Complications of Bone Morphogenetic Protein-2 Use in Open Transforaminal Lumbar Interbody Fusion

Allison Miller
Yale Physician Associate Program, allison.miller.amm359@yale.edu

Follow this and additional works at: https://elischolar.library.yale.edu/ysmpa_theses

Part of the Medicine and Health Sciences Commons

Recommended Citation
Miller, Allison, "Complications of Bone Morphogenetic Protein-2 Use in Open Transforaminal Lumbar Interbody Fusion" (2021). Yale School of Medicine Physician Associate Program Theses. 85. https://elischolar.library.yale.edu/ysmpa_theses/85

This Open Access Thesis is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale School of Medicine Physician Associate Program Theses by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
COMPLICATIONS OF BONE MORPHOGENETIC PROTEIN-2 USE IN OPEN TRANSFORAMINAL LUMBAR INTERBODY FUSION

A Thesis Presented to
The Faculty of the School of Medicine
Yale University

In Candidacy for the degree of
Master of Medical Science

May 2021

Allison Miller, PA-SII
Class of 2021
Yale Physician Associate Program

Arya Varthi, MD
Assistant Professor
Department of Orthopaedic Surgery
Yale School of Medicine
TABLE OF CONTENTS

LIST OF TABLES AND FIGURES........................................................................iv

ABSTRACT........................................................................................................v

CHAPTER 1: INTRODUCTION...........................................................................1

1.1 Background................................................................................................1
   A. Degenerative Disc Disease of the Lumbar Spine..........................................1
      i. Description and Diagnosis ......................................................................1
      ii. Prevalence .............................................................................................2
      iii. Causes and Risk Factors .......................................................................3
      iv. Treatment ............................................................................................4
   B. Transforaminal Lumbar Interbody Fusion .....................................................6
      i. Description ...........................................................................................6
      ii. Outcomes ............................................................................................7
   C. Recombinant Human Bone Morphogenetic Protein-2 ..................................8
      i. Description and Advantages ...................................................................8
      ii. Mechanism of Action ...........................................................................9
      iii. Indications .........................................................................................9
      iv. Off-Label Use in TLIF .........................................................................10
   D. Selected Potential Complications of rhBMP-2 Use in TLIF .........................10
      i. Introduction ........................................................................................10
      ii. Radiculopathy ....................................................................................11
      iii. Heterotopic Ossification ....................................................................12
      iv. Osteolysis ..........................................................................................13

1.2 Statement of the Problem..........................................................................13
1.3 Goals and Objectives..............................................................................14
1.4 Hypothesis ................................................................................................15
1.5 References...............................................................................................15

CHAPTER 2: REVIEW OF THE LITERATURE...............................................23

2.1 Introduction...............................................................................................23
2.2 Review of Empiric Studies .......................................................................23
   A. Radiculopathy .......................................................................................23
   B. Heterotopic Ossification .........................................................................27
   C. Osteolysis .............................................................................................30
2.3 Review of Confounding Variables ............................................................32
   A. Patient Variables ...................................................................................32
   B. RhBMP-2 Dose .....................................................................................33
   C. RhBMP-2 Placement .............................................................................34
   D. Choice of Interbody Device ..................................................................34
   E. Length of Follow-Up .............................................................................35
LIST OF TABLES AND FIGURES

Table 1: Proposed Timeline........................................................................................................47

Figure 1: Inclusion and Exclusion Criteria...............................................................................48

Table 2: Heterotopic Ossification Classification......................................................................51
ABSTRACT

Transforaminal Lumbar Interbody Fusion is commonly performed with the use of off-label recombinant human bone morphogenetic protein-2. Despite its widespread use, bone morphogenetic protein-2 has been linked with postoperative complications in observational studies. However, no experimental study has supported this association nor has the relative risk as compared to iliac crest bone graft been quantified. In this study, we will compare the incidence proportion of radiculopathy, heterotopic ossification, and vertebral endplate osteolysis following Transforaminal Lumbar Interbody Fusion in patients who receive recombinant human bone morphogenetic protein-2 versus patients who receive iliac crest bone graft. Specifically, we will carry out a single-blinded, randomized controlled trial to assess incidence through six months of follow-up. The insights gathered will guide patient and surgeon decisions as they weigh the risks of different graft materials.
CHAPTER 1: INTRODUCTION

1.1 Background

A. Degenerative Disc Disease of the Lumbar Spine

i. Description and Diagnosis

Degenerative disc disease (DDD) affects millions of adults today and is a common cause of life-altering disability.\textsuperscript{1,2} As described by Battié, there is variability in how DDD is defined across the literature; however, we define the diagnosis as the combination of radiographic evidence of degenerative changes in the spine and the presence of corresponding symptoms as the term is used in the surgical community.\textsuperscript{3,4} DDD can occur at a single level or multiple levels of the spine, most commonly occurs in the cervical and lumbar regions, and seldom occurs in the thoracic spine. This work will focus on DDD of only the lumbar spine.

The most common degenerative changes supporting a diagnosis of DDD are spondylosis, spondylololisthesis, spondylolysis, spinal stenosis, and disc herniation.\textsuperscript{5} All are diagnosed radiographically. Patients may have more than one of these changes at a single vertebral level or distinct changes at different levels. One pathology may provoke another; for example, the fracturing of the pars interarticularis that defines spondylolysis can result in the anterior slippage of one vertebrae relative to another characterized as spondylolisthesis.\textsuperscript{6} Spondylosis or degenerative arthritis describes a broad array of degenerative changes that can appear as a loss of disc height, facet joint arthropathy, osteophyte formation, or hypertrophy of the ligamentum flavum.\textsuperscript{7,8} Next, disc herniation occurs when part of the nucleus pulposus seeps through a tear in the annulus. Finally, if degenerative changes narrow the vertebral canal or neural foramina, spinal stenosis results.

Symptoms secondary to these anatomic changes manifest on a case-by-case basis and according to the underlying pathology. The most common symptoms include low back pain, leg
pain, and neurogenic claudication. Weakness, tingling, numbness, and sensory loss may all be manifestations of neurogenic claudication. Symptoms may be intermittent or consistent, bilateral or unilateral, or exacerbated by positional changes and physical loads. Also, patients may have symptoms and degenerative changes that are unrelated, so clinicians must be discerning in making the diagnosis. A careful history, physical exam, and judicious use of imaging and other diagnostics such as discography or nerve conduction studies are used to aid clinical reasoning. Other potentially serious causes of low back pain that must be ruled out include infection, neoplasm, fracture, cauda equina syndrome, referred pain, and rheumatologic disease. Ultimately the patient’s symptoms must be clinically judged to be related to their radiographic changes to be treated as DDD.

**ii. Prevalence**

DDD is commonly implicated as a cause of chronic low back pain, a leading cause of disability, healthcare utilization, and medical costs. The 2010 Global Burden of Disease study ranked low back pain as the number one cause of years lost to disability with 83 million disability-adjusted life years lost. Some studies suggest that chronic low back pain persisting beyond three months affects 15-45% of the U.S. adult population during their lifetimes and 8% of all U.S. adults at any given time are currently experiencing chronic low back pain of more than six months duration. Back pain was the principal reason for 13 million office visits in 2010. Mild to moderate cases of DDD frequently remain undiagnosed if the patient’s symptoms are managed successfully in which case imaging and surgical consideration are not recommended.

While the majority of cases of DDD are managed conservatively, severe DDD qualifying for surgical candidacy is a massive problem on its own. Ravindra et al. found that in 2015, 266
million individuals worldwide or 3.63% of the population had both degenerative spine disease and low back pain described to be neurosurgically relevant by the authors.\textsuperscript{1} Higher income countries were associated a higher density of cases; in the U.S. and Canada, the prevalence was 5\% or 16 million affected persons. This is likely an underestimate of the true number since the study did not include other common symptoms of DDD such as neurogenic claudication.

Many individuals with severe DDD do go on to receive surgery. DDD is the most common indication for both decompressive and fusion surgery, and these surgeries are performed widely.\textsuperscript{23} Using a nationally representative database, Martin et al. calculated that approximately 174,000 elective lumbar fusions were performed in the U.S. for a diagnosis of spondylolisthesis, spondylolysis, spondylosis, or disc herniation in 2015.\textsuperscript{24} In recent years, the number of lumbar fusions performed per year has more than doubled every decade, increasing by 220\% between 1990 and 2001 and by 118\% from 1998 to 2014.\textsuperscript{25-27} Decompressive surgeries remain common but appear to be decreasing, probably in favor of fusion.\textsuperscript{28} In 2011, around 525,000 laminectomies and disectomies were performed inpatient in the U.S. (although this number includes cervical cases), and many more were performed in an outpatient setting.\textsuperscript{29-31}

\textit{iii. Causes and Risk Factors}

There is evidence that normal aging of the spine causes degenerative changes.\textsuperscript{32,33} Aging is associated with cellular-level differences including decreased amounts of collagen and other proteoglycans even in healthy adults who never experience symptoms.\textsuperscript{34,35} These changes, along with minor injuries from daily stressors, cause a gradual weakening of the annulus fibrosus, a loss of disk height, and alteration of load distribution on the vertebral facets. This can lead to pathologic adaptations: facet joints may shift, osteophytes may form to limit excess motion, and annular tears may leak nucleus pulposus.\textsuperscript{36} In most adults, these adaptations are asymptomatic
and, if identified, incidental. However, in other cases they may cause compression of nerve roots (either directly or indirectly from increased intrathecal pressure secondary to spinal canal narrowing), release of inflammatory signals by injured structures, and grinding of arthritic facets.

In addition to old age, other risk factors for developing lumbar DDD are numerous. Those identified at this time include female gender, genetic factors, lower socioeconomic status, smoking, obesity, diabetes, arthritis, vibration (such as that imposed by riding in a motor vehicle), trauma, physically strenuous work, and sedentary lifestyle. Preventing DDD involves minimizing modifiable risk factors. A healthy lifestyle including engaging in regular activity, consuming a balanced diet, maintaining a healthy weight, and choosing not to smoke is recommended to reduce one’s risk. Utilizing proper body mechanics when lifting objects is also important.

iv. Treatment

The primary treatment for all patients with DDD is conservative. Nonsurgical treatment options shown to be helpful for varying durations include physical therapy, exercise-based interventions, CBT, mind-body techniques, and passive therapies like chiropractic care and massage. NSAIDs and acetaminophen are first-line pharmacologic treatments, and opioid use is not recommended in almost all cases. All patients should trial conservative management for a minimum of 12 weeks with 6 to 12 months considered a more appropriate duration in anything but the most severe cases. Patients experiencing symptoms lasting beyond 12 weeks rarely achieve a complete resolution of symptoms and should be counseled that the goals of treatment are to reduce disability and pain.

