Markerless Surface Topography for Adolescent Idiopathic Scoliosis Screening, Torso Aesthetic Appearance Assessment, and Postural Outcomes Evaluation

by

Nada Mohamed

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Mechanical Engineering

University of Alberta

Abstract

The primary goal of conservative treatment for adolescents with idiopathic scoliosis (AIS) is to prevent the spinal curve from worsening. Other goals of treatment are to improve aesthetic outcomes through postural correction, address respiratory dysfunction, and treat pain symptoms.

The likelihood of curve worsening during growth depends on several factors, such as age at diagnosis, skeletal age, and menarche presence for females. The spinal curvature may worsen during the rapid growth period, especially in girls. Screening for scoliosis aims to detect the condition early and treat it to avoid progression, potentially allowing for less invasive options. Screening for AIS involves conducting a surface assessment by a clinical expert in order to recommend a referral for radiographic exam evaluation for scoliosis confirmation. The Cobb angle measured from a standing posterior-anterior radiograph is used to confirm and diagnose scoliosis. However, current screening techniques like the Adam's forward bend test and the angle of trunk measurement have high referral rates with many false-positives and expose patients to unnecessary radiation from x-rays.

Additionally, aesthetic appearance of the torso is emphasized in scoliosis treatment since AIS patients are concerned with their appearance, which affects their self-image and their quality of life. Surface topography (ST) systems, trunk asymmetry scales, and back photography have been proposed to analyze aesthetic outcomes objectively. However, there is a lack of consensus and research on how to measure aesthetic outcomes effectively.

One of the highlighted goals of scoliosis treatment is to improve appearance and posture, which can be achieved through exercise regimens. The Schroth method is a physiotherapeutic scoliosis-specific exercise (PSSE) designed to correct posture. Although the Schroth method has shown

promising results in reducing the Cobb angle, improving muscle endurance, and improving quality of life aspects, its impact on external deformity and aesthetics requires further investigation.

Surface topography (ST) presents a noninvasive approach for assessing torso asymmetry, which was developed previously to estimate curve severity and progression. The ST method is an alternative tool for screening and detecting idiopathic scoliosis in adolescents. ST also offers the potential for objectively evaluating posture and aesthetic outcomes.

The purpose of this thesis was to develop the ST method for scoliosis screening, and to assess the ST method as a tool for aesthetic torso appearance evaluation to determine the effect of Schroth program on posture outcomes.

In the first objective of the thesis, patterns of asymmetry revealed from the ST method and extracted parameters were evaluated to develop a classification scoliosis screening model. This work developed a scoliosis screening model using ST deviation and torso depth maps to detect AIS from children with typically developing spines using convolutional neural networks (CNN, a deep learning algorithm). A sensitivity of 97% and a specificity of 90% for classifying positive and negative AIS cases were observed. The model was also validated with external data and compared with clinical screening tools, where preliminary results of the developed model obtained a higher specificity than the Adam's forward bend test and comparable specificity with the scoliometer test.

Another objective of the thesis, which was to determine the potential use of ST as an aesthetic assessment tool, parameters from ST analysis were also evaluated to assess responsiveness and association with patient perception of their back condition during six months of conservative treatment. The effect of the Schroth PSSE program added to the standard of care on external

asymmetries and aesthetics was evaluated using the ST method. This work determined the minimally important change (MIC) in surface topography asymmetry measurements for patients to perceive an improvement in their back condition. This work has shown an overall accuracy of 68%, a sensitivity of 62%, and a specificity of 74% for the MIC thresholds to detect perceived back improvement. In addition to routine care (observation or bracing), Schroth exercise therapy reduced asymmetry measurements in AIS and exceeded the estimated MIC of the ST parameters.

This thesis revealed the markerless ST technique as a potential method for scoliosis screening tool, and a posture and aesthetics evaluation tool, in addition to previous research of using ST to predict curve severity and progression.

