red flags.

Referral Appropriateness

Using our definitions for appropriateness, only 5.5% of lumbar CTs ordered for patients with LBP were found to be concordant with CWC recommendations or Bardin guidelines and therefore are considered to be appropriate (3,7). 16.8% were found to be non-concordant and thus inappropriate. The vast majority of referrals (75.8%) were ordered for reasons related to radicular syndromes and thus considered potentially appropriate. More information on these referrals

would be needed to determine appropriateness that was not available in the data

collected.

Table 4.2. Descriptive information and reasons for CT referral for all lumbar CTs by FPs for patients (over 18 years) with LBP in 2016 in EH RHA, NL, Canada.

Total number of CTs eligible for analysis	N= 3,596	
<u>Demographic variables</u>	**Mean (SD)**	
Percent female	54.5%	
Mean age	54.7 years (14 years)	
Referrals that mentioned previous history of surgery	5%	
<u>Reason for referral</u>	**Frequency**	**%**
Red Flag	197	5.5
Radicular Syndrome	2727	75.8
Non-specific LBP	604	16.8
Missing/No indications listed	69	1.9
Total	3,596	100

4.5 Interpretation

Findings

We found that of the 3,596 LS CT included referrals from the EH region, only 5.5% of them were ordered for a suspected red flag condition aligning with our definition of appropriateness based on evidence-based guidelines. Approximately 16.8% were for NSLBP and considered inappropriate. The largest proportion of referrals (75.8%) were ordered for patients with symptoms that indicated radicular syndrome; in most of these cases, however, there was insufficient or missing information pertaining to a complete clinical neurological exam to distinguish between radiating leg pain and radiculopathy, thus these cases were considered potentially appropriate.

This is the first study that has examined the appropriateness of LS CTs in Canada and addresses an important gap in the field of test overuse and patient safety (12,18). It also adds to the body of international research in this area, given there are only two other studies that have provided estimates of appropriate LS CTs (9,19). We used guidance from the STROBE and RECORD statements to improve rigour and ensure transparency in reporting (15,20). To our knowledge, it is the largest adult sample of LS CT referrals to have been reviewed.

Explanation of Findings

No publications on the appropriateness of LS CTs ordered in family practice settings exist making direct comparisons difficult. Recently, a meta-analysis of imaging appropriateness using a composite score from x-ray, MRI and

CT referrals found that approximately 34.8% were inappropriate due to the absence of a red flag indication (11). Only two other studies have been able to provide estimates of LS CT appropriateness, both of which were conducted in different settings with different definitions of appropriateness (9,19). For example, one study in a single hospital had a very specific definition of appropriateness where only CTs ordered for a suspected fracture were considered appropriate, and found 23% referrals to be appropriate (9). The second study was conducted in an ED setting and used a more sensitive definition where pain duration of six weeks and red flag indicators were appropriate reasons for a referral, and they found 56% appropriateness (19). These studies suggest that the percentage of appropriate referrals varies depending on study methodologies (e.g., definition and setting). Though campaigns to reduce DI have been present for years, large proportions of potentially inappropriate CTs are still occurring. This is indicative of overutilization and wastefulness in the healthcare systems and lack of due diligence for patient safety.

AHCPR, ACR, and AHRQ guidelines state that the presence of radicular syndrome is not a good indicator for ordering imaging unless the patient requires a change in treatment like surgery (5,21,22). We found that 92.6% of CT referrals that were reviewed in our current study were either ordered in contradiction of these recommendations or missing information that would indicate an intention to change treatment. The tendency for FPs to refer this group of patients for CT imaging is supported by previous survey research. Webster et al. and Negrini et al. showed that physicians are more likely to order imaging for patients when

there are symptoms of back-related leg pain regardless of guideline recommendations, and another study found that the presence of back-related leg pain was also a predictor of imaging (23-25). The high proportion of imaging in our study, combined with the findings from Webster, Negrini, and Kovacs, all demonstrate that while guidelines imaging is not necessary, physicians likely find some utility from it. Qualitative studies exploring physicians' reasons for using imaging in the absence of red flags often report that physicians use imaging as a reassurance tool, to satisfy patient demand or to expedite referrals to orthopaedic surgical consults (18,26). Further research is warranted to understand what clinical utility physicians gain from using CTs for patients presenting with leg-pain.

Limitations

The main limitation of this study was the quality of the information from the referral form, as this study relied on routinely collected data that was not collected for research purposes. There may have been important information missing from the referral form that was mentioned in the physical exam in the doctor's office but not written on the form or more serious terminology used to get a quicker response from the radiologists. The clarity in which the physician wrote the referral was also a limitation as the more unclear the handwriting, the more likely vital information was missed, causing misclassification. Another limitation is potential differences in data transcribing. There were three different research assistants working on data extraction, but it is unclear if they all were

92

able to read and extract information from the referral forms at the same quality level. However, all research assistants were provided with training and a codebook to limit discrepancies between the transcriptions. A third party provided the researchers with the dataset, so it is possible that there were mistakes made on the third party's side that could not be verified by the researchers. There is no way of knowing if many patients were receiving CT imaging due to conditions that excluded them from MRI magnetic field exposure (e.g., Pacemakers).

Future Research

This study aimed to better understand if physicians are following the red flag suspicion guidelines for ordering CT imaging and, if they are not following the guidelines, to determine for what they are ordering CT imaging. Now that there is a better understanding that a high proportion is ordered for radicular syndrome, research is needed to focus on the proportion of patients who have radicular syndrome and what the evidence states is the best care for these patients. Radicular syndrome treatments are inconsistent and often not evidence-based, thus research into what patients currently receive for treatment compared to the guidelines would be beneficial, as well as understanding the knowledge physicians have on radicular syndrome treatment. Studies that aim to assess physicians' knowledge and awareness of guidelines, as well as their ability to prescribe certain treatments and the patient's ability to access the treatment for radicular syndromes would be beneficial. It would also be valuable to understand

physicians' perspectives on the clinical utility of ordering CT imaging. Some

referral forms mentioned that patients themselves requested a CT, so it would be

useful to investigate how often a patient requests an image, and what benefits

patients expect to receive from it. Interventions to decrease unnecessary imaging

with timely referral to an allied health professional are important to provide to

patients to ensure that they are having their concerns addressed. Estimating the

potential cost-savings to the healthcare system would also be beneficial to

policymakers.