Patients with symptoms refractory to conservative management may consider surgical treatment. Laminectomy, foraminotomy, discectomy, nucleoplasty, facetectomy, arthroplasty,
spinal cord stimulator placement, and spinal fusion are all important options in the surgeon’s arsenal for the treatment of DDD.\textsuperscript{47} Spinal fusion techniques include posteriorlateral fusion (PLF) and lumbar interbody fusion (LIF). PLF fuses adjacent vertebrae by means of pedicle screws and overlying osteoconductive material which grows to connect the transverse processes.\textsuperscript{48} The LIF technique involves removing the intervertebral disc and then inserting a device filled with osteoconductive material in the empty space. If the fusion is successful, new bone will form between the vertebral bodies, joining them firmly together. The intraoperative goals of LIF are to secure instabilities, restore lordosis, correct deformity, and decompress neural elements all toward the ultimate goal of reducing the patient’s symptoms of DDD.\textsuperscript{49}

LIF is performed from a variety of different approaches that are named according to the relation of the operative site to the transverse process. The approaches relevant to this work include the Anterior (ALIF), Oblique (OLIF), Transforaminal (TLIF), and Posterior (PLIF) techniques. ALIF and OLIF involve a supine or laterally positioned patient and generally pose a risk of damage to intra-abdominal and/or retroperitoneal structures.\textsuperscript{50} Patients are positioned prone for PLIF and TLIF and approached by a midline or bilateral paramedian incision; damaged or at-risk structures include the paraspinal muscles, spinal nerve roots, dura, and ligamenum flavum.\textsuperscript{51} PLIF and TLIF can each be performed alone but are usually combined with the PLF technique for added stability. Each of these approaches has distinct advantages and disadvantages that are considered in the context of the patient and surgeon. In general, no approach is considered clearly better than another. Additionally, it is worth noting LIF is sometimes performed by means of minimally-invasive variants; however, some studies suggest advantages and disadvantages specific to these techniques and due this potential risk of confounding, this work will only consider open surgeries.\textsuperscript{52}
Apart from cases of progressive and severe neurologic deficits, surgical treatment is elective and the decision to operate should be made on a case-by-case basis by the patient and surgeon. Evidence exists that surgical treatment can reduce symptoms and improve functioning in the right patient.\textsuperscript{53-62} The American Pain Society and American Association of Neurological Surgeons (AANS) recommend fusion as an option for patients with DDD.\textsuperscript{63} Specifically, AANS recommends lumbar fusion for intractable low back pain due to spondylosis, disc herniation associated with spinal instability, or other severe degenerative changes and also for stenosis due to spondylolisthesis, deformity, or instability.\textsuperscript{22,64-66} The guidelines do not recommend one lumbar fusion technique over another; this is left to the surgeon’s judgement.

B. Transforaminal Lumbar Interbody Fusion

\textit{i. Description}

TLIF is a common method to achieve lumbar fusion. One study representing 25 million patients over a six year period ending in 2009 found that posterior interbody approaches (PLIF and TLIF) comprised 55\% of all fusions performed and were significantly more common in the elderly.\textsuperscript{27} Another report found that the annual number of posterior interbody fusions increased by 23\% between 2011 and 2014.\textsuperscript{67} Posterior approaches are popular because they are well known and comfortable to most spine surgeons and allow optimal visualization of nerve roots and direct neural decompression.\textsuperscript{68} TLIF is essentially a newer, revised version of PLIF that approaches the intervertebral space from one side rather than two, limiting paraspinal injury, neural retraction, and facet disruption.\textsuperscript{69} Therefore, TLIF is the preferred posterior technique for many surgeons.

TLIF is suitable for fusions between L1-S1. As mentioned, the patient is positioned prone and a midline or bilateral paramedian incision is made.\textsuperscript{51,70} After the incision, the paraspinal muscles are dissected to reveal the lateral margins of the vertebral facets. Then, a unilateral
laminectomy and facetectomy are performed to access the spinal canal. The surgeon now prepares to remove the intervertebral disc. Following discectomy, the vertebral endplates are prepared for the incoming implant which may be a cage, spacer, or structural graft. Osteogenic material is placed within the implant and also distributed around it to cover the vertebral endplates. Usually, bilateral pedicle screw fixation with posterolateral fusion is also performed to provide further stability and a posterior tension band. Radiographic confirmation of proper hardware placement and biomechanical restoration is obtained before closure.

The gold standard of osteogenic material used in fusion is autogenous iliac crest bone graft (ICBG). It is harvested as struts capable of bearing weight or as morselized bone from the patient minutes before fusion. Other options include local autogenous graft, which is repurposed bone fragments from the surgical site, and allograft. Local autograft and allograft are limited by availability, and allograft also carries a theoretical risk of provoking an immune reaction.

Finally, bone graft extenders and substitutes may augment these materials or replace them entirely; common ones include demineralized bone matrix, ceramics, and bone morphogenetic proteins (BMPs).

**ii. Outcomes**

Research supports TLIF as a safe and effective treatment. An early study of 55 patients with DDD found significant reductions in pain and disability as measured by the Visual Analog Scale (VAS) and Oswestry Disability Index (ODI), respectively, through an average of 46 months of follow-up. Operation time, complication rate, and estimated blood loss were comparable to other fusion techniques, and fusion rate was 89% with use of ICBG. Other studies have supported these findings: one meta-analysis of 19 studies involving TLIF confirmed significant reductions in back VAS, leg VAS, and ODI scores and another meta-analysis found a
fusion rate of 89.6% in 371 patients undergoing TLIF.\textsuperscript{74,75} The present-day fusion rate may be even higher; a recent meta-analysis of 21 articles including 432 patients who underwent a single-level TLIF at L5/S1 found a fusion rate of 99.3% using autograft. However, most of these studies did not use computed tomography which is more sensitive for non-union than radiograph.\textsuperscript{76}

C. Recombinant Human Bone Morphogenetic Protein-2

\textit{i. Description and Advantages}

Recombinant human bone morphogenetic protein-2 (rhBMP-2) is a popular substitute for ICBG in LIF, including TLIF, due to its association with several advantages. Perhaps the biggest advantage of using rhBMP-2 is the elimination of adverse events associated with harvesting ICBG. These events are significant and numerous; pain, hematoma formation, paresthesias, or infection occur at least 50\% of cases.\textsuperscript{77} One systematic review found that 60\% of patients experience donor site pain, and another study concluded that between 18\%-31\% of patients still experienced pain at 24 months postoperatively.\textsuperscript{78,79} Between 2\% and 5\% require re-operation due to wound complications. Other less common but important complications include iliac wing fracture and nerve and vascular injuries.

Next, rhBMP-2 likely increases fusion rates. One meta-analysis of ten RCTs involving ALIF or PLIF found that rhBMP-2 significantly decreased the risk of fusion failure at 24 months (RR=0.31) and decreased the rate of re-operation (RR=0.52) in comparison to ICBG.\textsuperscript{75} Another study independently re-analyzed the initial industry-sponsored data used to apply for FDA approval of rhBMP-2 and found a 12\% increased fusion rate versus ICBG in ALIF.\textsuperscript{80}

Finally, there are many patient populations in which rhBMP-2 may be particularly useful. The North American Spine Society recommends considering rhBMP-2 for use in patients at high risk for non-union or patients with poor-quality autogenous bone.\textsuperscript{81} In practice, these high risk
patients may comprise many if not most patients; those considered at high risk include patients with very common conditions such as hypertension, diabetes, osteoporosis, and tobacco use. RhBMP-2 is also useful in revision cases, multi-level fusions (to reduce extensive ICBG harvesting) and patients with rheumatoid arthritis or prior radiation therapy.\(^82\)

**ii. Mechanism of Action**

Bone morphogenetic protein-2 (BMP-2) is a dimeric protein molecule endogenously produced by mesenchymal stem cells and osteoblasts in the bone matrix.\(^83,84\) It is a powerful osteoinductive agent. It functions as a ligand, interacting with receptors on cell membranes and activating a signaling cascade which eventually regulates gene transcription.\(^85\) This cascade ultimately promotes the differentiation of mesenchymal cells into osteoblasts, allowing osteogenesis and bone formation. These processes occur throughout life.

RhBMP-2 is molecularly identical to BMP-2 but artificially produced from a hamster ovary cell line. The desired clinical effect of rhBMP-2, enhancement of solid bone fusion, is achieved through the same mechanism as endogenous BMP-2.\(^86\) Introducing exogenous rhBMP-2 to the body exponentially increases the natural physiologic concentration, powerfully stimulating osteoblast maturation and bone formation.

**iii. Indications**

RhBMP-2 has been approved for use by the FDA since 2002. The patent was given to Medtronic (Memphis, TN) which markets the product as Infuse\(^{TM}\) Bone Graft.\(^87\) Initial studies reported a benefit of an increased radiographic fusion rate with rhBMP-2 versus the ICBG control in single level ALIF between L4-S1.\(^77\) Currently, rhBMP-2 is FDA approved for use in spinal fusion at a single level between L2-S1. Use is only permitted in combination with one of the manufacturer’s intended interbody devices and the absorbable collagen sponge (ACS) carrier.
in which rhBMP-2 is dissolved. Approved surgical approaches include OLIF and open, retroperitoneal, or laparoscopic ALIF depending on the interbody device. The OLIF approach carries additional limitations on the level at which it may be used, either L5-S1 or L2-L5, again depending on the device. RhBMP-2 is not approved for use in TLIF.

iv. Off-Label Use in TLIF

Although it remains off-label, rhBMP-2 use in TLIF is very common. Between 40-50% of TLIFs in the U.S. were performed with rhBMP-2 in 2011.\textsuperscript{88} In fact, the most common use of Infuse\textsuperscript{TM}, a multi-billion dollar industry in the U.S. alone, is for off-label TLIF or PLIF, comprising 30% of all use of the product in 2009.\textsuperscript{89} Interestingly, the evidence so far has not shown increased fusion rates with use in TLIF. One recent retrospective cohort study found a 92.7% fusion rate with rhBMP-2, nearly identical to the control group of ICBG or local autograft.\textsuperscript{90} Another meta-analysis involving 17 studies including TLIF found rhBMP-2 minimally increased fusion with a rate of 95.0% (CI 92.8%-96.5%) in the intervention versus 93.0% (CI 78.1-98.0%) in the autograft control.\textsuperscript{91} Despite the lack of a clear fusion benefit, these studies do show that rhBMP-2 is effective and therefore can still retain the advantage of avoiding ICBG morbidity.

D. Selected Potential Complications of rhBMP-2 use in TLIF

i. Introduction

Since its approval, rhBMP-2 has often been the subject of controversy regarding its safety. In 2008, the FDA released a public health notification warning about potential life-threatening airway compromise due to rhBMP-2 use in the cervical spine, specifically from local inflammation or fluid collections caused by the product.\textsuperscript{92} In its on-label use, rhBMP-2 was recognized to pose a greater risk of retrograde ejaculation than was previously described.\textsuperscript{93} Many
suspected bias and lost confidence in the safety conclusions of the initial trials after it was made known that the authors of the studies received lucrative payouts from Medtronic, while others emphasized that the initial FDA statement noted some evidence for a variety of complications that was largely ignored.\textsuperscript{86} Finally, another rhBMP-2 product of a higher dose was submitted to the FDA for approval by Medtronic but rejected due to concerns regarding an association with malignancy.

Safety concerns for use of rhBMP-2 in TLIF have also been raised. Since its widespread adoption, reports of radiculopathy, heterotopic ossification (HO), and endplate osteolysis attributed to rhBMP-2 use in TLIF have surfaced.\textsuperscript{94,95} These complications are of concern to patients and surgeons as they can cause recurrent or new onset back pain or leg pain, hardware and fusion failure, or serious spinal cord compression and cases may require re-operation.\textsuperscript{96} There is anatomic rationale for why these complications may be more frequent in a TLIF approach than as reported in the initial trials using ALIF. In ALIF, the posterior annulus and posterior longitudinal ligament are left intact and shield spinal nerves and posterior structures from rhBMP-2, but in TLIF these structures are dissected. Yet, the risk of radiculopathy, HO, and endplate osteolysis in TLIF have not yet been quantified by any experimental study. Many have described the need for high-quality studies to promote safe rhBMP-2 use.\textsuperscript{97}

\textit{ii. Radiculopathy}

Radiculopathy following TLIF manifests as persistent or recurrent pain, numbness, weakness, or tingling in a dermatomal distribution and originates from compressive or non-compressive causes.\textsuperscript{12} Compression from heterotopic bone growth, fluid collection including hematoma and seroma, scar tissue, or recurrent stenosis can displace and irritate nerve roots, causing symptoms. Non-compressive causes of radiculopathy chiefly include radiculitis and
iatrogenic injury. Radiculitis refers to worsening pain in a dermatomal distribution ipsilateral to the surgical side that is secondary to an inflammatory process.\textsuperscript{98} A diagnosis of radiculopathy is made after incomplete decompression, infection, traction injury, implant migration, and chronic neuropathy have been ruled out. MRI and/or CT are used to identify or rule out structural pathologies.

