Spinal Fusion for Adult Degenerative Scoliosis: RCT of Robot-Assisted Versus Navigation Guidance

Kyle Hanks
Yale Physician Associate Program, Kyle.hanks@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
https://elischolar.library.yale.edu/ysmpa_theses/178

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.
SPINAL FUSION FOR ADULT DEGENERATIVE SCOLIOSIS: RCT OF
ROBOT-ASSISTED VERSUS NAVIGATION GUIDANCE

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

In Candidacy for the degree of
Master of Medical Science

May 2023

Kyle Hanks, PA-SII
Class of 2023
Yale Physician Associate Program

Dr. Peter Whang, MD, FACS
Associate Professor of Ortho & Rehab
YSM Dept affiliation
# Table of Contents

**Table of Contents**

List of Tables ........................................................................................................ iv

Abstract .................................................................................................................. v

**Chapter 1: Introduction** .................................................................................. 1

1.1 Background ........................................................................................................ 1
   i. Low back pain ................................................................................................. 1
   ii. Adult spinal deformities: Etiology, pathophysiology, and clinical presentation ................................................................................................. 2
   iii. Adult degenerative scoliosis ........................................................................ 4
   iv. Lumbar interbody fusion approaches .............................................................. 6
   v. Pre-operative and intra-operative guidance of pedicle screw placement .......... 7
   vi. Confirmation of pedicle screw placement ....................................................... 8
   vii. Radiation exposure ...................................................................................... 9

1.2 Statement of the Problem .................................................................................. 9

1.3 Goals and Objectives ....................................................................................... 12

1.4 Hypothesis ......................................................................................................... 13

1.5 Definitions ........................................................................................................ 13

1.6 References ....................................................................................................... 13

**Chapter 2: Review of the Literature** ................................................................. 18

2.1 Introduction ...................................................................................................... 18

2.2 Empirical Studies Review: Pedicle Screw Accuracy by Robot-Assisted Guidance and Secondary Outcomes ........................................................................................................... 19
   i. Analysis of pedicle screw placement accuracy from studies using similar treatment arms ................................................................. 19
   ii. Analysis of pedicle screw placement in correction of ADS with varying control groups ........................................................................................................... 21
   iii. Analysis of pedicle screw placement in correction of adolescent scoliosis ...... 22
   iv. Results from differing control groups and/or spinal pathologies ................. 22
   v. Summary ....................................................................................................... 24

2.3 Study Methodology Review ............................................................................. 25
   i. Study design .................................................................................................. 25
   ii. Number of centers, surgeons, and sample sizes of studies ......................... 26
   iii. Effect size and power .................................................................................. 27
   iv. Randomization techniques and concealment .............................................. 28
   v. Blinding ....................................................................................................... 29
Chapter 3: Study Methods

3.1 Study Design

3.2 Population, Sampling, and Recruitment

3.3 Randomization and Concealment Techniques

3.4 Subject Protection and Confidentiality

3.5 Treatment Arms and Outcome Measures

3.6 Blinding

3.7 Adherence

3.8 Data Collection

3.9 Sample Size Calculation and Institution Estimate

3.10 Analysis

3.11 Timeline and Resources

Chapter 4: Conclusion

4.1 Summary of the Literature

4.2 Summary of the Proposal

4.3 Advantages of the Proposed Study

4.4 Disadvantages of the Proposed Study

4.5 Clinical Significance

4.6 References

Appendices

Appendix A: IRB Consent Form

Appendix B: Recommended Participating Institutions

Appendix C: Sample Size Calculation

Appendix D: Patient Questionnaire

Appendix E: Tables

Bibliography
List of Tables

Table 1: SRS-Schwab Classification.................................................................5, 55
Table 2: Gertzbein and Robbins Classification of Pedicle Screw Accuracy........55
Table 3: Selection Criteria...........................................................................56
Table 4: Participant Baseline Characteristics with Method of Analysis..........57
Table 5: Primary Outcomes with Method of Analysis....................................57
Table 6: Secondary Outcomes with Method of Analysis...............................57
Abstract

Adult degenerative scoliosis is characterized by the side-to-side curvature of the spine due to the degeneration of facet joints and intervertebral discs. Life-limiting cases of adult degenerative scoliosis are corrected by spinal fusion, the success of which is determined by pedicle screw accuracy. Robot-assisted guidance has become more commonly used in spinal surgery to ensure pedicle screw placement. However, there is limited evidence for the role of robot-assisted guidance in the correction of adult degenerative scoliosis. Here, in this randomized controlled trial, we aim to investigate whether robot-assisted guidance can more consistently produce accurate pedicle screw placement than neuronavigation. One hundred and fifty participants of specific selection criteria will undergo spinal fusion with the guidance of robot assistance or neuronavigation with postoperative computerized tomography scans to measure the pedicle screw accuracy. This study may have a widespread, direct impact in supporting the use of robotics to correct degenerative scoliosis.
Chapter 1: Introduction
1.1 Background
i. Low back pain

Low back pain is one of the most common complaints amongst patients who visit their primary care physician. Pain and/or stiffness to the lower back may be accompanied by pain radiating down to the legs, thus presenting itself as a debilitating pain that decreases the quality of life of individuals. According to a Canadian study in 1998 interviewing 20-85 year old adults, low back pain is experienced by about 85% of the subjects sometime throughout their lifetime\(^1\) and varied in its intensity, resulting disability, and its chronicity. A global meta-analysis conducted from 1980 to 2009 concluded that, within one month, 23% of people experienced activity-limiting back pain\(^2\).

To diagnose lower back pain, providers need to understand the extensive medical history of their patients as well as conduct thorough physical exams. The subjective exam should include questioning on the severity, intensity, nature, and duration of their symptoms. The physical exam should rule out anatomic abnormalities such as scoliosis, vertebral tenderness, reflex abnormalities, and reproducible radicular pain. These findings will begin to form the differential diagnosis from the above categories. Additional objective information such as laboratory studies and imaging will help confirm certain diagnoses. Laboratory studies do not have much utility in diagnosing low back pain but do hold impact in cases where nonmechanical spine disease etiologies are suspected\(^3,4\). In a primary care setting, imaging is not usually warranted as most cases are nonspecific back pain and imaging of this could lead to irrelevant findings with subsequent
unnecessary treatments\(^5\). On the other hand, patients with “red flag symptoms” such as severe neurologic deficit, saddle anesthesia, incontinence, immunocompromised, injection drug use, constitutional symptoms, or trauma require emergent imaging. For these particular physical exams, a lumbar spine MRI is the best modality for diagnosing specific lower back pain\(^1\).

The etiology of lower back pain is rather vast. However, according to a study by Chou et al, approximately 85% of lower back pain patients who present to their primary care physicians will be diagnosed with nonspecific back pain\(^6\). Nonspecific low back pain is defined as an absence of an underlying condition that can be identified to be producing one’s symptoms\(^6\). The other 15% of patients with low back pain can be categorized as mechanical low back pain, nonmechanical spine disease, or visceral disease\(^7\). Mechanical low back pain includes muscular strain, degenerative diseases, herniated disc, spinal stenosis, fractures, and congenital diseases. Non mechanical spine disease includes neoplastic, infectious, and inflammatory etiologies. Lastly, low back pain from visceral disease could be conditions that arise from pelvic organs such as prostatitis, endometriosis, or chronic pelvic inflammatory disease. Renal disease, gastrointestinal disease, an aortic aneurysm and fat herniation of lumbar spine are also examples of visceral causes of low back pain\(^7\).

\textit{ii. Adult spinal deformities: Etiology, pathophysiology, and clinical presentation}

Of the patients that present with mechanical low back pain, only a small subset of these cases are due to degenerative changes. Degenerative disc disease is defined by which the spinal disk material becomes eroded and displaced into the spinal canal\(^8\). As the disc wears away, the bones become closer and eventually rub together. This
Degeneration of the vertebrae can cause pain and lead to complications such as spinal deformities, herniated disk, spinal stenosis, and spondylolisthesis. For the sake of this discussion, adult spinal deformities will be reviewed at length.

As average life expectancy and access to healthcare increases, prevalence of adult spinal deformities amongst the elderly population will also be expected to increase\textsuperscript{9-11}. Adult spinal deformities are defined as the spinal curvature or alignment that deviates from the normal limits due to degenerative changes\textsuperscript{12}. It is a heterogeneous group that is comprised of different spinal curves in the sagittal and coronal planes,\textsuperscript{9} known as kyphosis, lordosis, and scoliosis, and vertebral body slippage known as spondylolisthesis. These conditions arise from similar etiologies. Kyphosis can be progressive, such as postural kyphosis, congenital, or developmental such as in Scheuermann kyphosis\textsuperscript{13}. In fact, one of the leading causes for hyperlordosis is poor posture. However, on the other hand, adult scoliosis can be “de novo,” meaning it can be derived from primary degenerative changes or due to the progression of adolescent idiopathic scoliosis\textsuperscript{14}. Adult spinal deformities in general have multiple age-related factors that lead to these degenerative changes such as osteoporosis (or reduced bone mineral density), spinal degeneration, reduced mobility and balance, and neurodegenerative disorders\textsuperscript{11}. These degenerative changes cause a loss in the intervertebral disc height resulting in pathological changes of vertebral and facet joints. There tends to be an increase in load on the anterior aspect of the vertebral joint that may cause arthritic changes to the posterior aspect ultimately leading to instability and bone remodeling\textsuperscript{10,15}.

The clinical presentations of the aforementioned spinal deformities can include back pain, myotomal, dermatomal or reflex deficits, and neurogenic claudication.
However, the physical presentation of the curve shape can vary. Spondylolisthesis can present itself as chronic low back pain with associated gait problems known as a “waddling gait”, with a palpable step-off sign in advanced stages\textsuperscript{16}. Kyphosis is characterized as a rounded back with increased forward posture of the head\textsuperscript{13} (also known colloquially as “humpback”). Patients with lumbar lordosis have exaggerated lumbar curves anteriorly and are unable to lay their back flat on the ground (also known as “swayback”). Lastly, those with scoliosis may notice changes in posture, balance, or loss of height accompanied by chronic back pain. The Adam’s forward bend test allows a provider to visualize a scoliosis curve from behind the patient where kyphosis and lordosis can be observed when viewing the patient in standing from the side\textsuperscript{17}. In severe cases of all curves, patients could experience chest pain, shortness of breath, and bowel/bladder incontinence due to spinal cord compression\textsuperscript{13}. There can be combinations of curves in presentation due to compensations. One common presentation of this is known as flatback syndrome\textsuperscript{11} which occurs when a sagittal plane deformity results in a compensated loss of lumbar lordosis. With an understanding of low back pain and various adult spinal deformities, the next section discusses the specifics of adult degenerative scoliosis.