4.6 References

(1) Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet 2016 10/08; 2018/08;388(10053):1545-1602.

(2) Maher CG, Williams C, Lin C, Latimer J. Managing low back pain in primary care. Aust Prescr 2011 /;34(5):128-132.

(3) Bardin LD, King P, Maher CG. Diagnostic triage for low back pain: a practical approach for primary care. Med J Aust 2017 Apr 3;206(6):268-273.

(4) Chou R, Deyo RA, Jarvik JG. Appropriate use of lumbar imaging for evaluation of low back pain. Radiol Clin North Am 2012 Jul;50(4):569-585.

(5) Patel ND, Broderick DF, Burns J, Deshmukh TK, Fries IB, Harvey HB, et al. ACR appropriateness criteria low back pain. Journal of the American College of Radiology 2016;13(9):1069-1078.

(6) Davis PC, Wippold FJ,2nd, Brunberg JA, Cornelius RS, De La Paz RL, Dormont PD, et al. ACR Appropriateness Criteria on low back pain. J Am Coll Radiol 2009 Jun;6(6):401-407.

(7) Choosing Wisely Canada. Imaging Tests for Lower Back Pain: When you need them and when you don't. Available at:

https://choosingwiselycanada.org/imaging-tests-low-back-pain/. Accessed January 9, 2019.

(8) Traeger A, Buchbinder R, Harris I, Maher C. Diagnosis and management of low-back pain in primary care. CMAJ 2017 Nov 13;189(45):E1386-E1395.

(9) Oikarinen H, Meriläinen S, Pääkkö E, Karttunen A, Nieminen MT, Tervonen O. Unjustified CT examinations in young patients. Eur Radiol 2009 2009/;19(5):1161-1165.

(10) Downie A, Hancock M, Jenkins H, Buchbinder R, Harris I, Underwood M, et al. How common is imaging for low back pain in primary and emergency care? Systematic review and meta-analysis of over 4 million imaging requests across 21 years. Br J Sports Med 2019 Feb 13.

(11) Jenkins HJ, Downie AS, Maher CG, Moloney NA, Magnussen JS, Hancock MJ. Imaging for low back pain: is clinical use consistent with guidelines? A systematic review and meta-analysis. Spine J 2018 May 3.

(12) Charlesworth CJ, Meath THA, Schwartz AL, McConnell KJ. Comparison of low-value care in medicaid vs commercially insured populations. JAMA Intern Med 2016 2016/07;176(7):998-1004.

(13) Canadian Institute of Health Information. Medical Imaging in Canada 2012. 2013; Available at: https://www.cihi.ca/en/mit_summary_2012_en.pdf. Accessed January 9, 2019.

(14) Canadian Agency for Drugs and Technologies in Health. The Canadian Medical Imaging Inventory, 2017. 2018; Available at: https://cadth.ca/canadian-medical-imaging-inventory-2017. Accessed 05/21, 2019.

(15) Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Peteresen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLoS Med 2015 2015/;12(10).

(16) Eastern Health. About Us 2018; Available at: http://www.easternhealth.ca/AboutEH.aspx. Accessed May 5, 2019.

(17) Department of Statistics and Data Science. Inference for Categorical Data. Available at: http://www.stat.yale.edu/Courses/1997-98/101/catinf.htm. Accessed May 06, 2019.

(18) Busse J, Alexander P, Abdul-Razzak A, Riva J, Alabousi M, Dufton J. Appropriateness of spinal imaging use in Canada. Hamilton, ON: McMaster University 2013.

(19) Schlemmer E, Mitchiner JC, Brown M, Wasilevich E. Imaging during low back pain ED visits: A claims-based descriptive analysis. Am J Emerg Med 2015 2015/;33(3):414-418.

(20) Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. Int J Surg 2014 2014/12;12(12):1500-1524.

(21) John M. Eisenberg Center for Clinical Decisions and Communications Science. Noninvasive Treatments for Low Back Pain: Current State of the Evidence. Comparative Effectiveness Review Summary Guides for Clinicians Rockville (MD); 2007.

(22) Bigos S, Bowyer O, Braen G, Brown K, Deyo R. Acute Low Back Pain in Adults: Clinical Practice Guideline No. 14. AHCPR Publication No. 95-0642. Rockville, MD: Agency for Health Care Policy and Research. Public Health Service, US Department of Health and Human Services 1994.

(23) Kovacs FM, Fernández C, Cordero A, Muriel A, González-Luján L, del Real, María Teresa Gil. Non-specific low back pain in primary care in the Spanish National Health Service: a prospective study on clinical outcomes and determinants of management. BMC Health Services Research 2006;6(1):57.

(24) Webster BS, Courtney TK, Huang YH, Matz S, Christiani DC. Physicians' initial management of acute low back pain versus evidence-based guidelines. Influence of sciatica. J Gen Intern Med 2005 /;20(12):1132-1135.

(25) Negrini S, Politano E, Carabalona R, Mambrini A. General practitioners' management of low back pain: impact of clinical guidelines in a non-English-speaking country. Spine (Phila Pa 1976) 2001 Dec 15;26(24):2727-33; discussion 2734.

(26) Hall AM, Scurrey SR, Pike AE, Albury C, Richmond HL, Matthews J, et al. Physician-reported barriers to using evidence-based recommendations for low back pain in clinical practice: a systematic review and synthesis of qualitative studies using the Theoretical Domains Framework. Implementation Science 2019;14(1):49.

CHAPTER 5: REFLECTION ON PATIENT ENGAGEMENT

5.1 Patient-oriented research

Medical research is conducted by highly skilled and highly educated researchers making decisions about what matters for patients. However, not everything a researcher determines is important for a project aligns with what is important for patients who will be directly impacted by these decisions (1-3). There may be outcomes, treatments, and effects that have little meaning to patients with the disease or condition that is described (3). Recently, there has been a move to incorporate patients' ideas and priorities into research methods, called Patient-Oriented Research (POR), engaging them in various ways to ensure that the outcomes are important to patients, and the healthcare system is more ethical (1,4). There are various ways of addressing patient ideas and priorities into research designs, with high to low levels of engaging patients.