Laboratory in-vivo studies have successfully demonstrated a plausible mechanism for rhBMP-2 induced radiculitis. Two separate studies found that BMP-2 receptors are present on both central and peripheral nervous system cells, and rhBMP-2 is able to trigger these receptors directly.\textsuperscript{99,100} Other studies have found that the rhBMP-2 mechanism is associated with a significant inflammatory response involving cytokine release a million times greater than that triggered by physiologic BMP-2 concentrations.\textsuperscript{101} One experiment performing PLF using rhBMP-2 in a rat model found increased macrophage infiltration in and around the dorsal root ganglia on histology versus saline control.\textsuperscript{102} The rats also displayed increased mechanical allodynia between three and ten days postoperatively. Finally, a case study of a man with radiculopathy following TLIF with rhBMP-2 found a fibrovascular stroma with extensive lymphocytes and eosinophils typical of an inflammatory healing response on histologic examination of the affected nerve root.\textsuperscript{101}

\textit{iii. Heterotopic Ossification}

HO can occur if rhBMP-2 leaks from the interbody space into the spinal canal or neural foramina and forms undesirable new bone outside of the target area.\textsuperscript{103} As rhBMP-2 is capable of inducing bone formation on its own, the mechanism is assumed to be identical to that of rhBMP-2 in the intervertebral space. One published case study reported ectopic chondrocyte metaplasia among a fibrovascular stroma and immature bone on histology of a HO lesion, consistent with
the known mechanism of BMP-2. Ectopic ossification is sometimes used to describe the same process but to a more severe extent; however, this work will not make this distinction. A diagnosis of HO is made by imaging, either CT or MRI, and the patient may present with pain or radicular complaints.

iv. Osteolysis

Endplate osteolysis, or bony resorption, results from excessive osteoclastic activity. Osteolysis is part of a normal remodeling process in progression to fusion, but excessive resorption can destabilize the implant and cause subsidence, loosening, or migration. Endplate osteolysis is also associated with increased postoperative pain. Perhaps paradoxically, rhBMP-2 can induce unnecessary bone breakdown. In vitro and in vivo studies have shown that rhBMP-2 affects cells of the osteoclastic lineage by interfering with the balance of bone breakdown and bone formation, acting directly on the osteoclasts and their precursors, or both. Studies have noted the osteolytic phase induced by rhBMP-2 appears to be transient and osteolytic activity stops after 4-12 weeks; however, it may be consistent with decreased mineral densities seen on postoperative CT months later. Diagnosis is made with radiograph or CT.

Finally, it should be noted that radiculopathy, heterotopic ossification, and osteolysis are often related; for example, heterotopic bone growth usually becomes symptomatic by impinging on surrounding nerves. Or, excessive osteolysis allowing cage migration may induce heterotopic ossification in the spinal canal or neural foramina.

1.2 Statement of the Problem

Lumbar DDD is highly prevalent in the general population and its symptoms are one of the most common causes of disability in U.S. adults. Many studies have shown that patients are choosing lumbar fusion with increasing frequency every year, and TLIF is a popular choice
due to its association with several advantages and high fusion rates.\textsuperscript{27} About half of TLIF surgeries are performed with off-label rhBMP-2 which allows patients to bypass the morbidity associated with ICBG and may aid fusion especially in populations at risk for nonunion.\textsuperscript{89} Despite this widespread use, a number of recent observational studies describe increased rates of radiculopathy, heterotopic ossification, and endplate osteolysis in TLIF procedures performed in conjunction with off-label rhBMP-2, raising concerns about its safety.\textsuperscript{97} However, no experimental study has described the incidence of these complications and the true risk remains unclear. As hundreds of thousands of patients are likely to undergo TLIF in the coming decades, accurate safety information is needed to reliably guide patients and surgeons as they weigh the risks and benefits of different osteo
genic materials.

1.3 Goals and Objectives

The goal of this study is to evaluate the safety of rhBMP-2 use in TLIF using a high-quality, experimental design. Therefore, we propose a single-blinded, randomized controlled trial to quantify the incidence of radiculopathy, heterotopic ossification, and endplate osteolysis in patients who undergo TLIF with off-label rhBMP-2 in comparison to TLIF performed with gold standard ICBG. Our study will involve multiple surgical centers in the United States and enroll hundreds of patients. Participants will be randomized to receive either rhBMP-2 or ICBG and we will measure our primary outcomes through six months of follow up. Radiculopathy will be determined by the same medical team that performed the surgery as a clinical diagnosis of exclusion, and the presence of heterotopic ossification or osteolysis will be assessed by independent orthopedic surgeons who will review each patient’s CT imaging. We will also quantify a variety of intraoperative, perioperative, and secondary postoperative complications. Overall, this study will provide a thorough assessment of the relative risk of developing postoperative complications from electing to use rhBMP-2 over ICBG in TLIF.
1.4 Hypothesis

A. Among adults undergoing TLIF for DDD, patients who receive rhBMP-2 during surgery will not have a statistically significant difference in postoperative incidence of radiculopathy than patients receiving ICBG.

B. Among adults undergoing TLIF for DDD, patients who receive rhBMP-2 during surgery will not have a statistically significant difference in postoperative incidence of heterotopic ossification than patients receiving ICBG.

C. Among adults undergoing TLIF for DDD, patients who receive rhBMP-2 during surgery will not have a statistically significant increase in postoperative incidence of endplate osteolysis than patients receiving ICBG.

1.5 References


106. Toth JM, Boden SD, Burkus JK, Badura JM, Peckham SM, McKay WF. Short-term osteoclastic activity induced by locally high concentrations of recombinant human bone

CHAPTER 2: REVIEW OF THE LITERATURE

2.1 Introduction

We conducted a thorough review of the literature between December 2020-May 2021 using PubMed, Scopus, Embase, Web of Science, and Ovid. Only articles written in English were considered. We reviewed the titles, abstracts, and methods sections of each identified study to determine relevancy. Key terms employed to find articles relating to the study population and intervention included: lumbar fusion, posterior lumbar fusion, interbody fusion, transforaminal lumbar interbody fusion, TLIF, BMP, rhBMP-2, rhBMP, bone morphogenetic protein, recombinant human bone morphogenetic protein, Infuse™. The following terms were used to identify outcomes: radiculopathy, radiculitis, leg pain, back pain, radicular, heterotopic ossification, ectopic ossification, ectopic bone, osteolysis, osteolytic, complication, outcome. Confounding variables were identified by adding these key words: dose, demographic, narcotic, opioid, osteoporosis, cage, device, PEEK, titanium. Finally, additional terms to find studies with relevant methodology include: iliac crest bone graft, versus, randomized controlled trial, methods, design.

2.2 Review of Empiric Studies

A. Radiculopathy

Around 2009, reports of radiculopathy in TLIF began to surface. That year, Rihn et al. published a retrospective study of 48 patients undergoing single-level TLIF with rhBMP-2, reporting that 8 patients (16.7%) sustained new onset postoperative radiculopathy.¹ Patients were U.S. adults between the ages of 18 and 80 with a degenerative indication for surgery. All patients developing radiculopathy were imaged with CT. In our judgement, it appears two of these eight cases appear less likely to be related to rhBMP-2; one case was attributed by the authors to a
malpositioned screw and the other to a deep infection. However, in both cases radicular symptoms did not improve with appropriate treatment for these likely causes so an rhBMP-2 effect cannot be entirely excluded. The remaining six cases are more easily associated with rhBMP-2 use: a structural cause was not identified in five cases and the final case was deemed secondary to ectopic bone growth.

Later the same year, Rihn et al. expanded on this report by publishing another study that included a larger sample size and an ICBG control group. The study retrospectively examined 119 adult patients undergoing single-level TLIF for a degenerative indication. A similar incidence was found: 12 of 86 (14%) patients in the rhBMP-2 group developed radiculopathy. However, it appears that this study’s sample only expanded the sample described in the previous study, as for both studies the authors state their population was the same single institution during the same time frame. Again, two cases of radiculopathy were attributed to a malpositioned screw and a deep infection, respectively, but the rest could be consistent with rhBMP-2 use. In the ICBG control, 1 of 33 patients (3%) developed radiculopathy and no identifiable cause was found. This clinical trend lacked statistical significance (p=0.08).

Many recent publications have supported these findings with incidence rates near the low teens. Lubelski et al. reported new onset radiculopathy in 12.5% of 216 patients undergoing either TLIF or PLIF. All of the cases were deemed idiopathic; none were attributed to compression. The advantages of this study include a fairly large sample size and use of multiple surgeons, increasing generalizability. However, the descriptive design did not include a control group, limiting comparisons with other populations including ICBG. Also, in contrast to Rihn et al, degenerative indications were not the only indication for surgery in this sample and included
postlaminectomy syndrome, pseudoarthrosis, and osteomyelitis. The percentage of patients with these indications is unclear, but the authors note that they comprise a minority.

Next, Villavicencio and Burneikiene retrospectively described new onset radiculitis in 23 of 203 patients (11.3%) who underwent TLIF with rhBMP-2 at either at one or two levels. This number may underestimate the incidence of radiculopathy as we define it; we include both structural and inflammatory causes, but the authors of this study excluded all structural causes in their definition. Each subject experiencing radicular symptoms received an MRI to rule out structural causes, but it is not stated how many patients, if any, were found to have structural findings. However, we note that in each of the previously reviewed studies, the large majority of cases were consistent with an inflammatory etiology rather than a compressive one and this work’s incidence rate is consistent with these prior studies in a sizable sample.

Finally, a retrospective cohort study in 2018 authored by Khan et al. found a significantly greater risk of developing radiculopathy in rhBMP-2 treated patients undergoing TLIF than the autograft control (OR 4.53, CI 1.42-14.5). The incidence of radiculopathy was 9 of 83 (10.8%) in the rhBMP-2 group and 2 of 104 (1.9%) in the autograft control. All patients were U.S. adults undergoing TLIF for a degenerative condition with two years of radiographic and clinical follow-up. Each case of radiculopathy was attributed to an inflammatory cause, but the authors note that CT studies were not routinely attained to assess for structural causes. Overall, this study compares favorably to the incidence rates of the above studies in a sizable sample and also concludes a statistically significant higher incidence than that seen in autograft. Taken together, these studies suggest an increased risk of radiculopathy in TLIF performed with rhBMP-2 compared to autograft, with an incidence between 10.8-14.0%.
It must be noted that some studies have not found similar rates of radiculopathy with rhBMP-2 use in TLIF. In 2011, Owens et al. found an overall incidence of radiculopathy in 13 of 204 patients (6.4%) undergoing TLIF with rhBMP-2. Two of these cases appear unrelated to rhBMP-2 use; one case involved a malpositioned screw and the other an epidural hematoma. Five cases (2.5%) were associated with local inflammatory changes likely due to rhBMP-2 use. The remaining six cases (3%) were idiopathic; the authors argued these rates were consistent with the known incidence in TLIF independent of rhBMP-2 use. Indeed, as the retrospective study did not include a control group, the likely cause of these cases remains unclear. Nevertheless, even if all six cases were assumed to be due to rhBMP-2, the incidence rate of cases probably attributable to rhBMP-2 (5.4%) is still lower than previous studies.