\textit{iii. Adult degenerative scoliosis}

Adult degenerative scoliosis (ADS), also known as “de novo” scoliosis, is a common orthopedic complaint in individuals over 50. The prevalence of asymptomatic cases could be up to 65\% in individuals over 60 years old\textsuperscript{18}. In another study, the prevalence was consistently reported to be as high as 68\% of people with a mean age of 70.5\textsuperscript{19}. ADS can be classified according to etiology as Type I (primary degeneration
scoliosis), Type II (progressive adolescent idiopathic scoliosis), Type IIIA (lumbosacral scoliosis caudal to a previous spine fusion), or Type IIIB (secondary to metabolic bone disease). It is caused by asymmetric degeneration of facet joints and intervertebral discs leading to asymmetric loading, thus bone remodeling into a spinal curve. In many cases, patients with ADS may be asymptomatic. 90% of ADS patients commonly experience chronic low back pain and, as mentioned in previous section, can complain of postural imbalance, decreased height, neurologic claudication, and radiculopathy, depending on severity. ADS is typically diagnosed in individuals over 50 years old, if there is no prior history of idiopathic scoliosis, and when the coronal cobb measurement is at least 10 degrees in a radiograph or MRI, if indicated.

ADS is a dynamic disease with many factors contributing to severity and prognosis. Thus, treatment on ADS depends on its etiology, severity, and its corresponding limitations. The SRS-Schwab classification is the current standard for understanding the severity of the lumbar/thoracic deformity.

![Table 1. SRS-Schwab Classification](image)

Table 1. SRS-Schwab Classification. Describes the 4 coronal curve types and the sagittal modifiers of that include pelvic incidence (PI) minus lumbar lordosis (LL), global alignment, and pelvic tilt. “0” indicates normal, “+” indicates moderate modifier, and “++” indicates marked modifier.
A mild case of ADS, in patients without neurologic complications and stable curves, is defined as less than one degree per year and could be managed conservatively with physical therapy, analgesics, weight loss, and bracing\textsuperscript{26-28}. The goal of conservative management is to maintain quality of life and to live with the curve, rather than fix it. On the other hand, a more severe case is one that causes life-limiting symptoms and may require surgical intervention\textsuperscript{28}. Surgical correction of ADS is a major surgery that has specific qualifiers for patients to be candidates for the procedure. Patients typically qualify when the Cobb angle is greater than 30 degrees and/or greater than 6 mm of lateralolisthesis. Additional qualifications include experiencing peripheral neuropathy pain, such as intermittent leg claudication, and bowel or bladder symptoms if severe enough to compress the spinal cord\textsuperscript{18}. The current surgical intervention for ADS is decompression and fusion that can span multiple vertebral levels depending on the extent of the curve.

\textit{iv. Lumbar interbody fusion approaches}

Spinal surgery for ADS correction involves lumbar interbody fusion as the most used treatment. Different types of interbody fusion vary by their anatomical approaches which can be posteriorly, transforaminal, anteriorly, and lateral decubitus. Posterior lumbar interbody fusion (PLIF) approach is characterized by a midline incision directly over the spine. This approach avoids neurovascular injury from an anterior approach and allows for bilateral fusion with a single incision\textsuperscript{29}. Transforaminal lumbar interbody fusion (TLIF) is characterized by the transforaminal window that is made on one side of the vertebra and allows for less retraction of neural structures than PLIF\textsuperscript{29}. Anterior lumbar interbody fusion (ALIF) has obvious drawbacks as the approach is through the
abdomen but is a viable option for revision of PLIF surgery if needed. In a lateral decubitus lumbar interbody fusion (LLIF), like the ALIF, the surgeon accesses the spine via the retroperitoneal space.

Along with these approaches, the surgery can also be minimally invasive (MIS) or open surgery. There is a challenging and significant learning curve with MIS approaches. However, MIS approaches have been shown to decrease blood loss, operative time, and length of hospital stay with similar fusion failure rates as open surgery. In addition, a study conducted by Mobbs et al., has inconclusively debated which lumbar interbody fusion is superior for treating degenerative scoliosis while Kim et al has conducted a meta-analysis and reported most practitioners use PLIF. It is noted that idiopathic scoliosis has an adequate 10 year follow up from an anterior approach, but it did not translate to degenerative scoliosis. One study noted a combination of TLIF followed by PLIF had “reasonable” long term outcomes.

Regardless of the interbody fusion approach, the fixation will only be as strong as its anchor, the pedicle screw. These screws are typically made of titanium that are placed into the vertebral stem adjacent to the vertebral body, known as the pedicle. The screw’s placement is vital to the outcome of the procedure and will be the main interest of this study.

v. Pre-operative and intra-operative guidance of pedicle screw placement

Due to the vital nature of the placement of pedicle screws, the technique to ensure its accuracy is always advancing. Health teams have typically used pre-operative CT scans along with intra-operative fluoroscopy, such as the C-arm, to ensure the accuracy of the pedicle screw placement before drilling. This method has progressed and improved
with the implementation of neuronavigation, or brand name of O-arm based 3D image acquisition\textsuperscript{34}. One study concluded that the use of neuronavigation in cervical pedicle screw placement is highly reliable with only a 7\% breach rate\textsuperscript{34}. However, this current standard of care exposes the patient and the health team to large amounts of radiation throughout the case, creating an issue of long-term health safety for all involved. New technological advances such as robotics in surgery, specifically robot-assisted guidance, have presented a solution to help decrease the amount of radiation exposure. Operators are now able to upload a pre-operative CT scan and program the robot to guide placement of pedicle screws with minimal intraoperative imaging. It should be investigated if this technology is effective and safe, and whether it maintains consistent and accurate placement of the pedicle screw during spinal fusion cases for patients with ADS.

\textit{vi. Confirmation of pedicle screw placement}

Once the procedure is finished, patients will obtain another CT scan to assess the pedicle screw placement. The most widely used classification system for pedicle screw accuracy is Gertzbein-Robbins classification: Grade A indicates the screw is completely retained in the pedicle; Grade B, screw has cortical penetration but <2mm; Grade C, screw has cortical penetration more than 2mm but less than 4mm; Grade D, screw has cortical penetration more than 4mm but less than 6mm; Grade E, screw was cortical penetration of more than 6mm\textsuperscript{35,36}. To interpret this classification, Grade A is considered the optimal screw placement where Grades A and B are clinically acceptable screw position\textsuperscript{35} (see Appendix D, Table 2).
vii. Radiation exposure

Radiation exposure from neuronavigation presents potential long term health risks to the surgical team due to repeated exposures. These health risks can range from extreme cases of carcinogenesis to more minimal effects of hair loss\textsuperscript{37}. The average person living in the United States will be exposed to 3 mSv of radiation per year. A previous study determined the mean radiation dose from intraoperative neuronavigation during spine surgery as 5.15 mSv, with a range from 1.48 to 7.64 mSv\textsuperscript{38}. Ahmed et al combined the results of several studies and concluded that their patients were on average exposed to 8 mSv of radiation in total, with the surgeon only being exposed to 3 orders of magnitude of radiation lower of 11 mSv. The clinical yearly radiation exposure for surgeons is capped at 20 mSv\textsuperscript{39}. For robot-assisted guidance, multiple studies concluded significantly less radiation exposure but were compared to an older standard of care, fluroscopy\textsuperscript{40,41}. Further research needs to be conducted on radiation exposure during robot-assisted versus neuronavigation-guided procedures.

1.2 Statement of the Problem

As robotics are becoming more popular in medicine – and specifically surgery – there are questions regarding whether they should be adopted in certain practices and when would be the most appropriate times to incorporate them. To investigate, literature studies discussing the use of robotics in spinal surgery were reviewed.

There are many studies that broadly investigate various spinal pathologies and different combinations of control groups versus robot-assisted groups. In fact, some early studies have reported contradictory findings. In 2012, one prospective randomized trial concluded that conventional freehand (FH) technique was superior to the robot-assisted
technique in the accuracy of the placement of lumbar and sacral pedicle screws. As robotics were first introduced to spinal surgery in 2004, 8 years earlier to this first study, there have been significant advances in the technology as well as the providers’ competency with the systems. Newer retrospective studies investigating robot-assisted versus FH, for pedicle screw placement in various spinal pathologies, have pointed to robots offering similar screw placement to FH groups. One systematic review in 2020 had variable results on accuracy of pedicle screw placement based on which robot system was used (TiRobot, SpineAssist, Renaissance). However, very recently, a multi-center study followed the outcomes of robot-assisted spinal surgery over 5 years and validated the widespread use of robotics in spinal surgery. In the past decade, the trajectory of these studies seems to point to a more prevalent role of robotics in surgery. However, these poor study designs and broad inclusion criteria leave the results minimally applicable clinically.

With this newly widespread use of robotics for spinal surgery, it became of interest to investigate specific pathologies and if robot-assisted guidance performs better in certain situations. For example, spinal fusions to fix degenerative listhesis and central stenosis was studied in a randomized controlled trial. The researchers concluded that robot-assisted guidance was adequate during early application in 2015. When it came to spondylolisthesis in the lumbar spine, a prospective randomized controlled trial produced positive outcomes for the robot-assisted group. Additionally, when it comes to various degenerative and traumatic pathologies of the thoracolumbar spine of all age groups, pedicle screw placement is more accurate in robot-assisted versus fluoroscopy guided, as per a 2019 randomized controlled trial.
With more concrete evidence that robot-assisted guidance seems to produce similar success rates when compared to fluoroscopy/freehand techniques in simpler pathologies, it became more pertinent to explore whether robot-assisted guidance would have the same outcomes in more dynamic and complex processes such as adult degenerative scoliosis (ADS). A retrospective study determined robot-assisted guidance had a higher proportion of accurate pedicle screw placement than an FH group, but reported that a prospective randomized control trial, with a larger sample size, is warranted to increase the validity of their results\textsuperscript{48}.

Another retrospective study investigated robot-assisted guidance versus FH technique in three different pathologies such as spondylolisthesis, degenerative disc disease, and degenerative scoliosis. They did not have a significant difference in their results and concluded that robot-assisted guidance is similar to the FH technique\textsuperscript{49}. They also noted their limitation of a small sample size and low clinical application due to a single-surgeon study\textsuperscript{49}. Fan et al retrospectively investigated a similar question in patients with ADS which reported positive outcomes for the robot-assisted guided group\textsuperscript{50}. Finally, there has only been one study investigating the same question of interest with same treatment and control groups. A retrospective study by Mao et al concluded that there was no significant difference between robot-assisted guidance and neuronavigation. They did have multiple limitations and strongly urge for better study designs to investigate this topic.