Preface

This thesis is an original work by Nada Mohamed. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Boad, which are the following: project Name “Surface Topography”, No. Pro00118197, June 2022, project Name “Surface topography for scoliosis”, No. Pro00117065, March 2022, project Name “Surface Topography analysis of healthy participants”, No. Pro00118643, April 2022, project Name “Clinical validation of a Convolution Neural Network (CNN) approach based on Surface Topography (ST) for Idiopathic Scoliosis (IS) Screening in children enrolled in routine sports activity”, No. Pro00135870, December 2023, project Name “ Schroth Exercise trial for Scoliosis”, No. Pro00011552, September 2010, project Name “Multicenter Schroth Exercise Trial for Scoliosis”, No. Pro00043397, November 2013.

Chapter 6 of this thesis has been published as Mohamed, N., Acharya, V., Schreiber, S., Parent, E. C., & Westover, L. (2024). Effect of adding Schroth physiotherapeutic scoliosis specific exercises to standard care in adolescents with idiopathic scoliosis on posture assessed using surface topography: A secondary analysis of a Randomized Controlled Trial (RCT). PLOS ONE, 19(4), e0302577. https://doi.org/10.1371/journal.pone.0302577. I was responsible for data curation and analysis, as well as manuscript preparation. Dr. Schreiber and Dr. Parent were involved in the data collection. All the named authors contributed to the manuscript edits. Dr. Westover and Dr. Parent were the supervisory authors.

Acknowledgements

This thesis would not be possible without the help and support of many individuals.

I would like to express my deepest gratitude to my supervisor, Dr. Lindsey Westover for her support, and guidance throughout both my undergraduate and MSc programs. Having had the privilege of working under her supervision as a research student during my undergraduate studies back in 2018, and again in 2020, I have greatly benefited from her expertise, patience, and insightful feedback over the years. Dr. Westover has been instrumental in shaping my academic journey and I am profoundly grateful for it.

I would like to thank my supervisor Dr. Qipei Mei for his guidance and expertise throughout my program and giving valuable insights and feedback.

Special thanks to my fellow lab members and graduate students: Mostafa H., Jose, Maha, Mostafa M., Chester, and Eric. Their camaraderie, collaboration, and encouragement have made this journey both enjoyable and enriching.

I would like to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), the Women and Children's Health Research Institute (WCHRI), and the Faculty of Graduate Studies and Research (FGSR) for funding this research project.

I would also like to thank my siblings for their support and encouragement, and for always believing in me. Finally, I would like to express my deepest appreciation to my parents. The unconditional love, support, and encouragement from my mother and father have been my foundation throughout this journey. They have always been a constant source of strength and motivation. I want to thank them for their patience, understanding, and for always being there to listen and offer words of wisdom. I am forever grateful for everything they have done to help me reach this point.