Another study by Sebastian et al. using a prospective cohort model found 0% incidence of radiculopathy in U.S. adults undergoing TLIF with rhBMP-2 for degenerative diagnoses. However, sample size was small, including only 29 patients, and short-lived cases of radiculopathy may have been missed as the first clinical follow-up was six weeks following discharge. This study falls short of strong contradictory evidence.

The published incidence of radiculopathy in TLIF performed with autograft appears to be between 1.9-5.0%. As described above, this is supported by Rihn et al. and Khan et al., reporting an incidence of 1 of 33 (3%) and 2 of 104 (1.9%), respectively, in their control groups. In addition, Hackenberg et al. report an incidence of 1 of 52 (1.9%) in their prospective study of TLIF safety and efficacy using autograft. Similarly, Potter et al. retrospectively describe an incidence of new onset radiculopathy in 7 of 140 (5%) intervertebral levels treated by TLIF with autograft. Together, these studies suggest that new onset radiculopathy in TLIF with autograft is consistently present in low rates.
B. Heterotopic Ossification

The formation of heterotopic bone when using rhBMP-2 appears to be introduced by Haid et al. in a prospective, randomized trial published in 2004. Although the study intended to include hundreds, only 67 patients enrolled before the trial was suspended as 24 of 32 patients (75%) developed HO on imaging. In contrast, 4 of 31 patients (13%) randomized to the ICBG control developed such findings. Technical differences may be one factor contributing to these high rates; the authors do not describe an effort to place the rhBMP-2-soaked ACS as anteriorly as possible as became common in later studies. The rhBMP-2/ACS complex was placed centrally and no use of surrounding grafts was described. Perhaps surprisingly, the authors concluded their results were encouraging. They stated that no clinical correlation to HO was found through two years of follow-up, arguing that the investigational and control groups sustained similar rates of increased postoperative leg pain. Regardless, such a high rate of HO, even if asymptomatic, concerned some other surgeons.

In the following years, case reports surfaced that detailed clinically significant cases of HO, refuting the conclusion that HO was solely a radiographic phenomenon. Wong et al. reported five patients undergoing either PLIF or TLIF with rhBMP-2 who subsequently developed symptomatic HO. Three of these patients underwent revision to remove the growth and their symptoms subsided. Others criticized the techniques used for these patients noting that in each, the drain was placed deep to the fascia and the opening was irrigated after rhBMP-2 implantation. Later, Rosen, Kiester, and Lee reported 38 cases of symptomatic HO following TLIF with rhBMP-2. The author described consistent “pseudo-pedicle” formations across the cases that were only sufficiently identified with CT.
Next, two small, prospective studies offered contributions. Joseph and Rampersaud studied 33 patients, 22 of whom received rhBMP-2 and the remaining 10 autograft alone.\textsuperscript{13} Patients underwent TLIF (70\%) or PLIF and CT was obtained for all cases. Radiographic HO occurred at 5 of 24 levels (21\%) in the rhBMP-2 group and 1 of 12 levels (8\%) of the autograft at a mean eight months of follow up. The authors concluded this was a clinically insignificant difference, but admitted that the small sample size prevents strong conclusions.

Sebastian et al. found a statistically significant difference between the rhBMP-2 sample and ICBG control.\textsuperscript{7} The study included 77 single-level TLIF patients, with 29 in the interventional group and 48 in the control. HO was observed in 7 of 29 patients (20.7\%) using rhBMP-2 and 1 of 48 (2.1\%) using ICBG (p=0.013). All patients had radiographic follow-up through least six months, but mean follow-up is not specified. Radiographs and CT were used according to surgeon preference, but the proportion of diagnostic choice is not specified.

Another large study was conducted by Niu et al. in 2020. Using a retrospective cohort model that included 927 TLIF patients and routine use of CT during follow up, they reported a statistically significant increase in HO in the rhBMP-2 treated group than the autograft control after controlling for demographics and comorbidities (OR 9.56, p=0.026).\textsuperscript{14} The incidence was 13.5\% in the intervention and 1.6\% in the control. This study’s large cohort, superior imaging, and comparison with a control group provide strong support for their conclusions.

While these studies suggest an association between rhBMP-2 use and HO, several other studies have published very low or nonexistent rates of HO. First, Mummaneni et al. retrospectively described 40 patients undergoing TLIF: 21 received rhBMP-2 and 19 received autograft.\textsuperscript{15} None of the intervention group were found to have symptomatic foraminal bone formation at a mean nine months of follow up. Notably, this study was performed at the same
site with a similar author as Haid et al., as described above, and this time the surgeons placed either ICBG or local autograft posteriorly around the implant. This change in technique may have contributed to a decline in incidence; however, both studies reported that all cases were asymptomatic and radiographic incidence of HO was not reported.

Both studies by Rihn et al. reported cases of HO; the second study will be considered here since it likely includes the patient population in the first study. In 86 patients undergoing single-level TLIF with rhBMP-2, the authors found that two patients (2.3%) developed HO. Both cases were symptomatic. It appears the authors did consider asymptomatic cases of HO as they included only radiographic outcomes in their definition of a complication, and it appears that all patients had CT in follow-up. However, the authors did not report any cases of asymptomatic HO.

Finally, Crandall et al. retrospectively described a large cohort of 509 patients undergoing TLIF at 872 disc levels, all using rhBMP-2. No control group was considered. All patients were followed for a minimum of two years. The authors found ectopic bone formation in three patients (0.59% of sample or 0.34% of levels treated). Again, patients were only evaluated for HO if they were symptomatic; asymptomatic patients did not undergo postoperative MRI or CT. This study concludes the incidence of symptomatic HO appears very low in a large cohort.

The heterogeneity of these studies prevents strong conclusions regarding HO incidence rates. Our interpretation is that radiographic HO appears to occur between 0-21% of the time with rhBMP-2 use, although better studies are needed. We exclude Haid et al. in this estimation on the basis of their technical omission to avoid placing the ACS near the posterior opening. This technique appears routinely practiced today, perhaps due to the risk they identified. Symptomatic
HO occurs, but the large majority of cases appear to be asymptomatic through at least one year of follow-up.

The formation of heterotopic bone in TLIF procedures performed with autograft appears negligible. Although Haid et al. demonstrated a 13% incidence in 31 patients in their autograft control, Mummaneni et al. and Rihn et al. did not report any cases of HO in their samples of 19 patients and 33 patients, respectively.\textsuperscript{2,10,15} Additionally, a literature search did not provide any evidence of HO in TLIF without rhBMP-2.

C. Osteolysis

In 2006, McClellan et al. published the first reports of osteolytic changes occurring in TLIF with rhBMP-2 use. Their retrospective study reviewed 26 patients from their database. At a minimum of three months, CT showed bone resorption in 22 of 32 treated levels (69%) on 1mm cut imaging.\textsuperscript{17} Severe osteolysis, defined as defects in greater than 75% of the area of the interbody graft or greater than 1 square cm, occurred in seven levels (31%). Additionally, five patients with severe osteolysis showed graft subsidence due to loss of endplate integrity. The authors note that selection bias occurred in their inclusion criteria as all patients chosen to be in the study had CT, which is not routinely indicated postoperatively in their practice. Therefore, the study likely selected patients more likely to have complications.

Vaidya et al. supported these findings two years later, performing a prospective study that also found a high incidence of osteolysis.\textsuperscript{18} The study included 36 levels fused by TLIF in 26 patients. Follow-up was relatively lengthy as compared to other studies at a mean 26 months. All patients were imaged by radiography, with CT obtained only in a minority of patients, and assessments were made by three independent radiologists. In these 36 levels, 29 (81%) met their criteria for osteolysis. When these cases presented during follow-up was not specified. Further,
in seven patients with endplate osteolysis, the interbody device migrated, provoked symptoms, and required revision surgery. The authors attributed this to their technique of placing rhBMP-2 centrally in the implant, which caused osteolysis directly above and below and allowed subsidence. Of note, the authors used PEEK cages for all cases, which may be associated with increased rates of osteolysis. This possible relationship will be explored in the next section.

Next, Knox, Dai, and Orchowski reported a much lower but sizable incidence in 2011. The study enrolled 77 lumbar levels belonging to 58 patients, all treated with TLIF using rhBMP-2. Either PEEK or titanium cages were used. On retrospective review of CT images, they found 19 of 77 levels (25%) showed osteolysis at an average of 4.3 months of follow-up. The majority of cases with osteolysis (63%) were considered severe, or osteolysis greater than 1 cm. Six of these cases showed evidence of graft subsidence and the interbody device migrated in four cases, but none required revision.

Together, these studies suggest that endplate osteolysis occurs commonly following rhBMP-2 exposure in TLIF, but the true rate remains unclear. Incidence varies from 25-81% and the studies are limited by selection bias, differences in surgical techniques, and small sample sizes. We consider the study of Knox, Dai, and Orchowski to be the most sound of the group due to routine use of CT and enrollment of consecutive patients.

Three additional studies in the literature also measure osteolytic incidence but consider only symptomatic cases. If these studies found evidence of asymptomatic osteolysis during follow-up, they were not reported. Since this definition of osteolysis is incompatible with direct comparison to the definitions used in the previously reviewed studies, we now consider these reports separately. First, Rihn et al., as a part of the same study covered in the previous sections, found an incidence of 5.8% in its rhBMP-2 group of 86 patients. Patients presented with new
onset, worsening low back pain; two of these resolved at one year and the remaining three persisted beyond one year. Then, Owens et al. found only a single case in a consecutive cohort of 204 patients (0.5%). The patient became symptomatic from foraminal narrowing, but the symptoms eventually resolved without intervention. Finally, in the largest cohort, Crandall et al. found a 0% incidence of symptomatic osteolysis in 872 disc levels of open TLIF using rhBMP-2. No cases of cage subsidence nor migration were identified by routine postoperative radiographs. Overall, it appears that endplate osteolysis is rarely symptomatic. These studies provide important clinical context for the high rates of radiographic osteolysis seen above.

As a final note, it appears that the development of endplate osteolysis is not associated with autograft use. Similar to HO, both Mummaneni et al. and Rihn et al. did not report any cases of osteolysis in their autograft samples of 19 patients and 33 patients, respectively. Though these are small samples, the literature does not expand on these sample, and no studies have presented any additional evidence of endplate osteolysis in TLIF without rhBMP-2.