In summary, studies on robot-assisted guidance and neuronavigation over the years varied significantly in their results as technology and provider proficiency increased. Over the last decade, there has been a trend that robot-assisted guidance may
offer similar accuracy as fluoroscopy. When it comes to ADS, there has not been a randomized controlled trial that has studied the outcome of pedicle screw placement accuracy specifically between robot-assisted guidance and fluoroscopy. Additionally, with neuronavigation proving to be more accurate than fluoroscopy alone, robot-assisted guidance needs to be studied against this new gold standard. Therefore, we aim to determine the efficacy and safety of robotic-assisted guidance for pedicle screw placement of complex spinal pathologies, such as adult degenerative scoliosis, using a randomized control trial with a large sample size against an appropriate control group.

1.3 Goals and Objectives

We aim to test the efficacy and safety of robotics in spine surgery, specifically in spinal fusion, to correct adult degenerative scoliosis. In this study, we will investigate whether robot-assisted spinal fusion surgery is more accurate than the current standard of care, neuronavigation, while reducing the amount of radiation exposure. To achieve this goal, the first objective is to provide strict inclusion criteria to make the participants’ pathology a more homogenous set to reduce possible confounding variables while obtaining a large enough sample size. Secondly, a study design of a prospective randomized control trial will be novel for this topic and will strengthen the validity of our study to have widespread clinical application. Lastly, our study will only be as good as the surgical centers and surgeons performing each case. Therefore, it will be vital to partner with proficient centers and providers that will be screened by strict qualification criteria. If robot-assisted guidance is proven to be superior to neuronavigation, this study presented here may offer strong clinical implications for correction of complex degenerative spinal scoliosis, proving this to be the new gold standard in such cases.
1.4 Hypothesis

A multicenter, double-blinded, randomized controlled trial will be conducted to investigate how the accuracy of pedicle screw placement differs from robot-assisted (Mazor X Stealth Edition) guidance to neuronavigation (O-arm) in patients undergoing spinal fusion to correct adult degenerative scoliosis. To answer this question, we propose the following hypothesis: there is a statistically significant difference in the proportion of Grade A screw placement between robot-assisted guidance versus neuronavigation in spinal fusion cases to correct adult degenerative scoliosis in adults between 45-85 years old.

1.5 Definitions

Pedicle: Short, thick, cylindrical bony processes that project posteriorly from the superior part of the vertebral body and fuse with the laminae to form the neural arch, each vertebra has two pedicles

C-arm navigation: Mobile imaging unit used primarily for 2-dimensional fluoroscopic imaging during surgical and orthopedic procedures

Neuronavigation: Movable CT imaging structure developed for intraoperative 3D fluoroscopic imaging, typically referenced as O-arm

Spinal fusion: Surgery to permanently connect two or more vertebrae in spine eliminating motion between them

Mazor X Stealth Edition: A robotic guidance platform produced by Medtronic that allows for pre-operative and intra-operative planning

1.6 References


32. Mobbs RJ, Phan K, Malham G, Seex K, Rao PJ. Lumbar interbody fusion: techniques, indications and comparison of interbody fusion options including PLIF,


Chapter 2: Review of the Literature
2.1 Introduction

This chapter is an extensive literature review and critical analysis of similar studies exploring the differences between robot-assisted guidance and other modalities in spinal surgery. Initial research efforts were conducted on smaller platforms such as Cochrane during the summer of 2022. These initial broad searches included variations of keywords such as: “spine surgery”, “robot”, “robotics”, “robot-assisted”, “decompression”, “fixation”, “lumbar spine”, “thoracic spine”, “cervical spine”, “fluoroscopy”, which revealed many articles of interest. These articles demonstrated the advancement of technology in surgery over the last 20 years and how robot-assisted guidance has been becoming more popular in spinal surgery, proposing a general topic for our study. After this initial set of searches, there were results for articles investigating robot-assisted guidance and their accuracy on pedicle screw placement in the cervical spine\textsuperscript{51,52}.

To further investigate whether and how robotics-assisted guidance impacts pedicle screw placement accuracy, the search was then moved to bigger databases including Embase, Ovid, and PubMed in the fall of 2022. These databases with many combinations of keyword searches led to most of the literature discussed in Chapter 1 and will be discussed further in this chapter. Searches with more specific keywords such as “pedicle screw accuracy”, and “Mazor” resulted in recently published articles including a few prospective randomized controlled trials. These were either broad in their spinal pathologies, such as a study done by Han et al\textsuperscript{47} and others\textsuperscript{42,46,49,52,53,45,54}, or were using out-of-date comparison groups such as one by Kim et al\textsuperscript{31}. Even narrower searches using
the keyword “scoliosis” for various complex spinal pathologies resulted in a single adolescent population study\textsuperscript{55}. However, in adding a more specific key phrase “adult degenerative scoliosis” (ADS) in addition to the above keywords led to articles centered on retrospective study designs, including study done by Mao et al, which investigated the same question as our proposed study but has major limitations. Fan et al also investigated a similar hypothesis but is not a prospective randomized trial, thus having major limitations to clinical application and calling for more research on this topic\textsuperscript{50}.

The literature search was finally completed using the Scopus database during the winter of 2022 into spring of 2023. This revealed several more pertinent retrospective studies and several meta-analyses. In total, there are eight randomized controlled trials and two retrospective studies that will establish the foundation of this study. All these articles will be dissected in the following subcategories to provide reason for the study at hand and the foundation for Chapter 3. Many articles will be referenced throughout the literature review, but a handful of articles will be heavily scrutinized.

\textbf{2.2 Empirical Studies Review: Pedicle Screw Accuracy by Robot-Assisted Guidance and Secondary Outcomes}

Throughout this section, results from previous literature will be discussed with corresponding statistical significance and real data. Despite the wealth of research done on the pathogenesis, diagnosis, and treatment on ADS, research on robot-assisted intervention for ADS is lacking. This section provides a critical analysis of these studies categorized by the same control group as proposed.

\textit{i. Analysis of pedicle screw placement accuracy from studies using similar treatment arms}
This section will discuss studies that analyzed pedicle screw accuracy with similar treatment arms from the proposed study, robot-assisted guidance versus neuronavigation. From 2009 to 2016, Fan et al. conducted a retrospective three-arm study investigating pedicle screw accuracy by robot-assisted (Renaissance™), drill guide template-assisted, and Computed Tomography (CT) based guidance\textsuperscript{50}. The primary outcome of interest in the study was pedicle screw placement accuracy measured via the Gertzbein-Robbins classification system (discussed in Chapter 1.1.vi). The authors concluded that the robot-assisted group had a statistically significantly higher proportion of grade A (91.3\%, \( p < 0.001 \), \( p < 0.001 \)) and grade B (4.7\%, \( p < 0.001 \), \( p < 0.001 \)) pedicle screw placements than the drill guide template (81.3\%, 9.4\% respectively) and CT-based guidance groups (84.1\%, 8.9\% respectively)\textsuperscript{50}. Secondary results include, but are not limited to, intra-operation radiation dose and blood loss. Robot-assisted group emitted 0.41±0.39mSv intraoperative radiation which was significantly less than what CT-based guidance emitted (5.68 ± 2.66, \( p < 0.05 \)). Lastly, there was no difference in blood loss between the three groups\textsuperscript{50}. This study suggested that robot-assisted guidance had increased accuracy and decreased radiation doses compared to CT-guidance and drill guide template.

From years 2017 to 2019, Mao et al. conducted a similarly controlled trial that investigated the same outcome variable between the same predictor variables, robot-assisted (Mazor X robotics) guidance against neuronavigation (O-arm)\textsuperscript{56}. This study has some major limitations. First, their retrospective study design leads to possible weakened internal and external validity and opens the door to confounding variables. Additionally, their study was not blinded to the authors limiting their clinical implications. Finally, the inclusion of various spinal pathologies created a heterogeneous population which
introduces much room for confounding. With that said, the robot-assisted guidance group had 97.5% of screws deemed clinically acceptable whereas the neuronavigation group had 94.81%, which was not significantly different with a $p=0.076$\textsuperscript{56}. However, they did conclude that robot-assisted guidance had a statistically significantly higher proportion of grade A placements compared to neuronavigation (86.16%, 65.99% respectively, $p < 0.0001$)\textsuperscript{56}. One secondary outcome this study measured was operative times, which proved to not be significantly different\textsuperscript{56}. Similarly, to the study done by Fan et al, results from the study conducted by Mao et al. suggest that robot-assisted guidance may offer higher rates of accurate pedicle screw placement compared to neuronavigation. These two studies demonstrate how robot-assisted guidance may offer increased accuracy compared to neuronavigation in various spinal pathology indications, but not in ADS specifically.

\textit{ii. Analysis of pedicle screw placement in correction of ADS with varying control groups}

In this section, studies specific for ADS testing for pedicle screw placement accuracy but with differing control groups will be analyzed. Chen et al investigated the same primary outcome of pedicle screw placement accuracy\textsuperscript{48}. However, this study differed in their comparison group of freehand technique versus the robot-assisted group. This article is quite significant as it only included patients with a diagnosis of degenerative scoliosis undergoing surgical correction, a homogenous population in terms of pathology, opposed to studies with multiple pathologies. During procedures from 2018 to 2019, the study resulted in a higher proportion of pedicle screw placement accuracy in the freehand group (98.7%) as opposed to the robot-assisted group (92.2%, $p<0.001$)\textsuperscript{48}. Intraoperative blood loss was measured, and the robot-assisted group (499mL) recorded
significantly less blood loss than the freehand group (499 mL, 573mL, respectively, \(p<0.001\)). In conclusion, the only study observing solely ADS patients for spinal fusion resulted in the freehand group providing a higher proportion of accurate placements than the robot-assisted group. This section serves to demonstrate the question at hand, does ADS cause more difficulty for robot-assisted guidance, and will it be better than a newer standard of care.

**iii. Analysis of pedicle screw placement in correction of adolescent scoliosis**

One study done by Macke et al, performed during 2012 to 2013, investigated pedicle screw accuracy in spinal fusion to correct adolescent idiopathic scoliosis. This study simply recorded the placement of 662 pedicle screws and measured their accuracy. They retrospectively concluded that proper use of robot-assisted guidance can improve accuracy in this population\(^{55}\). This study provides insight that robot-assisted guidance does offer increased accuracy for scoliosis, but of a different etiology.