Table of Contents

Abstract ii

Preface v

Acknowledgements vi

Table of Contents vii

List of Tables xi

List of Figures xiii

Chapter 1 : Introduction *1*

1.1 Background and Motivation 1

1.2 Problem Statement 4

1.3 Thesis Objectives 4

1.4 Thesis Outline 5

Chapter 2 : Literature Review *6*

2.1 Scoliosis 6

2.2 Scoliosis Curve Types 8

2.3 Scoliosis Monitoring 9

2.3.1 Radiography 9

2.3.2 Surface Topography 10

2.4 Scoliosis Screening 12

2.5 Conservative Treatments and Scoliosis Specific Exercises 15

2.6 Trunk Aesthetics Evaluation 16

2.7 Summary 17

Chapter 3 : Adolescent Idiopathic Scoliosis Screening *18*

3.1 Summary 18

3.2 Introduction 19

3.3 Methodology 21

3.3.1 Data Acquisition 21

3.3.2 Surface Topography Analysis 22

3.3.3 Convolutional Neural Network 24

3.3.4 Statistical Analysis 25

3.4 Results 26

3.5 Discussion 29

3.6 Conclusion 32

Chapter 4 : Scoliosis Screening Validation *33*

4.1 Introduction 33

4.2 Methods 34

4.2.1 Measurements 34

4.2.2 Statistical Analysis 35

4.3 Results 35

4.4 Discussion 39

4.5 Conclusion 41

Chapter 5 : Aesthetic Appearance Assessment *42*

5.1 Summary 42

5.2 Introduction 43

5.2 Methods 46

5.2.1 Data Acquisition 46

5.2.2 Measurements 46

5.2.3 Statistical analysis 49

5.3 Results 49

5.3.1 Analysis by Curve Types 53

5.4 Discussion 54

5.5 Conclusion 57

Chapter 6 : Effect of Scoliosis Treatment on Torso Asymmetry *58*

6.1 Abstract 58

6.2 Introduction 59

6.3 Methods 61

6.3.1 Ethics Statement 61

6.3.2 Asymmetry Analysis 63

6.3.3 Longitudinal Asymmetry Evaluation 65

6.3.4 Statistical Analysis 65

6.4 Results 67

6.4.1 Intention to Treat Comparisons of Changes in RMS and MaxDev 69

6.4.2 Per Protocol Comparisons of Changes in RMS and MaxDev 70

6.4.3 Intention to Treat Group Comparison Based on the Classification 72

6.4.4 Per Protocol Group Comparison Based on the Classification 73

6.4.5 Intention to Treat Analysis of Control Group when Receiving Delayed Schroth Exercises 74

6.4.6 Per Protocol Analysis of Control Group when Receiving Delayed Schroth Exercises 75

6.5 Discussion 75

6.6 Conclusion 79

Chapter 7 : Association of Toso asymmetry with 3D Vertebral Position *80*

7.1 Introduction 80

7.2 Methods 81

7.3 Results 83

7.4 Discussion 90

7.5 Conclusion 91

Chapter 8 : Conclusion *92*

8.1 Summary 92

8.1.1 Development of Screening Model based on ST Analysis Using a CNN Algorithm _ 92
8.1.2 Validation and Comparison of ST Classification Model to Standard Clinical Screening Tools 93
8.1.3 Evaluation of the ST Method for Assessing Torso Appearance Aesthetics 93
8.1.4 Effect of Adding Schroth PSSE to Standard of Care on Posture Outcomes Using ST Method 94
8.1.5 Correlation of underlying spinal curve and the ST characteristics 95
8.2 Future Work 95
References 97
Appendix A: 116
Convolutional Neural Network Model Parameter Optimization Procedure 116
Dataset 116
Data Preprocessing 116
CNN Parameters Optimization 117
Optimized Parameters 120
Appendix B: 121
Linear Mixed Effects Model Analysis to Determine the Effect of Schroth PPSE on Torso Asymmetry 121

List of Tables

Table 3.1 Confusion matrix outcome variables comparing prediction classifier with ground truth to determine model performance. Sensitivity, specificity, false positive rate, false negative rate, positive likelihood ratio, and negative likelihood ratio were determined from the confusion matrix to evaluate the overall performance of the model 25

Table 3.2: Training, Validation, and testing set characteristics (mean ± standard deviation) 26

Table 3.3: Training and validation errors (dimensionless) and accuracy comparing CNN outputs to the ground truth for each cross-validation fold 27

Table 3.4: Performance assessments from the testing set of the CNN model obtained after each cross-validation fold 27

Table 4.1: Sample demographic characteristics with values represented as mean ± standard deviations for continuous variables and percentages for categorial variables 35

Table 4.2: Observed outcomes from ST, compared to Adams's forward bend, and scoliometer tests of the complete sample (n=70). Positive screening result indicated by (+) and negative screening results indicated by (-) 36

Table 4.3: Observed outcomes from ST, compared to the Adams's forward bend, and scoliometer tests of the sample (n=64), excluding participants with confirmed scoliosis from x-ray exams and those awaiting results. Positive screening result indicated by (+) and negative screening results indicated by (-) 37

Table 5.1: Baseline characteristics for the whole sample and split by curve type 50