2.3 Review of Confounding Variables

A. Patient Variables

Studies included in our literature search routinely collected demographic information on patient age, sex, and vertebral level(s) undergoing fusion. Most of these studies also considered BMI, diabetes, and smoking history and status as important confounders. Lubelski et al. identified smoking at the time of operation as a statistically significant confounding variable that increased the rate of radiculitis (OR 1.76, p=0.03). The same study also recorded baseline use of medications including antidepressants, anxiolytics, narcotic analgesics, stimulants, and antipsychotics. Wang et al. expanded their definition of comorbidity beyond diabetes alone, measuring burden with the Charlson Comorbidity Index.
Additional literature describes other variables which may be of relevance to our study’s outcomes. Preoperative steroid injections may alter the risk for developing radiculopathy; some studies have described long-term analgesic effects while others have found only a short-term benefit for a majority of patients.\textsuperscript{21,22} Regarding osteolysis, one retrospective review of 88 patients undergoing TLIF found that osteoporosis, as identified by Hounsfield Units on postoperative CT, was a significant risk factor for radiographic complications including cage subsidence.\textsuperscript{23} This finding was supported by a separate retrospective study following over five thousand patients.\textsuperscript{24} Bhattacharjee et al. found an association between chronic preoperative opioid use and both non-union and revision surgery.\textsuperscript{25} Also, Steinberg et al. found that patients in menopause were at greater risk for non-union and revision surgery which the authors attributed to bone metabolic imbalances in the menopausal state.\textsuperscript{26} Finally, some variables may be a non-issue; two different studies found that both marijuana use and pre-operative vitamin D levels did not affect complication rates or functional outcomes following fusion.\textsuperscript{27,28}

B. RhBMP-2 Dose

The literature records a variety of rhBMP-2 doses used in practice, and the exact dose providing a reliable fusion without increasing the likelihood of complications remains unclear. Although rhBMP-2 associated complications have been felt to be a dose-related phenomenon\textsuperscript{19,29}, this has not been proven. Villavicencio and Burneikiene, using a range of doses from 1.05-12.0 mg/level, did not identify any significant correlations between dose and incidence of radiculitis with logistic regression (p=0.60).\textsuperscript{4} Khan et al. similarly reported no dose-related relationship with TLIF-related complications up to 12 mg/level (p=0.91).\textsuperscript{5} Interestingly, another study of 502 multi-level spinal deformity cases found a low incidence of radiculopathy (1%) and no correlation with dose using total doses averaging 115 mg (range 40-351).\textsuperscript{30}
Regardless, studies show a trend of surgeons selecting smaller doses over time. Crandall et al., after studying a range of doses from 2-12 mg/level, suggest the optimal dose may be 4 mg/level as this dose was not associated with any postoperative complications and produced a 98% fusion rate in 65 patients. Niu et al. attributed the lower incidence of HO in their study than that of previous literature to their dose of 2-3 mg/level. Finally, Khan et al. gradually reduced their dose from 12 mg/level to 4 mg/level during their 17-year study, aiming to reduce complications. As these low doses are providing high fusion rates, there is not a compelling reason to use more. This logic is consistent with initial FDA safety data that reported no benefit beyond a dose threshold in animal models.

C. RhBMP-2 Placement

Surgical technique and rhBMP-2 containment likely play a significant role. Placement of the rhBMP-2/ACS complex near neural structures or the foramina may increase the incidence of radiculopathy or HO. As mentioned, the high rates HO reported by Haid et al. may have been due to placement near the posterior annulus and a lack of posterior packing. Further, Singh et al. demonstrated a rate of symptomatic ectopic bone formation approximately three times higher if rhBMP-2 was mixed with bone graft rather than separately placed anterior to the disc space or within the implant. In contrast, Crandall et al. reported a low incidence of complications after placing the rhBMP-2/ACS complex as far anterior and contralateral as possible within the intervertebral space and adding autograft or allograft for posterior backfill. With the intent to reduce risk of osteolysis, they also did not compress the rhBMP-2/ACS complex, avoiding hyperconcentration, and took care not to overdo endplate preparation. Similar techniques were employed by Niu et al., who reported relatively low rates of HO.

D. Choice of Interbody Device
Polyetheretherketone (PEEK) cages are a popular choice in TLIF due to their radiolucency and light material that decreases the rate of subsidence. However, a few small studies have suggested that PEEK cages may be associated with increased rates of osteolysis and decreased rates of fusion in TLIF compared to titanium cages. A retrospective study involving 48 patients undergoing single-level TLIF with titanium or PEEK implants demonstrated a 60% incidence of osteolysis in the PEEK group at 12 months, but this was not seen in the titanium group. Further, osteolysis was found to be significantly associated with non-union (p=0.007) which was disproportionately seen in the PEEK group. Another retrospective study reviewing 40 patients found osteolysis in 50% of cases using PEEK versus 10% of cases using metal, and PEEK also showed lower rates of fusion. However, this study was underpowered to achieve statistical significance.

**E. Length of Follow-Up**

Finally, time is an important variable that must be considered. The literature has shown varying quantities for radiculopathy and osteolysis depending on when it was measured. First, Helgeson et al. described a reduction of osteolysis with time, appearing in 54% of their cohort between three to six months and decreasing to 41% between one to two years. As a more stark example, another study observing 17 cases of PLIF performed with PEEK cages and a high dose of rhBMP-2 (12 mg/level) showed a 100% incidence of osteolysis at 3 months but 0% at 6 months. Next, rates of radiculopathy appear to decline with time after it is first detected. Rowan et al. showed a high rate (25%) of immediate postoperative radiculopathy strongly associated with rhBMP-2 over ICBG (OR=2.33); however, this decreased to 11.6% at a median follow-up of 4.1 months. Hofstetter et al. explained that the time of onset depends on the underlying pathology. Radiculopathy due to seroma formation can occur as early as ten days after surgery,
while symptoms from inflammatory reactions usually develop between two and four weeks and impingement from HO often manifests beyond two months.

2.4 Review of Relevant Methodology

A. Study Design

Prior studies comparing rhBMP-2 with ICBG have designed multi-center, randomized controlled trials to successfully measure meaningful differences in outcomes. Burkus et al. performed a multicenter randomized controlled trial to compare postoperative outcomes in patients receiving either rhBMP-2 or autograft in ALIF. They enrolled 131 patients across 18 clinical sites during a three-year period ending in 2001. Successful randomization was achieved using a statistical program that produced sequentially numbered envelopes for each enrollment site. Loss to follow-up was 3.8% at twenty-four months postoperative. Statistical significance was achieved in comparing intraoperative, radiographic, and functional outcomes.

Another successful study using this design was performed by Glassman et al. The study investigated outcomes over a two-year period in an elderly population undergoing PLF either with rhBMP-2 or ICBG. Enrolling 106 participants, the study included cases performed by five surgeons. Randomization was performed but the method was not specified. One hundred patients (94%) completed the study in accordance with their allocated treatment and provided data through 24 months of follow-up. Improvements in resultant fusion grades, complication rates, and revision rates in the rhBMP-2 group achieved statistical significance. Together these studies demonstrate feasibility for use of the RCT design to compare rhBMP-2 with ICBG.

B. Randomization

An effective means of randomization is illustrated by the influential Spine Patient Outcomes Research Trial (SPORT), a large, multi-center study consisting of three concurrent
randomized controlled trials. To assure that the number of subjects from each site were approximately the same for each group, the study employed a variably blocked randomization scheme.\(^{42}\) This means that group allocation was done sequentially within blocks of a predetermined size, with the sequence and size of the blocks further varied randomly. This was performed by a computer-generated, automated randomization program at the time of enrollment. The randomization codes were immediately and permanently embedded in the patient’s record, accessible to coordinators and investigators through a secured web site.

C. Recruitment

Both Burkus et al. and Glassman et al. as described above recruited participants from the population of existing patients at cooperating surgery sites.\(^{40,41}\) Consecutive sampling was used in both instances. Of interest, Birkmeyer et al. describe a unique addition in their method of recruitment for SPORT. After candidates for participation were identified by research nurses or participating surgeons from among their clinic’s patients, videotapes were sent to candidates.\(^{42}\) The videotapes provided information explaining gaps in current knowledge, risks and benefits of various treatments, and information about participation in layman’s terms. The authors state this was intended to increase both enrollment and acceptance of randomization by providing education about the importance of the research.

D. Setting

It is relevant for our study to note the caseload of open TLIF per surgical site in the literature in order to determine the number of sites we will need. We report the number of cases, number of surgical sites, and time frame of all studies included in our review that were published after 2010 in Appendix B. We include this restriction on publication date because of the rapid growth in TLIF caseload in recent decades as described in section 1.1 part A.
E. Inclusion and Exclusion Criteria

Our criteria for participation will be informed by indications and contraindications both for TLIF and for rhBMP-2 use. As mentioned, the AANS recommends lumbar fusion, including TLIF, for patients with DDD refractory to conservative treatment. Further, the FDA statement of approval for Infuse™ requires that patients must be skeletally mature, have a diagnosis of DDD as confirmed by history and radiology, and experience symptoms refractory to at least six months of conservative treatment.

Scoliosis and pseudarthrosis or other revision cases can also be indication for TLIF but appear to pose relevant differences in their risk for postoperative complications. Bridwell et al. found an overall complication rate of 36% in their cohort of 81 patients receiving fusion for scoliosis; Bradford, Tay, and Hu reported high rates of residual pain, neurologic injury, and thromboembolism, commenting that the difficulty in achieving coronal and sagittal balance in these cases can limit postoperative success. Dede et al. reported limited benefit in revision cases for patients with DDD, reporting an mean increase of 6.45 points in ODI postoperative scores, while 42% of the cohort reported their overall well-being was worse after surgery through a follow-up of 24 months. Additionally, the majority of patients undergoing revision already have epidural fibrosis at the surgical site which is associated with pain, radiculopathy, and an increased risk of damage to the neural structures during the revision.

The TLIF technique and rhBMP-2 use also have specific contraindications. In their review, Mobbs et al. state TLIF should not be performed in patients with extensive epidural scarring, arachnoiditis, active infection, or conjoined nerve roots in the operative location or in patients with damaged vertebral bodies from trauma or severe osteoporosis. Per the manufacturer’s instructions, rhBMP-2 use is contraindicated in patients with relevant
hypersensitivites or allergies, skeletally immature patients, patients with active malignancy or undergoing treatment for malignancy, pregnant patients, or patients with active infection, resected tumor, or extant tumor at the operative site.

In addition, Colom-Beauchamp et al., utilizing a similar design to our study, excluded candidates who had undergone any previous attempt at fusion and had prior exposure to any bone morphogenetic protein.49 Rihn et al. excluded multi-level cases, cases that included any bone graft substitute or extender other than rhBMP-2, and cases for an indication of tumor, infection, or trauma.2 Villavicencio and Burneikiene also excluded multi-level cases given the rationale that higher per-level doses of rhBMP-2 may be required for multiple level fusions.4

F. Measurement of Primary Outcomes

Radiculopathy: The literature widely agreed in their definition of radiculopathy as new onset radicular symptoms in a dermatomal distribution following surgery.2-4,6,20 The symptom of pain was sometimes measured by the Visual Analog Scale (VAS), including leg pain both ipsilateral and contralateral to the surgical side as described by Sebastian et al.7 In most cases, however, a method of measurement was not specified and radiculopathy was therefore presumed by us to be a clinical diagnosis. Owens et al. included “neurologic deficits” in their definition to emphasize symptoms of weakness, numbness, and tingling but did not mention that an additional measurement tool was used.6 Finally, Villavicencio and Burneikiene clarified that the clinical development of radicular symptoms must be between a few days to several weeks postoperatively after an initial period of improvement.4

As mentioned, there was variability in inclusion or exclusion of compressive etiologies in the definition which influenced the role of imaging. Wang et al. specifically quantified radiculitis, looking for and excluding compressive etiologies using MRI and CT on all patients
with radicular symptoms.\textsuperscript{20} Khan et al. and Rihn et al., on the other hand, did report compressive etiologies and reviewed only CT images for all patients.\textsuperscript{2,5} Other compressive causes not requiring imaging were determined using clinical judgement or laboratory tests. Owens et al. defined seroma as a clinically evident fluid collection with either a negative culture or showing resolution without the use of antibiotics.\textsuperscript{6} This definition distinguished seroma from infection, a separate complication.

\textit{HO:} Assessment of HO was made exclusively by appearance on CT. All studies designed to assess radiographic HO obtained CT at six months for all patients.\textsuperscript{2,10,13,14} Niu et al. presented a grading system to classify ossification as adapted from Joseph and Rampersaund.\textsuperscript{13,14} The scale operates on a grade between 0-3, and HO is defined as grade 3A, 3B, or 3C. Of note, grade 1 or 2 are suggestive of desirable fusion. This grading system is reproduced in Table 2.