**iv. Results from differing control groups and/or spinal pathologies**

Of the abovementioned studies, there are only four studies where: two have the same control groups of interest; one that investigates ADS; and the remaining study that discusses adolescent scoliosis. There is still a significant disconnect and lack of evidence of robot-assisted guidance for pedicle screw accuracy when compared to neuronavigation. Therefore, this section focuses on analyzing multiple studies, in chronological order based on publication date, to study the evolution of practice through the many variations of predictor groups and pathologies studied. This knowledge has created a strong foundation for development of robot-assisted instruments in spinal surgeries.
One of the first articles investigating robotics in spinal surgery was published in 2012. Ringel et al determined, in these early stages, that the freehand technique was superior to robot-assisted guidance when it came to pedicle screw placement in lumbar/sacral pathology (not significant, $p=0.019$). Similarly, a study by Kim et al studied pedicle screw accuracy with robot-assisted guidance versus freehand technique for patients undergoing posterior lumbar interbody fusion (PLIF) from 2013 to 2014. They reported no significant difference in the proportions of grade A and grade B placements between robot-assisted (99.4%) versus freehand technique (99.5%) ($p=0.534$). The lack of evidence for the benefits behind robot-assisted guidance in spinal surgeries continued through 2017 when Hyun et al conducted a randomized control trial amongst 60 patients with degenerative lumbar diseases. There was no significant difference in the Gertzbein-Robbin classification, but the fluoroscopic group had a shorter mean distance to facet, or closer to facet joint violation ($p<0.001$). Additionally, in that same year, another study by Malliqaj et al, investigated robot-assisted guidance versus fluoroscopic guidance for pedicle screw accuracy for multiple pathologies including tumors and traumatic indications. These authors concluded that the robot-assisted groups had a higher proportion of non-misplaced screws than the fluoroscopy group ($p=0.005$).

Over the next couple years, great advancements in the adoption of robotics in surgery were made, observed in the results from the following studies. In 2019, Han et al studied pedicle screw accuracy with robot-assisted guidance (TiRobot system) versus fluoroscopy in various thoracolumbar spinal pathologies. They concluded that the proportion of clinically accepted screws, or Grade A and Grade B, was significantly
higher in the robot-assisted group (98.7%) than in the fluoroscopy group (93.5%) \( (p < 0.01) \)\textsuperscript{47}. In 2020, robot-assisted screw insertion was proven more accurate than conventional fluoroscopy for cervical spine disease \( (p < 0.001) \)\textsuperscript{52}. Another group of clinical scientists observed facet joint violations in 2021, and demonstrated that the cohort of patients who received robot-assisted guidance had significantly decreased facet joint violations overall \( (p=0.0003) \), specifically in the upper thoracic spine \( (p=0.0209) \), and in lower thoracic spine \( (p=0.0455) \) but not in the lumbar spine \( (p=0.2340) \)\textsuperscript{59}.

At this point, robot-assisted guidance has increasingly been reported to be superior to fluoroscopy for pedicle screw placements. In 2021, a study by Cui et al quantified robot-assisted guidance pedicle screw accuracy in minimally invasive transforaminal lumbar interbody fusion (TLIF) to treat lumbar spondylolisthesis. This study concluded that robot-assisted guidance of screw placement led to more “grade A” placement compared to the open surgery group (P-value of 0.025) and significantly less intraoperative blood loss \( (p < 0.05) \)\textsuperscript{46}. Additionally, a very recent study in 2022 compared pedicle screw accuracy between robot-assisted guided versus that of fluoroscopy-guided cervical spine surgeries. Again, the robot-assisted guided group tallied significantly more “grade A” screw placements than the fluoroscopy-guided group\textsuperscript{51}.

v. Summary

In summary, it is important to note that only two results when using neuronavigation as the control group are not in agreement with each other. Additionally, of the studies looking into pedicle screw accuracy to fix ADS specifically, Mao et al concluded no difference in the proportion of Grade A and Grade B between treatment arms. Whereas Chen et al concluded that the freehand group had a higher proportion of
pedicle screw accuracy over the robot-assisted group. Thus, the question of pedicle screw accuracy by these two different modalities is still very much at large when it comes to ADS. However, of the remaining studies, five of them concluded that robot-assisted guidance had a higher proportion of accurate placements, two demonstrated no difference, and only one concluded that freehand was superior. Inconsistent conclusions across studies summarized here calls for the need for this type of study to definitively measure whether robot-assisted pedicle screw placements can indeed be the most reliable tool to treat adult degenerative scoliosis.

2.3 Study Methodology Review

From contradictory results reported in earlier studies that either support or not support the superiority of robot-assisted pedicle screw placement, each article will be evaluated to examine their validity for clinical applications. This section will scrutinize the study design of each study by discussing sample size, effect size, randomization techniques, blinding, selection criteria, statistical analyses, confounding, and loss to follow-up.

i. Study design

At this time, there were eleven pertinent prospective studies picked for the literature review. Of these eleven studies, eight were randomized controlled trials. These eight randomized controlled trials were very broad in their spinal pathology indication of study, and none specifically investigated adult degenerative scoliosis. However, these eight studies have the strongest study design with the most potential for meaningful clinical applications with the results listed in the previous
section. Other prospective studies that were not randomized\textsuperscript{49,51,53} allow for several issues such as having confounding factors, selection bias, and information bias.

There are several retrospective studies as well that are pertinent for the early steps of research but lack sufficient evidence supporting the specific use of robotics in ADS correction. In general, retrospective study designs have inherent bias that affects the study’s internal and external validity. All retrospective studies have predetermined exposures and outcomes which also raises a high possibility for selection bias of patients. Thus, the following studies’ results have weaker clinical implication. Fan et al researched a similar primary outcome of interest with three different treatment groups\textsuperscript{50}. Mao et al also investigated the same primary outcome of interest with the same treatment groups\textsuperscript{56}. Additionally, results from study done by Mao et al only support the association, not causation, between robot-assisted guidance and pedicle screw accuracy, which diminishes the use of this study for practical clinical applications. Their study does, however, offer substantial building blocks and evidence that additional randomized controlled trials, such as this study proposed, would improve on the vast limitations, and are needed to determine the efficacy of robot-assisted guidance in adult degenerative scoliosis.

\textit{ii. Number of centers, surgeons, and sample sizes of studies}

As will be discussed further in Chapter 3, the study being proposed will be a randomized controlled trial being conducted at multiple sites with multiple surgeons to produce a large sample size. The study with the largest sample size and the greatest number of surgeons included, is the retrospective study that investigated adult degenerative scoliosis conducted by Fan et al. They had five senior surgeons perform
surgeries for 267 patients. Since there were three control arms, 83 patients were in the robot-assisted group, 73 from the drill guide template group, and 109 in the CT-guided group. With enough surgeons at multiple sites, this sample size is achievable for a randomized controlled trial thereby strengthening the validity of their results.

As for the randomized controlled trials mentioned in the above section, there was not a single multi-centered trial. Six of them were single-center trials and two did not mention how many centers were used. Of the randomized controlled trials, despite only being a single-center trial, the study done by Han et al included patients seen by 5 surgeons, thereby generating a sample size of 234 patients – a relatively bigger inclusion population than other studies analyzed here – most amongst the RCTs analyzed. Their enrollment period was 21 months long, which is of a similar trial period length as proposed in our study. A different publication by Fan et al had the second biggest sample size of 127 subjects. However, they did not indicate the number of surgeons whose patients were included for the study. The remaining RCTs were single-center, single surgeon-performed studies with patient sample sizes ranging from 40 to 78. Each study claimed their results were limited due to a single center study design and called for a multi-centered trial. Additionally, studies with a smaller sample size also noted that as a limitation and urged future studies to figure out how to conduct with a larger population.

iii. Effect size and power

Mao et al will be the focus for its effect size which will help contribute to the sample size calculation for our study. Amongst 85 patients, there were 347 screws placed by neuronavigation and 318 by robot-assisted guidance. Of the 347 screws placed by neuronavigation, grade A classification was measured in 229 (65.99%). Of the 318 by
robot-assisted guided group, grade A classification was measured in 274 (86.16%) pedicle screw placements for robot-assisted guidance. This results in an effect size of approximately 20%. As for a different study, Fan et al did not include any further information on power nor their statistical significance. Additionally, Fan et al had a slightly larger effect size of 27% (87% robot-assisted vs 60% fluoroscopy) when only considering Grade A placement. Lastly, Su et al had an effect size of 20% for Grade A placement by robot-assisted guidance versus fluoroscopy in cervical spine fusions.

As for determining the most appropriate alpha value and power, we will focus on studies whose study designs were RCTs. Only three of the articles listed their power that estimated their sample size. Kim et al recorded b-value of 0.1 indicating a power of 90%, the highest amongst these RCT studies. Their alpha was 0.05 which concluded a need for at least 133 screws placed per group\textsuperscript{31}. Two other studies done by retrospective Fan et al and Ringel et al determined the need for 80% power to detect the outcome at a significance level of 0.05 in sample sizes of 106 and 60, respectively\textsuperscript{52,42}. Most of the studies had a commonality in their alpha values of 0.05\textsuperscript{42,46,47,57}. One RCT, Roser et al, failed to include their statistical analysis briefing in their article\textsuperscript{54}. As for the retrospective studies, Fan et al only listed an alpha value of 0.05 without denoting power or effect size. In conclusion, there is a call for a multi-centered RCT with a large sample size to increase power and significance of the results.

iv. Randomization techniques and concealment

Of the eight RCTs critically evaluated in this chapter, two of them have reported performing allocation to robot-assisted group or control group in a 1:1 fashion, but did not explicitly state their technique\textsuperscript{42,54}. The other six studies outlined their techniques for
patient group randomization, with five reporting allocation concealment as well. Cui et al did not mention concealment throughout their article\textsuperscript{46}. There were three different randomization techniques amongst them as there was some adoption of the technique by subsequent studies. Of those three techniques, three studies reported using an opaque envelope approach along with their randomization\textsuperscript{47,52,57}. The allocation sequence is performed via a computerized database which used blinded randomization blocks for their two intervention groups. This was concealed from the researchers. The participants were sorted with their assignments in sequentially numbered, opaque, sealed, and stapled envelopes. These envelopes were only opened after the enrolled patients completed all baseline assessments and were deemed ready to follow through with the intervention\textsuperscript{57}.

The second randomization technique was used in studies conducted by Kim et al. These researchers randomly assigned patients using a computer-generated randomization list. This list was concealed from the authors before the randomized allocation\textsuperscript{31,45}. Lastly, Cui et al utilized a distinct randomization technique. Patients were chosen consecutively and were assigned a group by a random number table. If it was an odd number, the patient belonged to the robot-assisted group. Even numbers were assigned to the comparison group. Despite the differences amongst these randomization techniques, they were very similar in their allocation stage. Only three of the eight studies analyzed in this section were thorough when describing how each subject’s allocation was concealed until the time of the intervention. This highlights the need for explicit instruction in the proposed study how randomization, allocation, and concealment will be performed.