5.2 Patient Engagement

POR involves patient engagement, which partners with patients at some or all stages of the research process to ensure that the patient's input is meaningful and timely (5). Patients, who are typically anyone who has the condition of research interest or is a family member or caregiver of someone with the condition of interest, that assist with research, and are not the subject that is being researched, are called patient partners. Patient partners are different from research participants as they contribute to the protocol and execution of a study

and do not have data collected from them as a part of answering a research question. Patient engagement is typically thought of as a continuum, where there are different levels of engaging patients, according to various factors, such as time, budget, project, and patient partner willingness to participate.

Patient engagement tools have been developed, and serve as a template for answering the who, what, when, how, where, and why of research method development (5-7). These plans involve patient partners at different levels of engagement on the continuum, that correspond with the goals of the research project. The levels of engagement form a pyramid and give patients different levels of decision-making power, with the basic level involving informing patient partners of the research being conducted in an objective, unbiased way, requiring nothing in return from the patient partners (5,7). The second level, "consulting patient partners," involves slightly more input from patient partners, allowing the patients to provide feedback on the research objectives, but in a more passive manner (e.g., after the objectives have already been mostly formed) (7). The third level is involving patient partners in the research project by directly giving patients the ability to consult on project objectives and allowing them to ensure that the relevant patient identified objectives are being considered (7). The fourth level of patient engagement allows patients to collaborate on the research objectives and involves them more in-depth in the research design decisions (7). The fifth and final level empowers patients to make all the decisions regarding the project objectives, allowing them to contribute in a manner that is most

meaningful to the patients (7). Each level comes with different pros and cons, and it is ultimately up to the research team to decide the best way to engage, depending on time, funding, and other constraints to the researchers.

Evaluation of the patient engagement plan is important to the process, but often can be challenging to incorporate in practice (8). It requires measuring whether or not the patients felt they made an impact, and measuring, from the research project, if the patient engagement significantly changed the project in any way. Methods of evaluation can be simple, such as getting feedback from partners, or more complicated, such as a pre- and post- surveys to all involved (6,7,9). Evaluating patient engagement is important because very little is known about the effectiveness of patient engagement.

5.3 Patient engagement for the Clinical Audit

For the low back pain medical record review for this thesis, (Chapter 4) there was a desire to include patient engagement since this is such a common disorder. The level of patient engagement used in Chapter 4 of this thesis was "involving" (the third level) patient partners to ensure that the objectives were framed with patient input. The patient involvement occurred December 2016, in the form of a world café, where all interested stakeholders, ranging from clinicians, allied health professionals, researchers, and patient partners, were given the ability to directly comment and advise on the research objectives and the research question. Different tables were set up with all stakeholders being represented at each table. The tables were presented with questions to respond to

and were asked to present their responses back to the larger group. All this information was collected and documented.

5.4 Patient and Public Engagement Activities and Reflection

The specific patient and public engagement activities that I participated in for this Masters thesis are as follows:

1) I was an observer and volunteer at the World Café event in December 2016.

2) I presented my project at TPMI Scientific Day October 2017.

3) I took a graduate student level Patient Engagement Course Fall 2017.

4) I attended a provincial Patient Advisory Council meeting December 2017.

5) I presented at the SHARE Summit in October 2018.

6) I volunteered at the Health for All Festival in Nov 2017 and Nov 2018.

I learned a great deal regarding the theory and practicality of public and patient engagement throughout the duration of my Masters program. I learned that it is important to maintain relationships with patients, provide frequent updates to patient advisors, and to communicate with the public for knowledge translation purposes. While there are many challenges with patient engagement, I learned that there is great value to be provided by engaging perspectives outside of academia and research.

This thesis engaged patients at the planning level and received patient guidance on our objectives. While patients did provide guidance at the planning stage of this study, there was limited time to receive further input from patients

as to how their perspectives would frame the results of this thesis. In the future, projects similar to this could utilise patients further in knowledge translation, dissemination, and manuscript drafting.

5.5 References

(1) Canadian Institutes of Health Research. Patient engagement. 2018; Available at: http://www.cihr-irsc.gc.ca/e/45851.html. Accessed 05/14, 2019.

(2) Manafo E, Petermann L, Mason-Lai P, Vandall-Walker V. Patient engagement in Canada: a scoping review of the 'how'and 'what'of patient engagement in health research. Health research policy and systems 2018;16(1):5.

(3) Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet 2009 Jul 4;374(9683):86-89.

(4) Lough S. Need to define patient engagement in research. CMAJ 2015 Sep 8;187(12):E385-6.

(5) Involve N. Briefing notes for researchers: involving the public in NHS, public health and social care research. INVOLVE Eastleigh, UK 2012.

(6) Canadian Institutes of Health Research. Strategy for Patient-Oriented Research - Patient Engagement Framework. 2014; Available at: http://www.cihr-irsc.gc.ca/e/48413.html. Accessed 05/14, 2019.

(7) Vat L. Patient and Public Engagement Planning Template. 2016; Available at: http://nlsupport.ca/getattachment/8dc1f539-d225-46fa-ba8d-2d06da934486/Patient-and-Public-Engagement-Planning-Template.pdf.aspx. Accessed 05/14, 2019.

(8) Abelson J, Humphrey A, Syrowatka A, Bidonde J, Judd M. Evaluating patient, family and public engagement in health services improvement and system redesign. Healthc Q 2018;21:61-67.

(9) Esmail L, Moore E, Rein A. Evaluating patient and stakeholder engagement in research: moving from theory to practice. Journal of comparative effectiveness research 2015;4(2):133-145.

CHAPTER 6: SUMMARY AND CONCLUSION

6.1 Summary of Findings

The objective of this thesis was to investigate the appropriateness of radiological imaging for low back pain in primary care settings according to evidence-based guidelines both globally and locally. Three different studies were undertaken for this thesis to determine what is known regarding radiation emitting imaging and imaging adherence to the guidelines for low back pain. Each study illuminated the picture of imaging appropriateness for modalities that expose patients to carcinogenic radiation.