Table 5.2: Correlation coefficients for the association between the change in ST parameters (ΔRMS, ΔMaxDev) with patient reported GRC scores 51

Table 5.3: Responsiveness results and minimally important changes (MIC) of ST parameters from receiver operating characteristic (ROC) analysis from the complete group and curve type analyses 52

Table 5.4: Accuracy results of classifying perceived improvement and no change/deteriorated (GRC <2) using RMS and MaxDev MICs for the entire sample, thoracic curves group, and lumbar curves group 54

Table 6.1: Patient characteristics at baseline for the Schroth and Control groups for the intention to treat and per protocol analyses 68

Table 6.2: Proportions of participants reported as N (percent of group sample) showing improvement, no change or progression in asymmetry after six months from the intention to treat and per protocol analyses 73

List of Figures

Fig. 2.1: Adolescent idiopathic scoliosis (Paria & Wise, 2015) 7
Fig. 2.2: Curve type classifications using the Schroth curve type algorithm (Schreiber et al., 2023) 8
Fig. 2.3: Measuring the Cobb angle from posterior-anterior radiograph (S. Lee et al., 2018) 9
Fig. 2.4: Torso asymmetry assessments using the ST method. DCM shown for AIS participants with a Cobb of 36° in the thoracic region (left), a Cobb of 25° in the lumbar region (middle), and a non scoliotic participant with an ATR of 4° (right) 11
Fig. 2.5: Scoliosis screening using the scoliometer to measure rib hump inclination of a participant with 30° Cobb (Weinstein et al., 2008) 13
Fig. 3.1: Data collection procedure with (a) scanners to obtain (b) 3D trunk surface. ST analysis processing procedure showing (c) best fit alignment of the original and reflected torsos, (d) deviation color map of the difference between original and reflected torsos, (e) back torso retained, representation of (f) back deviation and (g) surface torso depth maps 23
Fig. 3.2: Convolutional neural network (CNN) architecture 24
Fig. 3.3: Confusion matrices with corresponding ROC curve comparing CNN classification algorithm using ST with ground truth of the combined data, mild severity data (Cobb ≤ 25°) and moderate severity data (Cobb > 25°) 28
Fig. 4.1: Deviation color maps of participants screened positive with ST method. The CNN model output of each case with its sigmoid probability is reported 37
Fig. 4.2: Deviation color maps of participants screened positive with Adam's test. The CNN model output of each case with its sigmoid probability is reported 38
Fig. 4.3: Deviation color maps of participants screened positive with the scoliometer test. The CNN model output of each case with its sigmoid probability is reported 39
Fig. 5.1: ST analysis depicting A) best fit alignment of original (silver) and reflected torso (grey), B) deviation color map and C) isolated patch for obtaining RMS and MaxDev parameters 48
Fig. 5.2: GRC scores and the corresponding change in RMS and MaxDev with the correlation R value reported 51
Fig. 5.3: Receiver operating characteristics (ROC) curves of patient perceived improvement (GRC ≥ 2) at various cutoff points of change in RMS and MaxDev 52
Fig. 6.1: CONSORT flow-chart showing Intention to treat and per protocol group sample sizes 62

Fig. 6.2: Surface topography procedure(a) Alignment of the reflected (gold) and original torso (grey), (b) deviation color map where red reflects area of protrusion and blue to areas of depression relative to the other side, and (c) isolated patch of interest to calculate measurement parameters 64

Fig. 6.3: Examples of deviation colour maps, RMS and MaxDev measurements of the largest patch by area over time. (a) Patient experiencing progression, (b) patient with asymmetry improvement, and (c) no change in asymmetry 66

Fig. 6.4: Comparison between groups from intention to treat analysis (mean ± SE). Comparison of (a) RMS (mm) and (b) MaxDev (mm) outcomes. Linear mixed effects analysis produced the p-value presented 70

Fig. 6.5: Comparison between groups from the per protocol analysis (mean ± SE). Comparison of (a) RMS (mm) and (b) MaxDev (mm) outcomes. Linear mixed effects analysis produced the p-value presented 72