\textit{v. Blinding}
Here, we will assess how each RCT study performed when it came to blinding at all stages of data collection. One inherent challenge of a surgical study is that the surgeons cannot be blinded as they know what intervention group they are performing. Additionally, there is a challenge to blind patients ethically. Therefore, none of these randomized control trials blinded participants nor the surgical team. With all said, outcome assessors were blinded in five studies\textsuperscript{31,42,45-47}, which was most clearly outlined by Han et al. Additionally, pedicle screw accuracy was measured during the postoperative phase of the study. This is assessed by a postoperative CT scan that is then read by a radiologist who is blinded to the treatment groups\textsuperscript{47}. This same technique was used in the retrospective study by Fan et al. They report two spine surgeons evaluated postoperative CT scans were blinded to the three intervention groups\textsuperscript{50}. The proposed study will perform similar blinding techniques to strengthen the internal validity.

\textit{vi. Selection criteria}

As noted in Chapter one, our first objective in this proposed study is to clearly outline the inclusion criteria and strict exclusion criteria to produce a homogenous population. The retrospective study by Fan et al studied adult degenerative scoliosis and their selection criteria will be most pertinent to our study. Fan et al included patients who experienced degenerative scoliosis that is superimposed on preexisting scoliosis, imbalance in coronal and sagittal planes, Cobb angle greater than 30 degrees, and obvious lower back pain and radiculopathy\textsuperscript{50}. However, they did not mention a specific age range in their inclusion criteria. As part of their exclusion criteria, patients diagnosed with a vertebral fracture, infection, tumor, discitis, or vertebral tuberculosis were turned
away. Subjects were also excluded if they have had previous spinal surgeries or incomplete data in their review case\(^{50}\).

Because most of these previous studies investigated different pathologies in the context of robot-assisted spinal surgeries, their inclusion criteria sets differ from one another. What is valuable is understanding these studies’ exclusion criteria, which may be pertinent to incorporate into our study. Across these prospective studies, researchers included patients within similar age groups and obtained signed consent. In regards to exclusion criteria, these clinical trials have turned away subjects who have been diagnosed with the following as they may present confounding factors: osteoporosis\(^{47,52}\), old fractures\(^{52}\), previous spine surgery\(^{46,54}\), severe pedicle deformity\(^{47,51,52}\), severe systemic disease\(^{31,45,47,52}\), coagulation disorder\(^{47,52}\), cancer or spinal tumor\(^{31,45,46,49}\), spinal infection\(^{46,49}\), history of peripheral vascular disease\(^{31,45}\), and those who were unable to complete the initial questionnaire\(^{31,45,52}\). Of note, several of the RCTs including studies by Hyun et al and Ringel at al did not discuss exclusion criteria, thus giving way to confounding variables and weakening their results\(^{42,57}\).

vii. Outcome operationalization and statistical analysis

The primary outcome of each RCT is to quantify pedicle screw placement accuracy. This measurement is done during the post-operative phase when the post-op CT is read and, if the screw is malpositioned, the distance of the screw tip to the cortical bone where the screw breached is calculated in mm. This is then categorized by the Gertzbein and Robbins scale. Each RCT reports this data as a set of proportions. There is a proportion of all screws that are Grades A to E. Since the data are categorical, most of the studies proceeded with Chi-Square analysis\(^{31,42,46,47,52}\) and Fisher Exact tests\(^{57}\).
Retrospective study by Fan et al. measured categorical data via an ANOVA test\textsuperscript{50}. The use of the ANOVA test in this instance raises suspicion about the study’s results. They do have three independent variables; however, their primary outcome operationalization is presented as proportions. A suitable statistical test like a Chi-Square analysis, as used in most RCT studies, would be more appropriate for the data set of this proposed study.

Secondary variables included parametric continuous data such as blood loss, operative time, and radiation exposure. This was analyzed via an independent t-test (student t-test) across all studies analyzed in this paper\textsuperscript{31,46,47,52,57}. For non-parametric continuous data, studies used a Mann-Whitney U test\textsuperscript{52} or Wilcoxon test\textsuperscript{47}. Fan et al utilized an SNK test for secondary outcomes of blood loss and intra-op fluoroscopic dose\textsuperscript{50}. However, there is a flaw in this study’s use of the correct statistical analysis. The SNK test is a post-hoc test that follows a significant ANOVA test to determine the specific pairs of means that are different. It seems that they inappropriately ran an ANOVA for categorical data listed as proportions (as above) and again inappropriately ran this SNK test as their primary analysis for continuous data.

viii. Confounding variables

A previous multicenter study investigated the risk factors affecting pedicle screw accuracy for spinal fusions done with robot-assisted guidance\textsuperscript{60}. Researchers concluded that BMI, overweight and obese classifications, and female gender were associated with significantly more mispositioned screws\textsuperscript{60}. In fact, female gender is correlated to higher rates of osteoporosis which is known to have higher complications in spinal fusion procedures. This is due to weaker bone allowing for easier breach of the cortical wall and decreased holding strength post-op\textsuperscript{61}. 
The confounders of the Fan et al study will be discussed first. As mentioned in chapter 1, ADS is a very dynamic spinal pathology. The SRS-Schwab criteria outline the characteristics and severity of the disease, with some cases being significantly more difficult to treat and could lead to confounding results of the pedicle screw accuracy rate by treatment arm. This study does not explicitly exclude individuals with higher severity of curves such as those that extend to more proximal vertebrae or curves with corresponding sagittal deformities\(^{50}\). Without a homogeneous population amongst treatment groups, it is possible that some groups had “easier” cases than others thereby impacting the success rates for pedicle screw placement. This was observed in the robot-assisted group (7), which had significantly fewer screws placed within the thoracic region (T10 region) than the CT-guided group (24) (\(p = 0.011\))\(^{50}\). Thoracic pedicle screws tend to be the most challenging due to greater anatomical variation in thoracic vertebrae than that of lumbar vertebrae\(^{62}\). Additionally, Fan et al had no exclusions for BMI and concluded a statistically significant difference in baseline BMI between the robot-assisted group and CT-guided group. They also had a higher ratio of females to males (154:113). Despite no significant difference in the proportion of females between each group, females were the predominant gender in each group at 58% (robot-assisted), 55% (template guided), and 60% (CT guided). The study also contradicts itself noting that 5 senior surgeons were qualified to perform these cases in the methods section but also highlighting the fact that the cases were performed by over 40 different surgeons that included 4 different types of spinal surgeries in their discussion.

Additionally, study-limiting confounders will be discussed from the study by Mao et al. Similarly, to the study by Fan et al, there was no exclusion of BMI in their selection
criteria. In fact, the study does not mention BMI nor patient’s weight at all which can confound results if one group had a higher mean BMI than the other, due to longer distances and changes in angle of approach. Secondly, pathologies experienced by subjects included in this study was not homogeneous. Han et al included patients with degenerative, infectious, oncologic, and traumatic etiologies and the control group was not case matched to the treatment group. Therefore, some cases were inherently more difficult, or less likely to succeed, and could lead to worse outcomes. The study reports two confounding variables at the end of the article. They mention that several robotic cases were aborted and converted to neuronavigation due to technical issues, which causes some selection bias. Additionally, there was increased variability intra-operatively as the resident surgeon assisting the main surgeon was not consistent. Lastly, as mentioned previously in this chapter, the study design raises questions about selection bias and significantly limits the study.

Confounding variables noted above were well controlled for in the randomized controlled trials. These trials discussed their baseline characteristics and most had no difference in age, sex, BMI, and pathology which limits the greatest confounders. Han et al does mention they used a team of 5 experienced surgeons. However, they do not mention how they qualify the surgeon’s experience. Thus, each surgeon’s experience may be subjective creating variability in the proficiency of each surgeon. This may result in variability if the surgeon’s cases of each treatment group were not matched. This is certainly a confounder that could arise in our proposed multi-center study and certain objective measurements for surgeon proficiency must be noted to help limit such confounders.
ix. Loss to follow up

Since our study will have post-op results immediately after the procedure, we took note of the lack of reporting on follow-up risks amongst subjects across studies analyzed here. Only one RCT outlined how they lost several subjects in each treatment group but did follow up on patients who received pedicle screw placement surgery\textsuperscript{52}. In this study, 67 subjects were allocated to the robot-assisted group; however, only 61 of these subjects underwent treatment. Two of them had symptomatic improvement before the surgery and four denied the treatment group. In the fluoroscopy group, 68 subjects were allocated and only 66 of them underwent the procedure. In this group, one was lost due to symptomatic improvement and one due to denying the treatment group.

2.4 Conclusion

From our extensive literature review, there is insufficient evidence for determining whether robot-assisted guidance is superior to neuronavigation for pedicle screw placement accuracy, specifically in the correction of adult degenerative scoliosis. There exists two studies that measured this effect, Mao et al and Fan et al. However, both were greatly limited in their retrospective study design. These studies also have additional major confounding variables that diminish their application into clinical practice. On the other hand, several randomized control trials investigate similar questions, but none investigated neuronavigation (O-arm) specifically as their control group or studied other spinal pathologies. Additionally, several of these RCTs concluded that robot-assisted guidance offered superiority, however, were compared to an outdated control group and for different spinal pathologies. Nonetheless, this is evidence of the advancements made for robotics in spinal surgery and is motivating to test this novel
hypothesis. Learning from these studies, we aim to conduct a multi-centered randomized controlled trial to test the proficiency of robot-assisted guidance versus neuronavigation in spinal fusion cases treating adult degenerative scoliosis.

2.5 References
Chapter 3: Study Methods

This chapter will outline the structure and methodology of our study as it was justified by the literature summarized in Chapter 2. It will contain specific details such that individuals may replicate our study. Throughout each section of this chapter, the consort guidelines of surgical studies were considered and incorporated.

3.1 Study Design

This study will be designed as a prospective, multi-center, two-arm randomized control trial that will investigate the accuracy of pedicle screw placement between robot-assisted guidance (Mazor) and neuronavigation (O-arm) in spinal fusion cases to correct adult degenerative scoliosis. This study will only partner with surgeons who have had over 5 years of individual operative experience and must have performed each procedure at least 25 times. The type of lumbar interbody fusion approach will not be specified as it will be surgeon- and/or case-dependent. There will be measures taken intra-operatively to discern accuracy, thus making the variability between interbody fusion approaches negligible.

3.2 Population, Sampling, and Recruitment

The study population will be comprised of individuals from the ages 45-85 who meet the criteria for surgical intervention to alleviate symptoms due to adult degenerative scoliosis. To meet the sample size calculated below with such strict selection criteria, we will employ convenience sampling at each institution. Participants will be recruited to the study upon discussion with their healthcare teams at participating institutions, and have indicated they will proceed with surgical intervention to correct adult degenerative scoliosis. Participants who are interested and willing to participate will be selected for the
study and will be screened for the following selection criteria (Appendix E, Table 3). Specific dates of recruitment will be defined once the study is approved.

Inclusion criteria include: having diagnosed with adult degenerative scoliosis; experiencing life-limiting symptoms such as self-image; issues with posture; balance issues; leg claudication; and bowel and/or bladder issues. Additionally, participants must be able to sign a consent form written in English (official legal translators can be used, Appendix A).