6.1.1 Summary of Utilization trends

A report from CIHI stated that there was a high rate of CT imaging in Newfoundland and Labrador, so a study on the rates was undertaken in order to investigate local imaging trends more closely (1). The most recent rate, while only reflective of the largest health region in NL, shows that there were 1,395 CTs/100,000 persons (age-sex standardised) ordered in 2016. When looking at the rates of CT ordering over the years, there was a rate increase in CT ordering from 2013 to 2015, but the rate decreased in 2016. While rate-ratios indicated that there was a significant difference in the rates, it is likely that the difference was not relevant in a clinical sense. More data on more recent years is needed to determine if there is an absolute trend in increasing CT use.

6.1.2 Summary of SRMA

The second study utilized a PRISMA Systematic Review and Meta-analysis methodology, synthesizing research findings of what is known regarding x-ray and CT imaging appropriateness globally (2). Six articles were included that reported information on x-ray and CT appropriateness, but only five of these studies were combined in the meta-analysis using a single proportion model. There were three studies to individually report x-ray appropriateness outcomes, one study that reported x-ray and CT appropriateness outcomes individually, and one study that reported CT appropriateness outcomes. The GRADE approach was also used to evaluate the level with which we are confident in the meta-analysis and the evidence quality (3). Many different guidelines were used to determine, with low-quality evidence, that 44% (95% CI: 34%, 54%) of x-rays are appropriately ordered, and, with very low-quality evidence, that 54% (95% CI: 51%, 58%) of CTs are ordered appropriately.

6.1.3 Summary of Medical Record Review

The results of the medical record review show that only 5.5% of the 3,596 CTs performed in the Eastern Health (EH) Region in 2016 were appropriate to investigate a suspicion of a red flag condition. 75.8% of CT referrals were to investigate symptoms of radicular syndrome, which were considered potentially appropriate investigations (including disorders such as spinal stenosis, radiculopathy, sciatica, and radiating pain). 16.8% of CTs were performed to investigate non-specific LBP and considered inappropriate. While there is a low rate of appropriate imaging, determining the proportion of inappropriate imaging

is less clear, due to the lack of information provided on the CT referral forms. High-quality treatment options for patients with radicular syndrome are limited, and while surgery is a potential option, not all patients with radicular syndrome are candidates (4). There is also evidence that patients request imaging, so there needs to be an understanding as to why that occurs.

6.2 Findings in Context with other Literature

6.2.1 CT Age-Sex Standardised Rates in context

Very few studies have published data on the rates of CT image ordering over time to compare the rates of CT ordering in NL. Two studies have been published in the United States and Australia but cannot be directly compared to our context as the reference population used was unknown (5,6). No peer-reviewed articles have been published in Canada, though grey literature does exist on this subject (7). However, this type of data is important to report because it provides a metric for other health authorities to compare their own performance and provides further proof of the overutilization of CT imaging.

In Canada, a grey literature study has been published on the different rates of CT imaging in Ontario and Manitoba (7). It found that in Ontario in 2010/2011, the age-sex standardised rate of LS CTs was approximately 660/100,000 persons, and in Manitoba, it was approximately 1,000/100,000 persons. In our context, the data collected that was closest in time period to the Ontario and Manitoba data was from 2013, where the rate of CT ordering was found to be 1224.9/100,000. This number is double the rate in Ontario and is

greater than Manitoba's rate by 22%. This difference in rates shows that Newfoundland and Labrador, even with its smaller population, utilizes more CT imaging than much larger provinces.

While in a Canadian context our rates in NL are high, when comparing these rates to other countries, the rates are contextualised. In the US, one study found that rates varied between 320 LS CTs per 100,000 people (age, sex, and race standardised) and 2, 370 LS CTs per 100,000 people (5). A government-commissioned study in Australia found similar variation in rates of LS CTs across the country, which ranged from 209 LS CTs per 100,000 people to 2,464 LS CTs per 100,000 people (6). While they are not a perfect comparison due to different reference populations, these ranges show that in an international context with similar high-income countries, NL rates are neither the highest nor the lowest.

6.2.2 Systematic Review and Meta-Analysis

A synthesis of the published literature was conducted to determine how many radiation-emitting images have been performed according to guidelines. No other systematic review published has synthesized this type of data, thus it is challenging to compare it to other available literature. A SR & MA by Jenkins et al. (2018) investigated whether imaging is over and under ordered (8). The main difference between the two SR & MAs is that our SR & MA focused on appropriate imaging for CT and x-ray, while Jenkins et al. focused on inappropriate over-ordering and inappropriate under-ordering of the three main imaging modalities. Our study also had very specific inclusion criterion, whereas Jenkins et al. had a

much more sensitive inclusion criterion (8). For example, Jenkins et al. included providers such as chiropractors and imaging modalities, including MRI in their analysis. Chiropractors were excluded from our systematic review and meta-analysis because Chiropractors are not considered primary care in Canada and, depending on the province, may not be services covered by the Canadian public healthcare system. MRI was also excluded from our SR & MA because even though MRI is far more costly, only CT and x-ray expose patients to harmful carcinogenic ionising radiation.

The review by Jenkins (2018) using a meta-analysis, found that of all of the referrals, DI was ordered inappropriately 34.8% according to red flag suspicions criteria, while 31.6% were inappropriate according to the criteria of no suspicion of clinical pathology (8). In comparison, our study found that x-rays were ordered appropriately 44% of the time, and CTs were ordered appropriately 54% of the time. However, the main difference between our SR and Jenkins et al., is that Jenkins et al. combined all image types in their meta-analysis and stratified the data based on the denominator and the different definitions of appropriateness (8). Denominators were either based on referrals or patient presentation with back pain. It is noteworthy that the Jenkins et al. (2018) systematic review and meta-analysis had a much more robust search strategy that yielded more results than this thesis' systematic review and meta-analysis did (8).

6.2.3 Medical Record Review of CT imaging for LBP

As the SRMA suggested with low-quality evidence, globally, there may be a much higher proportion of appropriate imaging referrals for CT occurring elsewhere than here in Newfoundland and Labrador. However, since there were only two studies that met the inclusion criteria that specifically reported appropriateness of CT imaging, it is difficult to compare our local context to the global stage directly. This is likely because these two studies were very different in the guidelines and the sample sizes that they used. However, there is also a third study on CT imaging appropriateness that was conducted in France but was excluded from this thesis (Chapter 3) based on the guideline year. However, the findings of this study are also worth discussing in the context of our findings.