\textit{Osteolysis:} Knox et al. presented in detail their method of measuring osteolysis, a technique that has been emulated in subsequent works. The authors included in their retrospective study patients with CT scans at two time intervals—the first within two weeks postoperatively and the other between two and nine months, with most obtained at the six month follow-up visit. All CT scans were performed with 2.5 mm cuts and included sagittal and coronal reconstruction. Then, a reviewer compared the two images for signs of osteolysis, defined as bone resorption 1 mm or greater. Cases were then classified as mild (1-5mm), moderate (5-9mm), or severe (>1cm). Also, Saigal et al. observe that it can be difficult to differentiate rhBMP-2-related osteolysis from infection on imaging alone.\textsuperscript{50} They recommend that infection be excluded based on the patient’s clinical condition and if necessary laboratory tests such as ESR, CRP, and blood cultures.

G. Blinding
Due to the nature of the intervention, surgeons are not able to be blinded to the treatment group. However, the literature demonstrates ways to blind assessment of outcomes. The studies we have reviewed that consider a control group have consistently blinded evaluation of radiographic outcomes.$^{2,7,41}$ As an example, Khan et al. used two blinded orthopedic surgeons to separately evaluate all imaging, and a third independent observer resolved any discrepancies.$^{5}$

Another study demonstrated that it may be possible to effectively blind the participants in our study to their group allocation. This study enrolled 92 patients undergoing posterior LIF with ICBG and blinded them to which of their iliac crests would be providing the graft.$^{51}$ Donor sites were randomized 1:1 between the left and right iliac crest and harvest was performed through the primary incision for the fusion. At six weeks, three months, six months, and one year, only 22 patients (24%) were consistently correct in identifying the harvest site. Additionally, VAS pain scores did not differ between the affected and unaffected iliac crests at all time points.

H. Sample Size Calculation

Since each of our three primary outcomes will be assessed for statistical significance individually, our sample size for each was considered separately. Each calculation was performed using, in our judgement, the best incidence rates from our literature review.

Radiculopathy: To calculate our sample for radiculopathy, we used the incidence rates as published by Kahn et al. This work was selected because it reflects our study well by considering only patients undergoing TLIF for DDD, it quantifies radiculopathy rather than radiculitis alone, and it quantified the incidence in an ICBG control group as a part of the same study. Additionally, the authors found the lowest incidence rate among the studies that reported a difference and would therefore overestimate the sample size if the larger rates reported by the
other studies are in fact closer to the true value. The reported incidence in this study was 10.8% for the interventional group and 1.9% for the control.

**HO:** We selected the rates reported by Niu et al. for our calculation of HO. The study has the largest cohort, by far, of the studies considered, had superior methodology by routine use of CT, and considered a control group as a part of the same study. Similarly, this study also reports the lowest incidence among the studies that found a difference. This work found an incidence of 13.5% for the rhBMP-2 group and 1.6% in the autograft control.

**Osteolysis:** Finally, the work of Rihn et al. appears to provide the best estimate for incidence of osteolysis. First, the rates report the incidence of symptomatic cases, which bodes well for clinical significance. Next, the study compares incidence against an ICBG control. The sample size is smaller than some of the other studies but larger than all the studies considering radiographic cases alone. The incidence is also much lower than the other studies including radiographic appearance in their definition. Thus, this work will provide a large enough sample to consider both radiographic and symptomatic cases. The incidence published by this study is 5.8% for the intervention and 0% in the control.

### 2.5 Conclusion

While rhBMP-2 associated complications following TLIF have been described in the literature for over a decade, uncertainty continues to limit providers’ ability to fully inform patients. All existing studies are observational in design, and many involve small sample sizes. Most have a single-center focus, limiting generalizability to different practices. Heterogeneity exists in outcome definitions, execution, and reporting of adverse events. These limitations may explain, at least in part, the range of conclusions that are reported. Additionally, scientific flaws inherent to cohort studies such as a propensity toward selection bias resulting in baseline
differences between groups prevent definitive conclusions regarding the complication rate of rhBMP-2 from this data. This highlights the need for a prospective, multi-center, experimental study of sufficient sample size to characterize complication incidence. In our review of the literature, we have surveyed previous reports of our primary outcomes, identified confounding variables, and described methodology that will inform our study design. Our proposed study will add to the literature by using a high-quality study design to increase our understanding of this product that thousands of patients are already receiving.

2.6 References


CHAPTER 3: METHODS

3.1 Study Setting and Design

We propose a randomized, single-blinded, two-armed, controlled trial to determine the relative risk of developing three postoperative complications in patients who receive rhBMP-2 versus ICBG for TLIF. This study will include ten sites across the United States. The rationale for our number of included sites is provided in Appendix B along with a list of sites intended to be approached for participation. The study will be performed over a span of two years. Participants will be recruited on a rolling basis given that their surgery will occur by the end of month 17 of the study. After enrollment, participants will be randomly assigned to either the intervention or control group in a 1:1 ratio without stratification. Surgeons will be informed of the allocation before surgery is performed. Postoperatively, patients will be assessed for primary and secondary outcomes during follow-up visits at three weeks, three months, and six months. A timeline is shown in Table 1.

| Month | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  | 16  | 17  | 18  | 19  | 20  | 21  | 22  | 23  | 24  |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Events|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Enrollment| X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Surgeries Performed| X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Data Collection| X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Statistical Analysis| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Table 1: Proposed Timeline

3.2 Study Population, Sampling, and Recruitment

The target population of our study will be all U.S. adults between the ages of 21-75 with a diagnosis of lumbar DDD and symptoms refractory to an adequate trial of conservative management. Our study population will be all persons within the target population that are
patients actively seen at our participating medical centers within the study timeframe. Our sample will be recruited from our study population by means of consecutive sampling. Inclusion and exclusion criteria will be in accordance with guidelines and/or precedents for both TLIF surgery and rhBMP-2 use as detailed in our review in section 2.4 part C. For logistical reasons, patients must also report their intention to attend follow-up with the same medical team that performed the surgery for at least six months postoperatively. We also include that patients must have previously elected to undergo LIF with consent for a posterior approach. This was decided for ethical reasons as we want to avoid undue influence on patient choice of one approach over another. Patients who do not agree to undergo randomization will be excluded from the study. A full list of inclusion and exclusion criteria is included in Figure 1.

**Figure 1:** Inclusion and Exclusion Criteria

Candidates will be identified as eligible by their surgeons, other health care workers involved with the patient’s care, or research nurses at each site. Candidates will be approached by a research nurse for an explanation of the trial either in person, through video chat, or by phone. Candidates will then receive a link to a video further explaining the study’s rationale, design, and risks and benefits, with an invitation to participate. At this time, candidates may give
their consent, but they will also be encouraged to discuss enrollment with their surgeon or an advanced practice provider associated with the medical group either in person or via telehealth if they wish. New candidates will continue to be approached until sufficient enrollment is reached.

3.3 Participant Protection and Confidentiality

We will attain Institutional Review Board (IRB) approval for our study at each enlisted site prior to recruitment according to the 100 IRB Protocol for review of human subject research protocols or FDA-regulated activities involving human participants. In accordance with IRB requirements, candidates will receive our version of the Authorization and Consent for Participation in Research Project form which includes a description of the research project, risks and benefits, efforts to ensure confidentiality, and a statement of voluntary participation and withdrawal. A sample consent form is included in Appendix A. Medical interpretive services will be utilized for any patient whose primary language differs from that of the researcher providing this information. All candidates must provide written, informed consent prior to enrollment.

To protect patient confidentiality, all personnel involved in the study will complete a Health Portability and Accountability Act (HIPAA) training session and will submit the certification to their corresponding IRB. Participant data will be kept confidential within password-protected encrypted servers and according to HIPAA guidelines. Each participant’s file will be de-identified and codified prior to data analysis. De-identified information may be kept at the conclusion of the study if participants are separately consented to participate in any additional studies that are designed to elaborate on the present one. If participants are not enrolled in any further studies at the conclusion of this study, all data will be destroyed.

3.4 Study Variables and Measures
A. Baseline Assessment

Enrolled participants will be assessed for demographic characteristics, comorbidity, and baseline functional status before randomization. We will record age, sex, type of DDD diagnosis, vertebral level(s) requiring operation, BMI, presence of diabetes or menopause, comorbidities per the Charlson Comorbidity Index (CCI), smoking status and history, medications with emphasis on antidepressants, anxiolytics, narcotic analgesics, and oral or injectable corticosteroids, ODI score, VAS leg pain score, VAS back pain score, and Short-Form 36 score. The presence of osteoporosis will be assessed by Hounsfield Unit measurements on the first set of CT images collected for all patients within two weeks after surgery as described by Lee et al.¹

B. Intervention

The experimental group will undergo TLIF procedure utilizing rhBMP-2 (brand Infuse™) at a dose of 4.2mg/level. Surgeons will be instructed to place the rhBMP-2/ACS complex as anterolateral as possible, with care to avoid placement close to the posterior annulus, and to avoid over-compression of the ACS. The procedure will be performed using any titanium cage of the surgeon’s choice, and local autograft or allograft can be used for backfill at the surgeon’s discretion. Each surgery will be performed in conjunction with pedicle screw instrumentation and posterolateral fusion.

C. Control

Our control arm will receive TLIF in conjunction with ICBG. The autograft will be harvested through the primary surgical incision to avoid cutaneous damage that would compromise patient blinding. Titanium cages will be used for all cases. The primary graft may be combined with local autograft or allograft at the surgeon’s discretion. Pedicle screw instrumentation and posterolateral fusion will also be performed with each procedure.
D. Measurement of Primary Outcomes

Radiculopathy will be defined as new onset radicular symptoms in a dermatomal distribution following surgery. We include compressive etiologies resulting from hematoma, seroma, scar tissue, or heterotopic bone and non-compressive radiculitis in our definition. We will exclude symptoms from iatrogenic injury and incomplete decompression. Back and leg pain will be assessed preoperatively, perioperatively, and at each follow-up visit with VAS ratings as shown in Appendix C. Other neurologic symptoms such as weakness, numbness, and tingling will be described clinically. As radiculopathy is a diagnosis of exclusion, any laboratory tests or earlier imaging needed to rule out other potential etiologies will be used appropriately based on clinical suspicion.

Both HO and osteolysis will be evaluated by high-resolution axial CT with sagittal and coronal reconstruction. Images will be obtained within two weeks after surgery and again at six months. We define HO as any new bone growth that extends outside of the annulus and encroaches on the spinal canal or vertebral foramen. We adopt the grading system presented by Niu et al., reproduced in Table 2, where HO is defined as ossification of grade 3A, 3B, or 3C.²

<table>
<thead>
<tr>
<th>Ossification Grade</th>
<th>CT Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No ossification</td>
</tr>
<tr>
<td>1</td>
<td>Ossification limited in the disc space</td>
</tr>
<tr>
<td>2</td>
<td>Ossification extending to but not beyond the outer aspect of the annulus</td>
</tr>
<tr>
<td>3A*</td>
<td>Posterior ossification extending centrally into the spinal canal</td>
</tr>
<tr>
<td>3B*</td>
<td>Posterolateral ossification extending into the foramen</td>
</tr>
<tr>
<td>3C*</td>
<td>Posterior and posterolateral ossification</td>
</tr>
</tbody>
</table>

*Heterotopic Ossification is defined as a grade of 3A, 3B, or 3C

Table 2: Heterotopic Ossification Classification

Osteolysis will be defined as new bone resorption of greater than 1 mm at six months. We will use the same grading system as Knox et al., classifying cases as mild (1-5mm), moderate (5-
9mm) or severe (>1cm).³ Radiographic interpretation will be performed by two independent orthopedic surgeons for all images. Any discrepancies will be resolved by a third orthopedic surgeon. Review will occur at the time the images are obtained so that any radiographic ambiguities, such as infection, can be clarified with clinical assessment or laboratory testing.