Exclusion criteria are outlined as well. Specific SRS-Schwab classifications will be excluded in the study as follows. Since patients who experience lumbar curves with sagittal deformities have greater problems and disability,18 we will exclude patients with marked modifier sagittal deformities (Appendix D, Table 1) to increase the outcomes of the study. Additionally, the spinal location of the curve will be considered. To increase homogeneity of the population and decrease confounders, curves that require fusion proximally to vertebral level T10 will be excluded. However, there will be no exclusion for curves that require caudal intervention. Other exclusion criteria include: current infection, severe systemic disease, cancer, spinal tumor, prior spinal surgery, multiple pathologies of the adjacent vertebra, previous spinal fractures, severe pedicle deformity, coagulation disorder, BMI > 39.99, osteoporosis, decompression procedures alone, pregnancy, and if more than two of Charlson Comorbidities are present. Participants will also undergo routine pre-operative screening including electrocardiogram (EKG) vital signs, urine drug screen, COVID-19 PCR testing, and routine blood labs.
3.3 Randomization and Concealment Techniques

As patients are recruited into the study, they will be randomized into the control and treatment groups in a 1:1 fashion. Participants will be allocated into two groups using a computerized blinded randomization block that will randomly select each participant to be placed in either the robot-assisted group or neuronavigation group. This process will be concealed from the research group and participants. These results will be concealed in sequential envelopes as each participant is recruited, and will remain sealed until the patient is preparing to undergo the intervention. At that time, allocation will only be known to the lead surgeon.

3.4 Subject Protection and Confidentiality

Before the study is initiated and offered to participating institutions, this study will apply for approval from the Yale Institutional Review Board (IRB). This study will be subject to an IRB review since the university will be considered engaged in its research as the primary investigators are affiliated with Yale University. These are necessary measures as these implemented requirements will increase subject protection. Due to the multi-center nature of the study, we will refer to the section on investigator-initiated and multicenter research under the Yale University Human Research Protection Program (HRPP) policy 120. This outlines the steps that will need to be taken to complete this protocol.

This study proposal including the hypothesis of interest, proposed methods, and method of analysis along with evidence of funding, institutional agreement, and their training protocols will be presented to the HRPP and submitted to the IRB. This will serve as the central IRB submission since Yale New Haven Hospital will be the
coordinating center. This center will hold the responsibility for the conduct, administrative, and coordinating functions of multiple centers, known as the external institutions. Those external institutions will sign an IRB authorization agreement, follow the initial IRB submission, and honor it as the IRB Reliance or Yale New Haven Health as the relying institution.

The guidelines of the study will also follow IRB policy 200 to attain informed consent for human research. To satisfy these requirements, a statement will clearly and concisely explain the purpose of the proposed research, the expected duration of study, and details of each experimental procedure within the study. It will also outline any potential risks or discomforts and discuss the benefits that may be concluded from the research. There will be disclosures of alternative procedures that could be available to the participants and statements discussing how confidentiality will be maintained. Lastly, it will be known that the study is completely voluntary and may be discontinued at any time. Contact information of the lead investigators will be provided for any questions. All these requirements have been completed in the example informed consent form (see Appendix A).

Lastly, IRB policy 400 outlines the privacy and confidentiality of human research information. The study will take measures that will maintain the privacy and confidentiality of personal information accessed, created, used, or maintained during all phases of the study. The study will comply with HIPAA guidelines and take measures to secure private health information, or PHI. Only the minimum amount of information will be requested about each patient and will be limited to only the lead investigators. Their data, such as baseline characteristics, will be used to make sure they qualify for the study.
and that there is no difference between treatment groups. All data will be stored in locked devices that only the lead investigators can access. Once the study is completed and published, individual data will no longer be retained.

3.5 Treatment Arms and Outcome Measures

The treatment group of interest will be the robot-assisted group performed by Medtronic’s Mazor X Stealth Edition robotic guidance system. Procedure steps are followed by a study performed by Lieberman et al\textsuperscript{63}. In this group, participants will undergo a preoperative CT scan which will be uploaded to the robotic software. This software will be able to plan screw insertion points and trajectories. On the day of the surgery, patients will be appropriately prepped in a sterile fashion. As the procedure progresses, the robot will be fixated to the participant, typically via iliac spin or exposed spinous processes, depending on the approach of the surgery. The few intra-operative images needed during this technique will be taken at this time to orient the robot’s position based on the patient mount. At this stage, the robot will position its surgical arm on the patient as a guide, as it matches the intra-operative imaging from the pre-operative imaging and follows through with its planned programming. The surgeon will then use this trajectory to manually drill into the pedicle and insert the pedicle screw. Once all screws are placed, final imaging will be taken to verify screw positions.

The control group will be intra-operative neuronavigation with the use of O-arm navigation. Participants will be placed on a sterile field inside the O-arm device. Initially, the O-arm will be positioned over the vertebral level that will be operated on. A reference, such as StealthStation reference as described by the study by Park et al, will be drilled onto the patient at the iliac crest\textsuperscript{64}. At this point, the O-arm will obtain CT-type
imaging of the target level which will be produced into multiplanar images, meaning the navigation is ready. Once imaging is obtained, the O-arm is moved away to allow room for the surgeon. This imaging is then constructed by software into multiplanar images that create 3D navigation imaging on a screen. Once placed, the surgeon is ready to drill and place each pedicle screw. Immediate imaging to assess accuracy can be done by the O-arm.

The primary outcome of interest of the study will be the proportion of Grade A pedicle screw placement by the predictor group (robot-assisted guidance by Mazor X Stealth Edition guidance) versus the control group (neuronavigation by O-arm). This is determined by the Gertzbein-Robbins classification as discussed in Chapter 1 (also see Appendix E, Table 2).

One of the secondary outcomes of interest is the proportion of Grade A plus Grade B, known as “clinically acceptable,” pedicle screw placement by the predictor group (robot-assisted) versus the control group (neuronavigation). Other secondary outcomes that will be measured include operative time (min), blood loss (mL), radiation exposure time (seconds), and radiation exposure from the provider and patient (mSv).

3.6 Blinding

An inherent disadvantage of surgical studies is the difficulty to double-blind. The surgeon cannot be blinded because they will be the individual performing the procedure with certain tools of each treatment group. The study by Karanicolas et al discusses the difficulties of blinding surgical studies and possible ways to obtain the most blinding as possible while taking into account ethical considerations. Five groups of individuals
that need to be blinded are the participants, surgeons, data collectors, outcome adjudicators, and data analysts.

Data collectors, outcome adjudicators, and data analysts will be blinded. An independent radiologist will serve as the data collector to assess blinded post-op CT scans based on the Gertzbein-Robbins classification. The information will then be sorted and assessed by another blinded study investigator who will run the data analysis.

Difficulty in blinding lies in the participants and clinical teams. The procedures will be very similar and are both appropriate methods of treatment for this patient population. Variation only lies within the approach used for pedicle screw placement. Otherwise, incisions and post-op characteristics will be mostly similar. Patients will be blinded by not being told of their allocation pre- or post-operatively. This is an ethical design as each treatment group is an acceptable standard of care. As for the clinical team, the surgeon and the OR team will not be blinded as they will inherently know if a robotics system is being used or not. However, to best limit performance and ascertainment bias, those in pre- and post-op covering teams who are not present in the OR will not be informed of each participant’s treatment group.

3.7 Adherence

Primary and secondary outcomes of interest will be recorded either intra-operatively or in the immediate peri-operative time within the operating room (see below, Chapter 3.8. Data Collection). Thus, adherence will be very well controlled, and the study will have minimal loss to follow-up. Non-adherence to study may occur if a participant decides to leave the study prior to the allocated intervention. This was observed in the randomized control trial by Fan et al. These authors noted that six out of sixty-seven
subjects allocated to the robot-assisted group were lost due to symptomatic improvement before surgery or they did not accept their treatment group\textsuperscript{52}. Therefore, it may be appropriate to expect only a few participants to not follow up once enrolled and assigned a treatment group.

### 3.8 Data Collection

Data collection will begin with baseline characteristics. Each participant will fill out a survey that will ask for: age, sex, BMI, comorbidities (hypertension, diabetes mellitus, hypercholesteremia), pertinent past medical history, pertinent past surgical history, and expected number of vertebral levels to undergo fusion (see Appendix D). Once researchers obtain all baseline characteristics data, participants will be allocated to each treatment group which will be analyzed for homogeneity.

For pedicle screw placement accuracy, each participant will have a post-op CT scan. The scan will be electronically read by an independent blinded radiologist who will measure accuracy of pedicle screw placement and classify placement using the Gertzbein-Robbins classification system. The same procedure will be performed to assess the secondary outcome of interest of Grade A plus Grade B placement in both groups.

The secondary outcomes of interest will be measured intra-operatively by the surgical team. Blood loss is typically measured by a complete blood count lab draw and/or by weighing blood-soaked towels used during the procedure. The operative time is measured at standard in the OR and will be included in the data collection. Lastly, radiation exposure will be collected via personal radiation dosimeters that are placed on the surgeon and on the patient.
3.9 Sample Size Calculation and Institution Estimate

To calculate sample size, Power and Precision software version 4.0 was used. The effect size of Grade A placements between the neuronavigation group and the robot-assisted group was taken from multiple previous studies. In one retrospective study, Mao et al concluded that 86% of screws placed by robot assistance were Grade A, whereas only 66% of screws placed by O-arm navigation were Grade A\(^5\) (a difference of ~20\% in effect size). Additionally, Fan et al had a slightly larger effect size of 27\%, noting Grade A placement 87\% of the time by the robot-assisted group versus just 60\% of the time by fluoroscopy. Another study by Su et al also reported an effect size of 20\% for Grade A placement by robot-assisted guidance versus fluoroscopy – but for different spinal pathologies. From these studies, our study will aim to have an effect size of about 20\% for pedicle screw placements when performed by robot-assisted guidance versus neuronavigation groups.

A two-arm, chi-square analysis is selected, with an alpha value set to 0.05 and power of 80\%. Power and Precision software then prompted to enter the known values from previous literature, which computed a total sample size of 142 (Appendix C). Therefore, there should be 71 participants in each treatment group. Since the loss of follow-up remains to be a concern that would arise prior to intervention or denial of treatment, we plan to increase our sample size by 5\% (equivalent to 8 participants total, or 4 in each treatment group). Therefore, at least 75 participants per treatment arm are necessary for our study, yet more are welcome.