Fig. 6.6: Average RMS (mm) and MaxDev (mm) in the group serving as control during the first six months and receiving exercises in the last six months presented over time from the intention to treat analysis (mean ± SE). Linear mixed effects analysis produced the p-value presented 74

Fig. 6.7: Average RMS (mm) and MaxDev (mm) in the group serving as control during the first six months and receiving exercises in the last six months presented over time from the per protocol analysis (mean ± SE). Linear mixed effects analysis produced the p-value presented 75

Fig. 7.1: Data processing and acquisition of ST and 3D spine features for analysis 83

Fig. 7.2: Correlation of RMS of posterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D) 84

Fig. 7.3: Correlation of MaxDev of posterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D) 85

Fig. 7.4: Correlation of RMS of anterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D) 86

Fig. 7.5: Correlation of MaxDev of anterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D) ____ *87*
Fig. 7.6: Correlation of posterior torso surface patch centroid with the apical vertebra position in the lateral direction (A), superior direction (B), and posterior direction (C) ____ *88*
Fig. 7.7:Correlation of anterior torso surface patch centroid with the apical vertebra position in the lateral direction (A), superior direction (B), and posterior direction (C) ____ *89*

Chapter 1 : Introduction

1.1 Background and Motivation

Adolescent idiopathic scoliosis (AIS) is an abnormal curvature and rotation of the spine in three-dimensional space (Choudhry et al., 2016; Roach, 1999). Trunk asymmetry can be visually observed due to the spine deformity, which includes prominence of the scapula, uneven shoulders and hips, rib hump appearance when bending forward, and a sideways shift of the trunk and waist (Choudhry et al., 2016). Children and adolescents with idiopathic scoliosis can experience breathing difficulties, back pain, and decreased physical function (J. Danielsson & L. Nachemson, 2003). Due to the noticeable asymmetric appearance of their back, patients with AIS often experience low self-image (Belli et al., 2022; Essex et al., 2022). Scoliosis is monitored through radiographic examinations, and curve severity is assessed using the Cobb angle determined from posterior-anterior radiographs (Weinstein et al., 2008). A curve with a Cobb angle greater than 10° indicates a positive diagnosis of AIS (Weinstein et al., 2008).

One of the primary goals of the conservative treatment of AIS is to prevent the spinal curve from worsening (Negrini et al., 2018). The age at onset, skeletal age and the appearance of menarche in females are prognostic indicators of the likelihood of curve progression, all of which refer to the patients' growth potential (Bunnell, 1986). Progression is prevalent in girls during periods of fast growth (Choudhry et al., 2016; Rogala et al., 1978; Soucacos et al., 1998). The increased risk of the curve worsening during puberty highlights the need for early screening and continuous monitoring (Bunnell, 1986). Screening aims to identify scoliosis early when the deformity is likely undetected and when a less intrusive treatment option may be available (Bunnell, 2005). Standard screening techniques include the Adam’s forward test, which provides a subjective assessment of the external trunk, and the angle of the trunk rotation (ATR) measured using a scoliometer device that can provide an objective assessment (Choudhry et al., 2016; Negrini et al., 2018). However, these techniques have high referral rates due to their low specificity, which leads to unnecessary ionizing radiation exposure from radiography examinations (Negrini et al., 2018).