The first study by Oikarinen et al. (2009), included in the SR has limited applicability to our context due to differences in setting, guidelines, and population (9). The study population was patients who were younger than 35, and these patients came from one hospital in Finland. For defining LS CTs appropriateness, only images that were ordered for a suspected fracture were determined to be justified according to guidelines issued by the European Commission for this age range. Also, this study was investigating any type of CT image to determine if it was appropriate, or "justified", thus the total sample size that reported CT imaging for the lumbar spine was very small (9). The final number that they used in their analysis was 30 total LS CT referrals, and they did not provide a rationale as to why they used that number. Even so, 23% of the CT scans were determined to be appropriate.

Schlemmer et al. (2015) were also included in the systematic review, and though they had a much more robust sample, the guidelines that they used to determine appropriateness were much more inclusive (10). This definition of appropriateness allowed for any image that was ordered for a suspected red flag condition, or for back pain that persisted for longer than 6-weeks as appropriate. Thus, this study found that 56.2% of the 648 CT images included were appropriate (10). However, this study also included data from only Blue Cross insured patients who visited any Emergency Department across one state in the USA; thus, this high proportion of appropriateness may be to the broad definition of appropriateness and the fact that physicians who are caring for patients with LBP who are insured may have better documentation on which to determine appropriateness.

Hourcade et al. (2002) found that 25% of the 132 LS CTs that were included for LBP followed guidelines appropriately (11). This study further broke down the CTs by what the type of low back pain was and found that out of the 24 patients with NSLBP, 63% of them were imaged appropriately according to guidelines regarding a seven-week wait period for NSLBP (11). In patients with chronic LBP, none of the CTs were ordered according to guidelines which state that an MRI should be ordered if the pain is severely limiting a patient's activities. Patients with sciatica were appropriately imaged in 35% of the 80 LS CTs when following guidelines regarding imaging only if the patient is a surgical candidate and after waiting four weeks.

Hourcade et al. (2002) were excluded from the systematic review due to the outdated guidelines they used; however, the findings of this study are interesting to note in the context of the current clinical audit in NL (11). It is worth noting that our audit only determined that 5.5% of 3,596 CT referrals were appropriate according to red flag indications. None of the above discussed studies solely focused on the four main red flags for LBP and most included imaging that was done a certain length of time after the first complaint as appropriate. We did not include that guideline in our appropriate definition because: I) it was not always possible to determine the timeframe of the complaint from the CT referral due to insufficient charting, and II) Choosing Wisely guidelines discuss red flag indicators and to get imaging after self-treatment methods have been attempted, but determining self-treatment history from the referrals was also not possible (7). As we used a much more restrictive definition of appropriateness, it is interesting to see the proportion of appropriateness is reflective of this definition.

6.3 Limitations and Strengths

6.3.1 Limitations

Each of the three parts that make up this thesis had their limitations discussed in their individual chapters. A brief summary of each chapter's limitations is summarized below.

CT Utilization Rates.

The limitations of this study were largely due to the years selected for this study. Having more years available would have provided this study with more

data to draw reliable conclusions if there is an increasing trend in ordering CTs in NL. The study was also missing age-sex standardised data for 2016, thus these calculations may not be completely accurate. There was no current data from other provinces or countries to which we could compare our rates of LS CTs. Rates that are available from Manitoba and Ontario, USA, and Australia cannot be used in a direct statistical comparison due to the difference in statistical reference standards used in these different populations (5-7). Thus, the ability to draw conclusions from the comparisons between locations is also limited in scope.

Systematic Review and Meta-Analysis on LS CT Appropriateness.

The largest limitation to this study was the search strategy used to find eligible studies. It is likely that the search strategy was not sensitive enough. This limited search strategy was done mainly due to limited resources but introduced the possibility that relevant studies were missed. To minimise the effects of this limitation, forward and backward tracking was done on all relevant studies.

Another limitation was the lack of a protocol registered on Prospero. Since a rigid protocol was not registered, there was a lack of transparency as to where the review diverged from the protocol. Though this is not a requirement for a Systematic Review, it provides rigour to the study, which is always beneficial. Without registering this review, there is less reported rigour.

CT Imaging Clinical Audit.

The most important limitation for this study was the quality of information that was provided on the CT referral forms. The information came from free-text and was often handwritten. Many physicians did not provide accurate patient histories or accurate suspected diagnoses, and it is likely that this was due to having to handwrite all the information. This also meant that when directly transcribing the data from the referral, it is possible the information contained on it was diluted due to the different transcribers' ability to read the information contained in the referral accurately. Since there were three transcribers and limited resources to validate the referrals, there is no way to ensure the transcriptions were accurate.

We identified three main categories for patients in order to determine which proportion of CTs were appropriate (4). There are very clear guidelines for Family Physicians on when to order CT imaging for patients with non-specific low back pain. However, for situations where a patient was categorised as radicular syndrome, the guidelines are far less clear and evidence-based tests, and treatment options are harder to find. Often, guidelines agree that if a patient is a surgical candidate, then CT imaging is appropriate. However, with the data we were provided on CT referrals, determining if the patient was being considered for spinal surgery was not clear. Thus, we used a strict definition of appropriateness and did not label images as inappropriate due to this grey area in the radicular syndrome category.

Another limitation to the study is that verification of the accuracy of the data provided by the third party who provided us with 2016 data was not possible due to resource and time limitations. This is a limitation because we can't be fully confident that all patients who received a CT were captured in the initial dataset.

6.3.2 Strengths

CT Utilization Rates.

The strength of the brief study on CT utilization rates in Newfoundland and Labrador was that no study like this has really been published in a peer-reviewed journal for a region like ours. By conducting this study and publishing it, we are contributing valuable information in the form of baseline data. Improving rates of CT utilization is impossible if there are no metrics to which to compare ongoing data.

Systematic Review and Meta-Analysis on LS CT Appropriateness.

Though there were limitations in this study, many methods were undertaken to ensure the results were trustworthy and performed with rigour. As mentioned, to combat the limitations with the search strategy, forward and backward tracking was performed on all included and applicable studies. This resulted in finding more articles that were eligible for inclusion. A recent, related systematic review and meta-analysis was also reviewed for articles that were eligible.