To obtain our sample size of 150 participants, there will have to be multiple centers enrolled in the study. Approximately 6.5 procedures per month will be conducted.
throughout this study period of 23.5 months. Given that each site will perform at least two procedures per month, surgeons will each perform approximately 47 cases in the 23.5-month period. Thus, we will need approximately 3.3 institution sites. Therefore, we will cautiously plan for participant over-enrollment and partner with 4 institutions with a few extra sites for contingency. There will be, however, a low threshold to increase the number of sites as the study is conducted if the participant amount is not being met. See Appendix B for the list of sites.

3.10 Analysis

Baseline characteristics of BMI, age, sex, comorbid conditions (hypertension, diabetes), and smoking will be measured accordingly. BMI will be measured as mean with standard deviation, to be analyzed via a student t-test across treatment groups. Age will be nonparametric data, to be presented as medians and interquartile ranges, analyzed via the Mann-Whitney U test. Sex and the comorbid conditions will be noted as proportions of the sample population and measured by a Chi-Square analysis (see Appendix E, Table 4).

As for the primary and secondary outcomes, similar tests will be used as for baseline characteristics. The primary outcome (proportion of Grade A placement per treatment group) will be analyzed for differences in proportions across groups, thus a Chi-Square analysis will be used to quantify significance (Appendix E, Table 5). The same applies to the secondary outcome of the proportion of Grade A plus Grade B placement between treatment groups. As for blood loss, operative time, and radiation exposure, all variables will be parametric continuous data, thus will be analyzed via a student t-test (Appendix E, Table 6).
3.11 Timeline and Resources

This study will begin enrollment as soon as approval from the Yale IRB is given. Since the study design partners with proficient surgeons and centers, there will only be a week-long orientation period for all personnel to understand their roles and the protocol. The study will begin enrolling, recruiting, and treating patients as soon as possible. The study period will occur over 24 months and recruitment will continue to occur until the latest date, which will be at 23.5 months. Data collection will be ongoing throughout the study with an emphasis during month 23, as the last participants are treated. There will be a two-week buffer at the end of the study period to gather all remaining data to organize and finalize the data to run the appropriate analyses.

Resources to conduct this study will be carefully distributed and allocation will be ensured across institutional sites. There will be a lead investigator per site that will oversee operations, with the lead and co-lead investigators at Yale New Haven Health in charge of the operation. There will also be one research assistant at each site to help organize the data monthly. As for the clinical team, due to proficiency qualifications, 1-2 surgeons per site will be expected along with an independent radiologist who can perform post-operative CT scan reads.

3.12 References


Chapter 4: Conclusion

In conclusion, this chapter will tie the preceding work and will highlight important aspects of the literature and proposed study. It will emphasize the advantages and disadvantages of its study design and the clinical significance of the study if executed correctly.

4.1 Summary of the Literature

Robotics were first introduced into spinal surgeries in the early 2000s, and since have advanced remarkably fast. Robot-assisted guidance now offers a more streamlined technique of performing spinal fusions while reducing potential risks of higher doses of radiation exposure emitted from other standards of care, such as neuronavigation; however, it is still of question whether robot-assisted guidance offers superior pedicle screw accuracy rates than neuronavigation. Through the literature analysis, there were many studies investigating the general use of pedicle screws during spinal fusion cases and how the accuracy varies with the use of several modalities being compared to robot-assisted guidance. These studies would vary in study design, pathology indication, and standard of care comparison. There was an overall trend since the early adoption of the technology that robot-assisted guidance continued to improve and would offer similar, if not better, pedicle screw accuracy than other modalities. There were several randomized control trials that investigated pedicle screw accuracy, but none compared neuronavigation (O-arm) specifically as their control group and some studied other spinal pathologies other than adult degenerative scoliosis. In search of literature on pedicle screw accuracy specifically in spinal fusion cases to correct adult degenerative scoliosis, it is unknown if robot-assisted guidance offers increased accuracy rates than
neuronavigation. Two studies, Mao et al and Fan et al, specifically investigated pedicle screw accuracy with adult degenerative scoliosis as the indication for the procedure, however, both were greatly limited in their retrospective study design. They additionally have major confounding variables that further diminish their application into clinical practice. It has been demonstrated that there is evidence for the general advancement of robot-assisted guidance during spinal fusion cases, but there is a lack of evidence if robot-assisted guidance can provide similar, if not better, pedicle screw accuracy rates in spinal fusion cases to correct adult degenerative scoliosis.

4.2 Summary of the Proposal

To address this gap in literature as summarized above, the following study has been proposed. The proposed study is designed to be a prospective, multi-center, two-arm randomized control trial that will investigate the accuracy of pedicle screw placement between robot-assisted guidance (Mazor) and neuronavigation (O-arm) to correct adult degenerative scoliosis. There will be 150 participants who will be comprised of individuals from the ages 45-85 who meet the criteria for surgical intervention to alleviate symptoms due to adult degenerative scoliosis. Strict inclusion and exclusion criteria are outlined to produce a homogenous population to limit confounding variables (See Table 3).

The primary outcome of interest of the study will be the proportion of Grade A, from the Gertzbein-Robbins classification, pedicle screw placement by the robot-assisted group versus neuronavigation group. Secondary outcomes of interest include Grade A plus Grade B placement, operative time, blood loss, radiation exposure to surgeon, and
radiation exposure to patient. The data will be analyzed by Chi-Square and student-t test analyses when appropriate as indicated in chapter 3.

4.3 Advantages of the Proposed Study

This study offers advantages when compared to the existing literature for robot-assisted guidance in the setting of adult degenerative scoliosis. First, the study design of a randomized controlled trial would be the only one performed to investigate pedicle screw accuracy between robot-assisted guidance versus neuronavigation in correction of adult degenerative scoliosis. The randomized controlled trial study design decreases the possibility of selection and information bias; thus, increasing the study’s internal and external validity.

Secondly, this study will have access to more resources compared to other literature investigating pedicle screw accuracy since the primary institution will be Yale New Haven Health. Most of the previous studies within the literature review were lacking in their sample size, especially the other randomized controlled trials. With access to more resources, this study would allow to be appropriately powered and be able to enroll many patients for the study. Also, proficient surgeons will be available and recruited to minimize the confounding of the robotics learning curve.

Additionally, the multi-center approach (See Appendix B for Centers) has some advantages. The advantages are that it could produce more generalizable results across spine centers within the United States. With at least 4 sites participating in the study, they could reinforce each other’s results producing more convincing evidence for widespread adoption.
Lastly, previous studies such as Cui et al., Kim et al., and Fan et al., used C-arm fluoroscopy as their control group\textsuperscript{45,46,52}. The newer technology that neuronavigation, such as O-arm, offers has been proven to be more reliable than C-arm when it comes to accuracy\textsuperscript{34}. Thus, our control group of neuronavigation will be more relevant with bigger implications including being more clinically relevant for more widespread use of results.

### 4.4. Disadvantages of the Proposed Study

Despite being a novel study with a strong study design, multi-centered offering a large sample size, and comparing to a more current standard of care, there are, however, several disadvantages of our proposed study. First, to ensure sufficient enrollment during our study period, the type of approach of lumbar interbody fusion will vary from surgeon to surgeon and perhaps from case to case. This will introduce variability into the study. With that said, this variability is inherent to surgical studies as each case is typically different even with the same surgeon due to anatomical differences from patient to patient. Therefore, we expect this disadvantage to not affect the outcomes greatly since each modality for accuracy will be used intra-operatively the same regardless of technique.

Additionally, the multi-center approach presents itself as a disadvantage. Different centers may have variation in their protocols and techniques in how to perform certain procedures. For example, some centers may train their surgeons to perform a posterior lumbar interbody fusion where other centers may favor the transforaminal lumbar interbody fusion approach. There may also be variations in the surgical skill despite strict criteria for participating surgeons.
4.5. Clinical Significance

Adult degenerative scoliosis is already a very common musculoskeletal disease in the older population. As the population of the United States grows and the percentage of elderly increases, adult degenerative scoliosis will become even more prevalent. Thus, it is imperative that the method of how it is corrected is sufficiently studied. This proposed study, with its design and well-powered sample size, will offer information that is otherwise lacking in the literature with widespread clinical application. This would prove to be beneficial for patients, as there could be better outcomes, and for providers, as they could limit their radiation exposure. The continued gain of technology in the workforce may shift employment opportunities in the future, but that is a discussion for another paper.

4.6. References
Appendices

Appendix A: IRB Consent Form

CONSENT FOR PARTICIPATION IN A RESEARCH PROJECT
YALE UNIVERSITY SCHOOL OF MEDICINE
Policy 200

Study Title: Spinal Fusion for Adult Degenerative Scoliosis: Robot-Assisted Versus O-arm Guidance
Principal Investigator: Peter Whang, MD, FACS, FAAOS

Invitation to Participate and Description of Study

We are pleased to invite you to participate in the research study designed to assess pedicle screw placement accuracy for those undergoing spinal fusion for adult degenerative scoliosis. You have been asked to participate because your surgical team denoted you as a patient with adult degenerative scoliosis who is a candidate for spinal fusion. The study will include at least 150 participants over a minimum 4 different sites.

In your decision making of deciding to be a part of this research study, you should know the 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, and possible benefits. 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

The surgical team has diagnosed you with life-limiting adult degenerative scoliosis deemed you a candidate for spinal fusion. If you decide to proceed with this study, you will be added to the computerized random block generator to allocate you to either the robot-assisted (Medtronic’s Mazor X Stealth Edition) or neuronavigation (O-arm) treatment group.

Baseline information will then be collected including age, sex, BMI, health comorbidities, past medical history. If you qualify under the selection criteria, you will then proceed to a routine pre-op exam to obtain an EKG, blood labs, urine tox screen, and COVID testing.

After all pre-operative testing, a procedure date will be scheduled by the surgical team. On that date, the surgeon will open your envelope revealing your assigned treatment group. The technique and approach of spinal fusion will vary by surgeon thus will be discussed specifically and in detail with your surgical team. As for the pedicle screw accuracy modality, if you are in the robot-assisted group, pre-operative imaging will be used by the
robot software. The surgical team will plan and program screw trajectory from the pre-operative imaging. During the procedure, the robot will be attached to your hip and will act as a guide for the surgeon’s drill. Several more intra-operative images will be needed during the technique. If you are allocated to the neuronavigation group, you will be prepped by the surgical team on the day of your procedure and the O-arm device will be placed around the operative table and a reference will be fixated also to the hip. A series of imaging will be taken until the reference is able to construct a 3D navigation on a monitor that the surgeon will use as the guide. In both groups, you will undergo post-operative imaging, most likely a CT scan, to measure the accuracy. Other intraoperative measurements will be taken such as blood loss, time of operation, and radiation exposure.

Once you wake up from the procedure, your participation in the study will be completed as all outcomes of interest will have been recorded.

A description of this study will be available on https://clinicaltrials.gov/, as required by U.S. Law. This site will not include information that can identify you. The purpose of this database is to provide information on what studies are currently being done and what studies have been done. At most, the site will include a summary of the results. You may search the site at any time.