E. List of Secondary Outcomes

Data will be collected on all patients for intraoperative, perioperative, and postoperative outcomes. Intraoperative and perioperative variables will be assessed up until the patient leaves the hospital, but postoperative outcomes will be assessed during follow-up. The intraoperative outcomes of operative time, estimated blood loss, volume of transfusion required, neurologic injury, and vascular injury will be recorded. Perioperative outcomes include length of hospital stay, minor adverse events, and severe adverse events. Minor adverse events will consist of any of the following: wound complications, abscess formation, acute kidney injury, deep vein thrombosis, pneumonia, urinary tract infection, anemia requiring transfusion, and ileus. Cardiac arrest, stroke, sepsis, myocardial infarction, renal failure, pulmonary embolism, ventilator use, unplanned intubation, return to operating room, and death will be considered severe adverse events. Finally, postoperative outcomes will be divided into secondary postoperative complications and indicators of surgical success. These secondary postoperative complications include iliac crest pain, iliac crest infection, interbody cage subsidence, interbody cage migration, change in spinal curvature or height, revision surgery, and development of new malignancy. Indicators of surgical success are postoperative ODI score, VAS score for back pain, VAS score for leg pain, Short-Form 36 score, fusion at six months, and reoperation rate.

Although we recognize surgical success is routinely assessed beyond six months, we include these postoperative outcomes to allow our cohort to easily transition to any additional
studies that may wish to follow our participants beyond our proposed time frame. It is possible our study, which will require substantial effort and resources, would be attractive to further analysis of outcomes relevant to the rhBMP-2 debate such as overall fusion rate and long-term incidence of malignancy.

3.5 Methodology

A. Assignment of Intervention

Computer-generated randomization will assign patients in a 1:1 ratio to either the intervention or control group using a variably blocked randomization scheme. The allocation will occur at the time of enrollment and will immediately be imbedded in the patient’s medical record. Surgeons will view the allocation before operating but the patient will remain blinded.

B. Blinding

Due to the nature of the intervention, each surgeon and their operating team will not be blinded to their participants’ group allocations. The same providers are encouraged to follow the patient postoperatively, so all clinical assessments will also not be blinded. However, all radiographic interpretations will be made by blinded orthopedic surgeons. Patients will remain blinded to their group allocation throughout their participation in the study.

C. Data Collection

Baseline assessment will be conducted by a combination of patient interview, patient self-administered survey, and physician survey. Intraoperative outcomes will be collected during routine postoperative documentation by the surgical team. Perioperative outcomes will be recorded either by the surgical team or by research nurses that will follow patients during their hospital course. Research nurses will ensure an assessment is entered for each outcome and provide continuity if patients are transferred to a medical service during their stay. Radiculopathy
and secondary postoperative outcomes will also be assessed by patient interview, patient self-administered survey, and physician survey each follow-up visit. The clinical primary outcome, radiculopathy, must be collected and entered to the study’s data system by the surgeon or an associated medical doctor or advanced practice provider. Secondary outcomes may be collected by any health care provider with the surgical team or a research nurse. The other primary outcomes will be assessed by imaging. High-resolution CT will be obtained within two weeks of surgery and at six months postoperatively. Interpretation will be made within one week after imaging is obtained by the independent assessors who will then directly enter their findings into the study’s data system.

3.6 Statistical Analysis

Data analysis will be conducted during the final month of the study according to the intention-to-treat principle. If the cross-over rate between the intervention and control groups is greater than 3%, per-protocol analysis will also be performed; however, we do not anticipate that this should occur as participants must agree to be blinded to their group to enroll in the study. The incidence proportion of each primary outcome will be statistically analyzed using Fisher’s exact test. Further, osteolysis will be analyzed for statistical differences in severity between our intervention and control groups using the Wilcoxon Rank Sum test. Other baseline information and secondary outcomes will be analyzed with an unpaired t-test for continuous variables and Fisher’s exact test for categorical variables. If differences are found in baseline characteristics between the intervention and control groups, single or multiple logistic regression will be performed for the primary outcomes and secondary categorical outcomes, and single or multiple linear regression will be performed for secondary continuous outcomes. Measurements of function, including ODI, VAS, and SF-36 scores, will be compared pre and postoperatively.
within each allocation group using paired t-tests. Then, the mean difference of pre and postoperative scores will be compared between groups using an unpaired t-test. A p value < 0.05 will indicate a significant result for all tests. Finally, if our study is performed with industry sponsorship, we will share our raw data with independent reviewers for statistical analysis to limit industry influence.

3.7 Sample Size Calculation

The sample size was calculated using what we consider the best incidence proportion published in the literature for each of our three primary outcomes as previously described in section 2.4 part G. The calculation was made with G*Power 2 software, and each calculation accounted for an alpha of 0.05, power of 80%, and a 1:1 allocation between the intervention and control groups. Two-sided tests were used for calculations of radiculopathy and HO since the literature reports these complications can exist with the use of our control, ICBG. Although we concluded that the risk of HO with autograft appears to be negligible, since two studies reported a very low incidence we employed a two-sided test so that we may err on the side of an overestimation rather than undermine our study with an insufficient sample. However, since our literature review did not reveal any cases of endplate osteolysis following autograft, we employed a one-sided test for our third calculation. We also, we believe, overestimated the incidence proportion of osteolysis in our control for this calculation in order to use a non-zero value; our literature review found an incidence of 0%, but we used 0.5%. This value would only overestimate our final sample calculation. The calculation for each complication is shown in Appendix D.

The sample size for radiculopathy, heterotopic ossification, and osteolysis was calculated to be 256, 164, and 302, respectively. We will use the largest calculation, 302, as our sample size.
for the entire study so that we may achieve statistical significance for each outcome. We note the benefit that this sample size will increase the power in the analysis of the other two primary outcomes. Finally, we will account for a 15% drop out/loss to follow up rate for a final sample size of 356. We will allocate 178 participants in each treatment arm.

3.8 Resources


CHAPTER 4: CONCLUSION

4.1 Study Advantages and Limitations

Our study’s biggest advantage is its experimental design. The RCT format minimizes selection bias and known and unknown confounding variables to a degree not previously seen in the literature. We are also the first study of those measuring our primary outcomes to blind patients to their intervention which will reduce respondent and recall bias. Our use of an equally sized control group will allow direct comparison between our intervention and the gold standard treatment. Finally, while most other studies have retrospectively reviewed available patient records, our prospective model has been designed to enroll a sufficient sample size to achieve statistical significance.

Additional advantages exist as a result of our design and methodology. Our routine use of high-resolution CT will provide superior detection of HO and osteolysis compared to radiograph. By limiting our fusion approach to TLIF, we will help characterize the complications of this technique independent from PLIF. Finally, our use of multiple sites and surgeons across the United States will increase the generalizability of our results to a broad population.

Despite these strengths, we recognize our study will have several limitations. First, our lack of blinding of clinical assessments leaves the possibility for observation bias. Next, some of the studies in our literature review have suggested a lower incidence of our primary outcomes than the figures used in our sample size calculation, and our study will not be powered to detect differences this small. Our inclusion of both compressive and non-compressive causes in our definition of radiculopathy may complicate clinical application of our results as these two etiologies may produce different courses of symptoms. Also, our six months of follow-up may
not be long enough to describe the course of some symptoms that may resolve spontaneously, and it will not report long-term values of fusion rates nor reductions in disability.

There are also limitations of generalizability. Our study will not be generalizable to minimally-invasive TLIF which is rapidly increasing in use and may broadly become the technique of choice in the coming decades. Likewise, our results will not be generalizable to the minority of TLIF cases that are performed for non-degenerative diagnoses.

Finally, we recognize our study is intensive and will require a significant amount of personnel, coordination, and resources. The necessary financial investment will likely require the support of industry, providing a potential source of bias.

4.2 Clinical Significance

Hundreds of thousands of patients have turned to TLIF to alleviate their symptoms from lumbar DDD.\(^1\) Hundreds of thousands more are likely to undergo TLIF in the coming decades, with the number of cases boosted by the aging population, higher expectations of functioning, and the continued advancement of surgical techniques. Future procedures have the potential to be performed with rhBMP-2, a product that can help patients avoid the comorbidity of ICBG and may improve their chances of a successful surgery.\(^2,3\) These meaningful advantages explain why rhBMP-2 is currently used in about half of TLIF surgeries despite remaining as an off-label use.\(^4\)

Unfortunately, the drive to use rhBMP-2 in TLIF has outpaced our knowledge of its safety. The literature has described cases of radiculopathy, heterotopic ossification, and vertebral osteolysis after this use of rhBMP-2 which can cause patients further back pain, leg pain, and radicular symptoms or even destabilize hardware and require reoperation.\(^5,6\) Yet, the risk of these complications has not been studied well, and patients and surgeons must weigh the risks and benefits of rhBMP-2 with incomplete information.
Simply put, patients deserve to have high-quality studies that examine the product in the way it is being used. Our study will meet this need by performing a randomized controlled trial, the premier study design of scientific research, and will enroll patients from across the United States for this purpose. Since rhBMP-2 is already used so widely, our findings will be of great relevance regardless of the direction of our final conclusions. If our findings support the claim that postoperative complications appear more common with rhBMP-2 use, patients will benefit as they may avoid these complications by choosing alternative options. If rhBMP-2 appears safe, patients will also benefit as they may be reliably encouraged to enjoy the advantages of its use. Finally, if our study finds no difference in postoperative complication incidence, our findings could help support FDA approval for rhBMP-2 in TLIF. This formal approval would reduce financial and logistical barriers to rhBMP-2 use and increase accessibility for patients who may benefit.

4.3 Resources

APPENDIX A: Consent and Privacy Rule Authorization Form

CONSENT FOR PARTICIPATION IN A HUMAN RESEARCH PROJECT

Study Title: Complications of Recombinant Human Bone Morphogenetic Protein-2 Use in Open Transforaminal Lumbar Interbody Fusion
Principal Investigator: Arya Varthi, MD

Invitation to Participate and Description of Project

You are invited to participate in a research study designed to quantify adverse events that may occur with use of recombinant human bone morphogenetic protein-2 (rhBMP-2) in transformaminal lumbar interbody fusion (TLIF). You have been asked to participate in this study because your provider identified you as being an adult age 21-75 with a diagnosis of degenerative disc disease that has not improved with conservative management, and you have decided to undergo surgical treatment using lumbar interbody fusion. RhBMP-2 is a product that causes bone formation and can be useful in fusion surgery. There will be approximately 356 participants in this study across ten clinical sites in the United States.

To decide whether or not you wish to be a part of this research study you should know enough about its risks and benefits to make an informed decision. This consent form gives you detailed information about the research study, which a member of the research team will discuss with you. This discussion should go over all aspects of this research: its purpose, the procedures that will be performed, any risks of the procedures, possible benefits and possible alternative treatments. Once you understand the study, you will be asked if you wish to participate; if so, you will be asked to sign this form.

Description of Procedures

If you decide to participate in this study, you will be asked questions about your demographic information and medical history before you have surgery. Information will be collected about your age, gender, spinal diagnosis, medications, and other medical conditions you may have. You will also complete a few surveys on your own that will help us learn about the symptoms you experience and how it impacts your life. These surveys are often completed by patients before undergoing spine surgery, even if they are not a part of a research study. The surveys you will complete are the Visual Analog Scale, Oswestry Disability Index, and Short-Form 36 score.