We adhered to the PRISMA checklist for conducting a systematic review and meta-analysis, ensuring our reporting was transparent. We also used the RECORD checklist on all included studies to identify limitations in the included studies. Since many studies did not fulfill the RECORD Checklist criteria, we were able to recommend that this be an area a focus for future similar studies. We also performed a GRADE analysis on the included studies in order to confidently make a statement regarding the findings of the systematic review.

CT Imaging Clinical Audit.

No published study has looked at this amount of CT imaging data for LBP and reported it according to strict, transparent guidelines. This clinical audit was reported according to the RECORD statement, which is an extension of the STROBE. This reporting checklist was created specifically for studies reporting data that was not initially collected for health research, such as medical records. By following this checklist, we ensured that the methodology was transparent, and all aspects of the study were properly reported. No study that has been done on appropriateness has mentioned following a reporting checklist and thus is missing critical information that affects the quality of the study.

Another strength, as briefly mentioned, was the large sample size included, which was collected over a year. Similar studies have only included a small sample size, with Schlemmer et al. having the largest at 648, and the smallest sample size of 30 belonging to Oikarinen et al. (9,10). By auditing a full year of

data, we can be confident that there was no selection bias or unintended, random or seasonal variations that may influence the proportions of appropriateness.

Finally, this study was strengthened by the clear way that appropriateness was reported. One of the main findings of the systematic review and meta-analysis was that most studies on appropriateness only reported vague definitions as to how referrals or images were categorized as appropriate, and would usually just mention the type of guidelines used. By reporting our criteria explicitly, we are providing readers with our precise reasoning as to how we classified the referrals and the ability to reproduce our study. As our dataset was limited in quality, classifying our data as appropriate, potentially appropriate, and inappropriate opens up opportunities to research reasons for potentially appropriate imaging more in-depth in the future.

6.4 Future research

Though this thesis illuminated many aspects of imaging for patients with low back pain, more research is needed to further our understanding of this global issue. We previously mentioned that there is no globally accepted definition of appropriateness, which limits our ability to generalise result findings from one region to the next. Researchers and experts in the field of back pain and imaging should attempt consensus in order to provide a unified definition to allow hospitals, clinics, and regional health authorities the ability to audit their own data in order to work towards quality improvement and patient safety. If

more published studies on appropriateness were available with a unified definition of appropriateness, it would be easier for study comparison.

Though we have an understanding of the proportion of CT referrals that are appropriate, there is no understanding as to why physicians refer in contradiction to the guidelines. Evidence from the free text in the referrals indicated that patients might have been requesting imaging from their physicians. While patient autonomy and empowerment are paramount to the effectiveness of the healthcare system, when a patient requests a test that is unnecessary and potentially harmful, it would be useful to understand the motivations behind this behaviour better. It is not known how often this type of request is asked of a physician, but collecting data on patient requested imaging would also be useful in this situation. We also need to better understand why physicians concede to their patients despite guidelines stating otherwise. This would allow for solutions to both of these problems to be created, and with patient engagement, deploy strategies that would allow for the reduction of unnecessary imaging, both in our local context of Newfoundland and Labrador and nationally.

6.5 Dissemination

The findings of this study have been presented and disseminated at multiple conferences. Abstracts have been submitted to the Choosing Wisely Canada National Meeting and the International Forum for Back and Neck Pain. Abstracts and materials for recent conferences can be found in the Appendices 4 & 5.

6.6 Conclusions

This thesis' objective was to investigate the appropriateness of radiological imaging ordered by family physicians and general practitioners for patients with low back pain.

Results from the systematic review show that globally, the proportion of appropriateness for imaging modalities that emit radiation (e.g., x-ray and CT) is 44% for x-rays and 54% for CT imaging. This suggests that there are many radiological images that are ordered for back pain unnecessarily, wasting the public healthcare system's resources and time, exposing patients to carcinogenic radiation. However, our ability to trust the results is somewhat limited due to the poor-quality studies that are available.

In our local RHA, there is a high rate of CT imaging occurring, especially when compared to other provinces in Canada. This also has the same negative connotations as previously discussed, with exposure to radiation being the most harmful, especially to patients. When focusing on one year's CT referral data, it is clear that many images are ordered because there are no clear guidelines for patients with radicular syndrome back pain. Only 5% of the CTs ordered in 2016 were for a suspected red flag condition. While it is likely that more of the CT images were appropriate, it is difficult to separate those images that were done for a potential surgical candidate from the images that were truly unnecessary due to the limited information on the referral forms. Patients and physicians require

solid recommendations in ordered to decrease these unnecessary images in order

to provide better patient care and to improve healthcare expenditure.

6.7 References

(1) Canadian Institute of Health Information. Medical Imaging in Canada 2012. 2013; Available at: https://www.cihi.ca/en/mit_summary_2012_en.pdf. Accessed January 9, 2019.

(2) Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151(4):264-269.

(3) Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011 Apr;64(4):383-394.

(4) Bardin LD, King P, Maher CG. Diagnostic triage for low back pain: a practical approach for primary care. Med J Aust 2017 Apr 3;206(6):268-273.

(5) Lurie JD, Birkmeyer NJ, Weinstein JN. Rates of advanced spinal imaging and spine surgery. Spine 2003;28(6):616-620.

(6) Australian Commission on Safety and Quality in Health Care. Computed tomography of the lumbar spine. 2015; Available at: https://www.safetyandquality.gov.au/wp-content/uploads/2015/11/SAQ201_03_Chapter2_v9_FILM_tagged_merged_2-3.pdf. Accessed May 1, 2019.

(7) Busse J, Alexander P, Abdul-Razzak A, Riva J, Alabousi M, Dufton J. Appropriateness of spinal imaging use in Canada. Hamilton, ON: McMaster University 2013.

(8) Jenkins HJ, Downie AS, Maher CG, Moloney NA, Magnussen JS, Hancock MJ. Imaging for low back pain: is clinical use consistent with guidelines? A systematic review and meta-analysis. Spine J 2018 May 3.

(9) Oikarinen H, Meriläinen S, Pääkkö E, Karttunen A, Nieminen MT, Tervonen O. Unjustified CT examinations in young patients. Eur Radiol 2009 2009/;19(5):1161-1165.

(10) Schlemmer E, Mitchiner JC, Brown M, Wasilevich E. Imaging during low back pain ED visits: A claims-based descriptive analysis. Am J Emerg Med 2015 2015/;33(3):414-418.