You will be told of any significant new findings that are developed during your participation in this study that may affect your willingness to continue to participate. Research results will not be returned to your doctor. If research results are published, your name and other personal information will not be given.

Risks and Inconveniences

As one of our outcomes of interest is radiation exposure, you may have a slight difference in radiation exposure from one treatment group to another. If there is a difference in radiation, it will not be a large enough exposure to cause harm as both treatment groups are approved and are performed commonly with adequate results.

Benefits

There are no additional benefits associated with this study.

Economic Consideration

There will not be any compensation provided for participation. You also will not be subject to any additional costs. Your insurance will be appropriately billed as undergoing a spinal fusion procedure. You will also still be responsible for any other payments for your visits/procedure including co-pay, co-insurances, and deductible.

Treatment Alternatives
If you choose not to participate in this study, you may search for other techniques of spinal fusion that may include other hardware components such as interbody cages or spacers. You may also find other surgeons who perform spinal fusion via pedicle screw placement with other techniques such as C-arm fluoroscopy or freehand, although studies show to be less accurate than our control group. You also may decline surgical intervention all together and proceed with chronic symptomatic treatment.

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. Examples of information that we are legally required to disclose include abuse of a child or elderly person, or certain reportable diseases. Physical documents will be stored in a locked cabinet or a room only accessible by medical staff and/or principal investigator. Data stored electronically will be protected via encryption software and will only be accessed by lead research investigators. This data will be stripped of all identifiable information and will be assigned a number associated with your treatment. 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. Additionally, once all analyses are run and the study is completed, the raw data obtained will be destroyed or deleted.

All health care providers, members of the health team, and members of the research staff are subject to HIPAA. All members are mandated to protect the privacy of your information. Additionally, anyone accessing your electronic health records and using PHI will be mindful of HIPAA.

Voluntary Participation and Withdrawal

The participation of this study is completely voluntary. It is up to you to decide freely if you would like to enroll or not in this study. If you refuse to participate in the study, there will no penalty or loss of benefits. You do not give up any legal rights by signing this form.

If you do decide to enroll in the study, you are able to withdraw yourself from the study at any time, again without any penalty or loss of healthcare outside of the study. There will be no hindrance in the relationship between your own primary care team. At this point, no new information will be gathered but data already gathered may still be used. To formally withdraw, please call the contact information below and report that you no longer want to be a subject to this study at any time.

Questions
We have used some technical terms in this form. Please do your due diligence about this information/research and ask about anything you don't understand and to consider this consent form carefully, as long as you feel is necessary, before you make a decision.

**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 and I understand it in its entirety. My signature also indicates that I have received a copy of this consent form.

Name of Subject (print): ________________________

Signature: __________________________________

Relationship: _________________________________

Date: ______________________________________

___________________________________________

Signature of Principal Investigator                      Date

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, Dr. Peter Whang at (203) 867-5309.

If, after you have signed this form you have any questions about your privacy rights, please contact the Yale Privacy Officer at (203) 432-5919. If you would like to talk with someone other than the researchers to discuss problems, concerns, and questions you may have about this research, or to discuss your rights as a research subject, you may contact the Yale Human Investigation Committee at (203) 785-4688.
Appendix B: Recommended Participating Institutions

1. Yale New Haven Hospital Spine Center, New Haven, CT
2. Johns Hopkins Neurosurgical Spine Center, Baltimore, MD
3. Stanford University Medical Center, Palo Alto, CA
4. Columbia University Irving Medical Center, New York, NY

Extra sites for contingency
1. Mayo Clinic Spine Center, Jacksonville, FL
2. UCLA Spine Center, Santa Monica, CA
3. Duke University Hospital, Durham, NC
4. Arizona Spine and Joint Hospital, Mesa, AZ
5. University of Washington Medical Center, Seattle, WA
Appendix C: Sample Size Calculation

Sample Size was calculated using Power and Precision software

Proportion positives were extracted from retrospective study Mao et al. Very similar effect sizes were observed across several studies including a randomized controlled trial. These studies had agreeing effect sizes of approximately 20%.

<table>
<thead>
<tr>
<th>Group</th>
<th>Proportion Positive</th>
<th>N Per Group</th>
<th>Standard Error</th>
<th>95% Lower</th>
<th>95% Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>O-Arm</td>
<td>0.66</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robot assisted</td>
<td>0.86</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rate Difference: -0.20  142  0.07  -0.34  -0.06

Alpha= 0.050, Tails= 2  
Power = 0.803

The program displays power

For the given effect size (population proportions of 0.66 vs. 0.86), sample sizes (71 and 71), and alpha (0.050, 2-tailed), power is 0.803. This means that 80% of studies would be expected to yield a significant effect, rejecting the null hypothesis that the two population proportions are equal.

With a loss of follow up estimated to be 5% due to death or improvement only, will plan to have atleast 150 subjects (142 ÷ 0.95 = 149.5)
Appendix D: Patient Questionnaire

Patient Legal Name: ________________________

Date of Birth: ______________________________ (Age: ____). (Sex: _____)

Today’s Date: _____________________________

Height: _________________________________

Weight: _________________________________

BMI: _________________________________

Medical conditions (including but not limited to high blood pressure, diabetes, high cholesterol)

Current Medications:

Previous Surgeries:

Expected Number of Vertebral Levels to Undergo Fusion: _____
Appendix E: Tables

Table 1: SRS-Schwab Classification

<table>
<thead>
<tr>
<th>4 Types of curves</th>
<th>3 Sagittal modifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T</strong> Thoracic</td>
<td><strong>Fl minus LL</strong></td>
</tr>
<tr>
<td>lumbar curve &lt; 30°</td>
<td>0: up to 10°</td>
</tr>
<tr>
<td><strong>L</strong> TL / lumbar</td>
<td>+: moderate 10-20°</td>
</tr>
<tr>
<td>thoracic curve &lt; 30°</td>
<td>++: marked &gt; 20°</td>
</tr>
<tr>
<td><strong>D</strong> Double curve</td>
<td><strong>Global alignment</strong></td>
</tr>
<tr>
<td>T and T/L curve, both &gt; 30°</td>
<td>0: SVA &lt; 4 cm</td>
</tr>
<tr>
<td><strong>S</strong> Sagittal deforLA</td>
<td>+: SVA 4-9.5 cm</td>
</tr>
<tr>
<td>mity</td>
<td>++: SVA &gt; 9.5 cm</td>
</tr>
<tr>
<td>coronal curves &lt; 30° L</td>
<td><strong>Pelvic Tilt</strong></td>
</tr>
<tr>
<td>moderate or</td>
<td>0: PT &lt; 20°</td>
</tr>
<tr>
<td>severe modifier(s)</td>
<td>+: PT 20-30°</td>
</tr>
<tr>
<td></td>
<td>++: PT &gt; 30°</td>
</tr>
</tbody>
</table>

Table 2. Gertzbein and Robbins Classification of Pedicle Screw Accuracy

<table>
<thead>
<tr>
<th>Grade</th>
<th>Breach distance (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>&lt;2</td>
</tr>
<tr>
<td>C</td>
<td>&lt;4</td>
</tr>
<tr>
<td>D</td>
<td>&lt;6</td>
</tr>
<tr>
<td>E</td>
<td>&gt;6</td>
</tr>
<tr>
<td>Inclusion Criteria</td>
<td>Exclusion Criteria</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Ages 45-85</td>
<td>SRS-Schwab classification of marked sagittal deformities</td>
</tr>
<tr>
<td>Adult Degenerative Scoliosis Diagnosis</td>
<td>Fusions involving spinal levels proximal to T10</td>
</tr>
<tr>
<td>Life limiting symptoms</td>
<td>Specific Current and Past Medical History</td>
</tr>
<tr>
<td>• Significant dysmorphia of self-image</td>
<td>• Current infection</td>
</tr>
<tr>
<td>• Issues with posture</td>
<td>• Severe systemic disease</td>
</tr>
<tr>
<td>• Balance issues</td>
<td>• Cancer</td>
</tr>
<tr>
<td>• Leg claudication</td>
<td>• Spinal tumor</td>
</tr>
<tr>
<td>• Bowel and/or bladder issues</td>
<td>• Prior spinal surgery</td>
</tr>
<tr>
<td>Complete consent form</td>
<td>• Multiple pathologies of adjacent vertebra</td>
</tr>
<tr>
<td></td>
<td>• Previous spinal fractures</td>
</tr>
<tr>
<td></td>
<td>• Severe pedicle deformity</td>
</tr>
<tr>
<td></td>
<td>• Coagulation disorder</td>
</tr>
<tr>
<td></td>
<td>• Osteoporosis</td>
</tr>
<tr>
<td></td>
<td>• Pregnancy</td>
</tr>
<tr>
<td></td>
<td>BMI &gt; 39.99</td>
</tr>
</tbody>
</table>
### Table 4: Participant Baseline Characteristics with Method of Analysis

<table>
<thead>
<tr>
<th></th>
<th>Robot-Assisted</th>
<th>O-Arm</th>
<th>Type of Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Medians and IQR</td>
<td>Median and IQR</td>
<td>Mann-Whitney U</td>
</tr>
<tr>
<td>BMI</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Student T-test</td>
</tr>
<tr>
<td>Sex</td>
<td>N (%)</td>
<td>N (%)</td>
<td>Chi-Square test</td>
</tr>
<tr>
<td>Hypertension</td>
<td>N (%)</td>
<td>N (%)</td>
<td>Chi-Square test</td>
</tr>
<tr>
<td>Diabetes</td>
<td>N (%)</td>
<td>N (%)</td>
<td>Chi-Square test</td>
</tr>
<tr>
<td>Smoking status</td>
<td>N (%)</td>
<td>N (%)</td>
<td>Chi-Square test</td>
</tr>
</tbody>
</table>

### Table 5: Primary Outcomes with Method of Analysis

<table>
<thead>
<tr>
<th>Grade</th>
<th>Robot-Assisted</th>
<th>O-Arm</th>
<th>Type of Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>N (%)</td>
<td>N (%)</td>
<td>Chi-Square test</td>
</tr>
</tbody>
</table>

### Table 6: Secondary Outcomes with Method of Analysis

<table>
<thead>
<tr>
<th></th>
<th>Robot-Assisted</th>
<th>O-Arm</th>
<th>Type of Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade A+B</td>
<td>N (%)</td>
<td>N (%)</td>
<td>Chi-Square test</td>
</tr>
<tr>
<td>Blood Loss</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Student T-test</td>
</tr>
<tr>
<td>Operative Time</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Student T-test</td>
</tr>
<tr>
<td>Radiation exposure</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Student T-test</td>
</tr>
</tbody>
</table>
Bibliography