After this information is collected, a computer program will randomly assign you to either Group A or Group B. The group you are assigned to will be made known to your surgeon, but you will not be told which group you are in until after the study is completed. We temporarily withhold this information from you because it helps increase confidence that our results are accurate. If you are in Group A, your surgeon will use rhBMP-2 during your TLIF procedure. If you are in Group B, your surgeon will not use rhBMP-2. Instead, your surgeon will take material from your hip bone (also called your iliac crest) to be used in your TLIF procedure. If material is required
to be taken from your iliac crest, this will be performed in the first few minutes after you are put under anesthesia for your surgery.

The study will also collect information about how you do in the hospital after your surgery and before you go home. If you have any surgical or medical complications, this will be recorded.

After you leave the hospital, you will attend follow-up appointments with your surgeon or other medical doctors or advanced practice providers associated with your surgeon. You will be asked to attend appointments three weeks, three months, and six months after your surgery. These follow-up appointments are mainly routine visits where your provider will be able to assess your progress and monitor for complications. Your provider will ask you about any symptoms you may be experiencing. You will also be asked to fill out the same surveys that you did at the beginning of the study (VAS, ODI, SF-36). The surveys should only take a few minutes.

Finally, you will also be asked to obtain radiographic images of your lower spine using high-resolution computed tomography (CT) scans. Sometimes they are also called CAT scans. You will be asked to obtain these scans within two weeks following your surgery, and then again around six months. These images will be reviewed by your doctor in order to treat you. However, the images will also be shared with independent surgeons who will assess them to see if you have developed any complications. These surgeons will not able to identify you and will not be given any of your personal or medical information.

In this study we are evaluating if adults who receive rhBMP-2 with TLIF experience different amounts complications after surgery than patients who use bone harvested from their iliac crest.

A description of this clinical trial will be available on http://www.ClinicalTrials.gov, as required by U.S. Law. This website will not include information that can identify you. At most, the website will include a summary of the results. You can search this website at any time.

**Risks and Inconveniences**

You should know that using rhBMP-2 is not approved by the United States Food and Drug Administration (FDA) for use in TLIF. Using rhBMP-2 in this manner is considered an off-label use. Adverse events that have been reported with rhBMP-2 include the following: back pain, leg pain, inflammation, bone formation outside of the intended space, fluid collections, weakening of the vertebral surfaces, development of antibodies, cancer, implant loosening or movement, non-union, retrograde ejaculation, allergic reaction, anaphylactic reaction, bone fracture, bowel or bladder problems, change in mental status, damage to internal organs, dural tears, gastrointestinal complications, incisional complications, infection, itching, localized edema, and scar formation.

Alternatively, taking bone from the iliac crest to use in your surgery is considered the standard treatment. Known risks associated with iliac crest graft include: nerve damage, blood vessel damage, infection, pelvic fractures, pain, bleeding, hematoma, sensory loss, and defect hernias.

Participation in this study may involve risks that are currently not known.
Benefits

If you receive rhBMP-2 as a part of your surgery, you may experience a benefit from not having to give a part of your iliac crest to be used in the surgery. You may also have a slightly increased chance of having a solid fusion. If you participate in this study in either group, you will help future patients understand if rhBMP-2 is associated with more risks or if it is as safe as standard treatment.

Economic Considerations

There is no direct compensation associated with this study. There are no costs associated to participate as the intervention will be given free of charge should you be randomized to receive it. You will be still responsible for any deductibles or copays required by your insurance company for your TLIF surgery. If you have any questions regarding your insurance coverage, please contact your insurance company directly. There will be no financial penalty for withdrawing for the study.

Confidentiality

Any identifiable information that is obtained in connection with this study will remain confidential and will be disclosed only with your permission or as required by U.S. or State Law. When the results of the research are published or discussed in conferences, no information will be included that would reveal your identity unless your specific consent for this activity is obtained.

Information about your study participation will be entered under a unique identification number in a password-protected software and stored on secure servers until needed for statistical analysis. Health Insurance Portability and Accountability Act (HIPAA) standards will be met and maintained for all devices and personnel. Only approved research personnel will have access to your medical records in order to verify information required for the study. Any information that is not relevant will not be extracted from your medical records. Data auditing will be performed at random points throughout the trial to ensure no inappropriate viewing or disclosure of protected health information has occurred. All records no longer needed for research purposes will be destroyed in accordance with HIPAA requirements. All data and records used in the study will be destroyed after data analysis has concluded, unless you have given your specific consent to share your data with separate research study. Representatives from committees at each participating site that review, approve, and monitor research on human subjects may inspect study records during internal auditing procedures. However, these individuals are required to keep all information confidential.

Voluntary Participation and Withdrawal

Participating in this study is voluntary. You are free to choose not to take part in this study. Refusing to participate will involve no penalty or loss of benefits to which you are otherwise entitled (such as your health care outside the study, the payment for your health care, and your health care benefits). You will not be able to enroll in this study if you do not consent to be
randomized to either Group A or Group B. You must also acknowledge that you will not be informed of which group you are placed in until after your six month follow-up visit.

If you do become a participant, you are free to stop and withdraw from this study at any time during its course. To withdraw from the study, you can call a member of the research team at any time and tell them that you no longer want to take part. Withdrawing from the study will involve no penalty or loss of benefits to which you are otherwise entitled. It will not harm your relationship with your providers or the hospital at which they are employed.

When you withdraw from the study, no new health information identifying you will be gathered. Information that has already been gathered may still be used and given to others until the end of the research study, as necessary to ensure the integrity of the study and/or study oversight.

Questions

We have used some technical terms in this form. Please ask about anything you don't understand and consider this research and the consent form carefully, as long as you feel is necessary, before you decide to participate.

Authorization

I have read (or someone has read to me) this form and have decided to participate in the project described above. Its general purposes, the particulars of my involvement and possible hazards and inconveniences have been explained to my satisfaction. My signature also indicates that I have received a copy of this consent form.

Name of Subject:_______________________________________
Signature:_____________________________________________
Relationship:___________________________________________
Date:___________________________

Signature of Principal Investigator: ____________________________
Or Signature of Person Obtaining Consent: _______________________
Date: ___________________________________________________

If you have further questions about this project or if you have a research-related problem, you may contact the Principal Investigator Arya Varthi, MD.

If, after you have signed this form you have any questions about your privacy rights, please contact your site's Privacy Officer at XXX-XXX-XXXX. If you would like to talk with someone other than the researchers to discuss problems, concerns, and questions you may have concerning this research, or to discuss your rights as a research subject, you may contact your site’s Human Investigation Committee at XXX-XXX-XXXX.
APPENDIX B: Estimation of Number of Participating Sites

This table includes information from all studies included in our literature review that performed exclusively TLIF and were published after 2010. The rate of cases per site per month was calculated in order to estimate how many participating sites we will need to enroll for our study.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Includes</th>
<th># Cases</th>
<th># Sites</th>
<th>Duration (Months)</th>
<th>Rate&lt;sub&gt;(Cases/Site/Month)&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crandall et al.</td>
<td>rhBMP-2 only</td>
<td>509</td>
<td>1</td>
<td>96</td>
<td>5.3</td>
</tr>
<tr>
<td>Khan et al.</td>
<td>both</td>
<td>191</td>
<td>1</td>
<td>216</td>
<td>0.88</td>
</tr>
<tr>
<td>Lubelski et al.</td>
<td>rhBMP-2 only</td>
<td>216</td>
<td>1</td>
<td>107</td>
<td>2.0</td>
</tr>
<tr>
<td>Owens et al.</td>
<td>rhBMP-2 only</td>
<td>204</td>
<td>1</td>
<td>48</td>
<td>4.3</td>
</tr>
<tr>
<td>Sebastian et al.</td>
<td>both</td>
<td>77</td>
<td>1</td>
<td>52</td>
<td>1.5</td>
</tr>
<tr>
<td>Villavicencio and Burneikiene</td>
<td>rhBMP-2 only</td>
<td>204</td>
<td>unclear</td>
<td>31</td>
<td>NA</td>
</tr>
<tr>
<td>Wang et al.</td>
<td>both</td>
<td>223</td>
<td>1</td>
<td>78</td>
<td>2.9</td>
</tr>
<tr>
<td>Niu et al.</td>
<td>rhBMP-2 only</td>
<td>927</td>
<td>1</td>
<td>126</td>
<td>7.4</td>
</tr>
<tr>
<td>Knox et al.</td>
<td>rhBMP-2 only</td>
<td>58</td>
<td>1</td>
<td>67</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Average= 3.1/month = 49.6 per site for length of study = 7.2 sites = 8 sites
We add 2 additional sites for margin = 10 sites

This estimate may be an overestimate of the true value of sites we need for our study. Each site may have performed more TLIFs than were included in their studies. We do know that several studies excluded several cases due to insufficient follow up or not having the preferred method of imaging available. Further, studies that reported only cases which used rhBMP-2 may or may not have performed many other TLIFs at that site that did not use rhBMP-2. Finally, the present-day amount of cases are likely much higher than those reported by these studies since the amount of fusion surgeries has doubled in recent decades. These studies describe the caseload between 1997-2008.

Conversely, this estimate may be an underestimate of the true value needed. This estimate calculates the number of cases that may be performed at an average clinical site, but it is not known how many of these patients would be willing to be enrolled in a blinded RCT.

Sites to be approached for participation: Yale University School of Medicine, New Haven, CT; Northwestern University Feinberg School of Medicine, Chicago, IL; Washington University in St. Louis, St. Louis, MO; Cleveland Clinic Lerner College of Medicine, Cleveland, OH; University of Louisville School of Medicine, Louisville, KY; Emory University School of Medicine, Atlanta, GA; Sonoran Spine Center, Phoenix, AZ; NYU Langone Orthopedic Hospital, New York, NY; Wake Forest School of Medicine, Winston-Salem, NC; Walter Reed Army Medical Center, Washington, DC.
APPENDIX C: Visual Analog Scale for Back and Leg Pain

Name: _______________________________

Date of Birth: ______ / ______ / ______

Date:_____/_____/_____

Visual Analog Scale: Back and Leg Pain

1. Please mark on the line below how much pain you have had from your back, on average, over the past week:

   ![Visual Analog Scale for Back Pain]

   0          10
   no pain    worst pain
   imaginable

2. Please mark on the line below how much pain you have had in your worst leg, on average, over the past week:

   ![Visual Analog Scale for Leg Pain]

   0          10
   no pain    worst pain
   imaginable

3. If you have pain in the other leg, please mark on the line below how much pain you have had on average, over the past week:

   ![Visual Analog Scale for Other Leg Pain]

   0          10
   no pain    worst pain
   imaginable
APPENDIX D: SAMPLE SIZE CALCULATION

The following calculations were made using G*Power 2 software.

Alpha (level of significance) = 0.05
Beta (type II error) = 0.20
Group allocation ratio= 1:1

This study is powered at 80% for our hypothesis regarding osteolysis, meaning that 80% of studies would be expected to produce a statistically significant result for the given effect size. Our study’s power is greater than 80% for our hypotheses regarding radiculopathy and heterotopic ossification, respectively.

Radiculopathy calculated sample size: 256

HO calculated sample size: 164

Osteolysis calculated sample size: 302

By using the largest calculated sample size, 302, and accounting for a 15% loss to follow-up/drop out rate, our total N = 356, with 178 participants in each group.
BIBLIOGRAPHY


72