(11) Hourcade S, Treves R. Computed tomography in low back pain and sciatica. A retrospective study of 132 patients in the Haute-Vienne district of France. Joint Bone Spine 2002 Dec;69(6):589-596.

APPENDICES

Appendix 1. Systematic review and meta-analysis search strategy.

<u>Pubmed</u>

("Back Pain"[Mesh] OR "back pain"[tiab] OR backache[tiab] OR "back pains"[tiab] OR backaches[tiab] OR "back aches"[tiab] OR dorsalgia[tiab]) AND ("Guidelines as Topic"[Mesh] OR "Practice Guideline"[Publication Type] OR advice[tiab] OR treatment[tiab] OR options[tiab] OR policy[tiab] OR protocol[tiab] OR Guidelines[tiab] OR "decision tool"[tiab] OR "decision aid"[tiab] OR algorithm[tiab]) AND ("Guideline Adherence"[Mesh] OR "guideline adherence"[tiab] OR "policy compliance"[tiab] OR "protocol compliance"[tiab] OR "protocol adherence"[tiab] OR "Institutional adherence"[tiab] OR "Institutional compliance"[tiab] OR comply[tiab] OR compliant[tiab] OR conform[tiab] OR conformance[tiab] OR appropriateness[tiab] OR justif*[tiab])

<u>Embase</u>

(('backache'/exp OR 'backache'/de OR backache*:ti,ab OR 'back'/exp OR back) AND pain*:ti,ab OR 'back pain syndrome':ti,ab OR backpain*:ti,ab OR dorsalgia*:ti,ab OR 'pain, back':ti,ab) AND (('clinical practice guidelines':ti,ab OR guideline*:ti,ab OR 'guidelines as topic':ti,ab OR 'practice'/exp OR practice) AND guideline*:ti,ab OR 'practice guidelines as topic':ti,ab) AND (adherence*:ti,ab OR compliance*:ti,ab OR conform*:ti,ab OR justif*:ti,ab)

(MH "Back Pain+" OR TI "back pain*" OR AB "back pain*" OR TI backache* OR AB backache* OR TI "back ache*" OR AB "back ache*" OR TI dorsalgia OR AB dorsalgia) AND (MH "Practice Guidelines" OR PT "Practice Guideline" OR TI advice OR AB advice OR TI treatment OR AB treatment OR TI options OR AB options OR TI policy OR AB policy OR TI protocol OR AB protocol OR TI guidelines OR AB guidelines OR TI "decision tool" OR AB "decision tool" OR TI "decision aid" OR AB "decision aid" OR TI algorithm OR AB algorithm) AND (MH "Guideline Adherence" OR TI "guideline adherence" OR AB "guideline adherence" OR TI "policy compliance" OR AB "policy compliance" OR TI "protocol compliance" OR AB "protocol compliance" OR TI "protocol adherence" OR AB "protocol adherence" OR TI "institutional adherence" OR AB "institutional adherence" OR TI "institutional compliance" OR AB "institutional compliance" OR TI comply OR AB comply OR TI compliant OR AB compliant OR TI conform OR AB conform OR TI conformance OR AB conformance OR TI appropriateness OR AB appropriateness OR TI justif* OR AB justif*)

Appendix 2. List of excluded studies from chapter 3 with reasons for exclusion.

Study	Reason for Exclusion
Fullen 2007	Outcome reporting and patient population
Raja 2018	Aggregate number which included MRI
Charlesworth 2016	Aggregate number with MRI
Foo 2017	Aggregate number with MRI
Kost 2015	Aggregate number with MRI
Lin 2016	Aggregate number with MRI
Rao 2015	Aggregate number with MRI
Rego 2016	Aggregate number with MRI
Bishop 2003	Guideline year
Buller-Close 2003	Guideline year
Day 1995	Guideline year
Deyo 1986	Guideline year
Eccles 2001	Guideline year
Espeland 1999	Guideline year
Espeland 2001	Guideline year
Gonzalez-Urzelai 2003	Guideline year, data collection method
Halpin 1991	Guideline year
Hourcade 2002	Guideline year
Richards 2002	Guideline year
Schectman 2003	Guideline year
Schroth 1992	Guideline year
Suarez-Almazor 1997	Guideline year
Tacci 1999	Guideline year

Appendix 3. RECORD and STROBE Checklist Items for Included studies in descriptive synthesis.

Author	Item																																		
	1	1.1	1.2	1.3	2	3	4	5	6	6.1	6.2	6.3	7	7.1	8	9	10	11	12	12.1	12.2	12.3	13	13.1	14	15	16	17	18	19	19.1	20	21	22	22.1
Baez*	✓	NR	✓	NR	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Culleton*	✓	✓	✓	NR	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Muntion-Alfaro~	✓	✓	?	NR	✓	✓	✓	✓	✓	✓	NR	NR	?	?	?	✓	✓	?	?	?	✓	?	✓	?	?	✓	✓	NA	?	✓	?	✓	✓	NR	NR
Oikarinen	✓	✓	NR	NR	✓	✓	✓	✓	✓	NR	NR	NR	✓	✓	✓	✓	NR	NR	?	✓	?	NR	✓	NR	NR	✓	?	NA	✓	✓	✓	✓	✓	NR	NR
Schlemmer	✓	✓	✓	NR	✓	✓	✓	✓	✓	✓	NR	NR	✓	✓	✓	?	✓	✓	✓	NR	✓	NR	?	?	✓	✓	✓	NA	✓	✓	✓	✓	✓	✓	✓

Tahvonen	✓	✓	✓	NR	✓	✓	✓	✓	?	NR	NR	NR	✓	✓	?	?	NR	?	?	?	NR	NR	NR	NR	NR	✓	?	NA	?	NR	NR	✓	✓	NR	NR

* Abstract only
~ Abstract published in English, Full study published in Spanish
✓ Reported adequately
? unclear; missing information to make a clear decision, but partial information is reported
NR: not reported
NA: Not applicable

Appendix 4- One-page summary submitted to Choosing Wisely Canada

National Meeting in Montreal May 27th, 2019.

Appendix 5- First abstract submitted to International Forum for Back and Neck Pain in Quebec City, July 3-6.

Title:
Are Lumbar Spine CT referrals from General Practitioners Appropriate According to the Guidelines? A Medical Record Review in Newfoundland using Linked Databases

Abstract:
I. STUDY DESIGN
A retrospective medical record review.

II. OBJECTIVE
To evaluate the appropriateness of lumbar spine Computed Tomography imaging referrals from family physicians for patients with low back pain by categorizing the reason for referral according to evidence-based imaging guidelines.

III. SUMMARY OF BACKGROUND DATA
Choosing Wisely guidelines recommend ordering CT imaging for patients with low back pain if there are red flag indicators present. However, it is not known how many referrals are made to conform to a suspected red flag condition.

IV. METHODS, RESULTS, CONCLUSIONS
A medical record review of electronic health records was performed in the largest health region in Newfoundland and Labrador. All adult lumbar spine CT referrals included for analysis were ordered by a family doctor between January 1st, 2016 and December 31st, 2016. Each CT referral was collected from two linked databases (Meditech and PACs). Free-text data on the reason for referral were extracted, cleaned, and categorised into three groups: red flag indicated (appropriate), radicular syndrome (potentially appropriate), or non-specific LBP (inappropriate). 3623 lumbar spine CTs were included. The mean age of patients was 54.7 (SD 14 years) with 54.5% of referrals for female patients. 5.4% (95% CI 4.3%-5.7%) of lumbar spine CT referrals were for a suspected red flag condition and judged to be appropriate. 75.3% of LS referrals were for radicular syndromes, 16.7% for non-specific LBP, and 2.6% were missing/unable to code. This audit found very few CTs are being ordered appropriately in concordance with evidence-based guidelines regarding red flags. More research is required to understand why CTs are being ordered for radicular syndrome and non-specific LB P in order to reduce these potentially unnecessary referrals.

Appendix 6- Second abstract submitted to International Forum for Back and Neck Pain in Quebec City, July 3-6.

Title:
Appropriateness of CT and X-ray Imaging for Patients with Low Back Pain in Primary Care Settings: A Systematic Review and Meta-Analysis

Abstract:
I. STUDY DESIGN
Systematic review and meta-analysis.

II. OBJECTIVE
To determine the proportion of CT and x-ray images for low back pain that are appropriate.

III. SUMMARY OF BACKGROUND DATA
Evidence-based guidelines recommend diagnostic imaging for low back pain if physicians suspect red-flag spinal pathology. Otherwise, imaging should be avoided to decrease unnecessary testing and radiation exposure. Previous reviews provide a single pooled estimates of appropriateness for multiple imaging types ordered by multiple providers. However, no systematic review has provided an estimate of appropriateness for physician-ordered x-rays and CTs which is important for developing targeted behaviour-change interventions.

IV. METHODS, RESULTS, CONCLUSIONS
Pubmed, CINAHL, and Embase were searched for "low back pain", "guidelines", and "adherence". Independent screening, data extraction and study quality were conducted in accordance with PRISMA guidelines. Studies were included if they reported the proportion of appropriate CT or x-ray images ordered in family-practice or emergency-department settings after 2000. A random effects, single-proportion model meta-analysis was used and synthesized with GRADE. Six studies were included in the descriptive synthesis, and five studies were pooled for meta-analysis. Four studies reported x-rays appropriateness, one study reported CT appropriateness, and one study reported on both modalities. Risk of bias was high in 4 studies, moderate in one, and low in one. The pooled estimate of x-ray appropriateness (n=5010) was 44% (95% CI: 34%, 54%) and judged to be low-quality. The pooled estimate for CT appropriateness (n=678) was 54% (95% CI: 51%, 58%) and judged to be very-low quality. The quality was downgraded largely due to high risk of bias due to possible misclassification of population and outcome and imprecision of the estimates. Importantly, there were heterogeneity and overall lack of reporting on how appropriateness was defined. Explicit guidance on defining "appropriateness" is necessary to advance future work in this area.

126

Appendix 7- Health Research Ethics Board Approval Letter

Ethics Office
Suite 200, Eastern Trust
Building
95 Bonaventure Avenue
St. John's, NL
A1B 2X5

June 01, 2017

Faculty of Medicine
Disciple of Epidemiology

Dear Miss Logan:

Researcher Portal File # 20172108
Reference # 2017.079

RE: "Investigating the Appropriateness of and Influencing Factors for Radiological Tests Orders for Low Back Pain"

Your application received a delegated review by a sub-committee of the Health Research Ethics Board (HREB). *Full approval* of this research study is granted for one year effective **June 1, 2017**.

This is your ethics approval only. Organizational approval may also be required. It is your responsibility to seek the necessary organizational approval from the Regional Health Authority (RHA) or other organization as appropriate. You can refer to the HREA website for further guidance on organizational approvals.

This is to confirm that the HREB reviewed and approved or acknowledged the following documents (as indicated):

- Application, approved
- List of variables, approved
- Letter of request, approved

MARK THE DATE

This approval will lapse on June 1, 2018 . It is your responsibility to ensure that the Ethics Renewal form is submitted prior to the renewal date; you may not receive a reminder. The Ethics Renewal form can be found on the Researcher Portal as an Event form.

If you do not return the completed Ethics Renewal form prior to date of renewal:

- *You will no longer have ethics approval*

- *You will be required to stop research activity immediately*
- *You may not be permitted to restart the study until you reapply for and receive approval to undertake the study again*
- *Lapse in ethics approval may result in interruption or termination of funding*

You are solely responsible for providing a copy of this letter, along with your approved HREB application form; **to Research Grant and Contract Services** should your research depend on funding administered through that office.

Modifications of the protocol/consent are not permitted without prior approval from the HREB. **Implementing changes without HREB approval may result in your ethics approval being revoked, meaning your research must stop**. Request for modification to the protocol/consent must be outlined on an amendment form (available on the Researcher Portal website as an Event form) and submitted to the HREB for review.

The HREB operates according to the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2), the Health Research Ethics Authority Act (HREA Act) and applicable laws and regulations.

You are responsible for the ethical conduct of this research, notwithstanding the approval of the HREB.

We wish you every success with your study.

Sincerely,

Patricia Grainger

Ms. Patricia Grainger (Chair, Non-Clinical Trials Health Research Ethics Board)
Dr. Joy Maddigan (Vice-Chair, Non-Clinical Trials Health Research Ethics Board)

CC: Dr. Patrick S. Parfrey