In recent years, aesthetics has become one of the main goals of scoliosis treatment, according to the International Society on Scoliosis Orthopaedic and Rehabilitation Treatment (SOSORT) (Negrini et al., 2006). Improving aesthetic and posture appearance is a concern for AIS patients and parents (Essex et al., 2022). Despite being the most widely used indicator for monitoring the scoliosis curve, the Cobb angle ranks behind aesthetics as a priority when treating AIS (Negrini et al., 2006). Although aesthetics is considered an important outcome for patients and physicians, no clinical instruments are available to measure how aesthetics evolves over treatment (Negrini et al., 2006, 2018). Methods to evaluate aesthetics have been proposed, which can be broadly categorized as subjective and objective tools. Subjective tools include visual inspection and photography comparison of the back over time and patient-reported questionnaires. A commonly used questionnaire is the Scoliosis Research Society 22 (SRS-22), which includes a portion on self-image (Asher et al., 2003). The Walter Reed assessment scale, and the trunk appearance perception (TAPS) scale are additional methods developed using standardized drawings for patients to describe their perception of their spine deformity (Sanders et al., 2003; Zaina, Negrini, & Atanasio, 2009). Objective aesthetic evaluation tools such as the posterior trunk symmetry index (POTSI) and the anterior trunk symmetry index (ATSI) based on measures from photographs have been proposed (Stolinski et al., 2012). The POTSI scale is used to evaluate the asymmetry of the posterior trunk. The POTSI parameter combines measurements of six indices taken from specific anatomical landmarks of the trunk. Similar to the POTSI scale, the ATSI scale was proposed to analyze the asymmetry of the trunk surface. The ATSI scale specifically assesses the anterior aspect of the trunk since the scoliosis deformity also affects the front of the trunk and can be easily noticed by patients. Both the POTSI and ATSI scales accuracy and reliability depend on the torso landmarks' identification, which can result in measurement error. Additionally, the POTSI indicator has a poor standardized response mean, suggesting that it may not be a reliable tool for monitoring the progression or improvement in trunk symmetry in response to treatment interventions (Parent, Damaraju, et al., 2010). Surface topography can also be used to objectively evaluate a person's aesthetic profile (Hackenberg et al., 2003; Su et al., 2022; Theologis et al., 1993).

Due to a lack of consensus on measuring aesthetics, additional research is needed to evaluate the efficacy of AIS treatment on cosmetic appearance (Negrini et al., 2006; Romano et al., 2013). Invasive spine surgeries can improve cosmetic outcomes, and non-invasive means such as bracing

and exercises may also improve aesthetics (Hawes, 2003; Negrini et al., 2006; Zaina, Negrini, Fusco, et al., 2009). A notable exercise technique is the Schroth Physiotherapeutic Scoliosis Specific Exercises (PSSE) method to improve posture (Weiss, 2011). Schroth exercises involve breathing techniques, muscle strengthening, stretching, and postural correction exercises (Weiss, 2011). These exercises are typically performed in various positions, including lying down, sitting, and standing, to target different muscle groups and promote spinal alignment. The Schroth method has shown improved Cobb angle, breathing function, back muscle strength, and slowing curve progression in several studies (Kuru et al., 2016; Otman et al., 2005; Schreiber et al., 2015, 2016). However, quantitative evaluation of the effect of Schroth exercises on external deformity and aesthetics has yet to be adequately studied (Romano et al., 2013).

Although the Cobb angle is the gold standard indicator for quantifying the spine curve severity, it has several limitations. The Cobb angle measures the curvature of the spine in two dimensions in only the posterior-anterior plane. However, scoliosis is a complex 3D condition, and the Cobb angle may not fully capture the extent of rotational deformity or the actual severity of the curve (Ramirez et al., 2006; Thulbourne & Gillespie, 1976). Scoliosis patients, especially children and adolescents, require frequent monitoring of their spinal curvature using x-rays over an extended period (Negrini et al., 2018). Cumulative exposure to ionizing radiation from repeated x-rays can increase the risk of developing cancer (Levy et al., 1996; Nash et al., 1979). Assessment of the spine in 3D can be valuable in planning an effective scoliosis management plan and monitoring progression over time. Capturing additional x-ray views provides comprehensive data for 3D reconstructions, aiding with treatment decisions. However, additional x-ray projections are not recommended unless necessary due to radiation exposure. An alternative approach has been proposed to estimate the spine profile from surface trunk analysis.

Surface topography (ST) asymmetry analysis has been introduced as a non-invasive technique to assess torso asymmetry along the best plane of symmetry (Hill et al., 2014; Komeili et al., 2014). ST captures the external torso surface using only visible light and does not rely on manual landmarks (Komeili et al., 2014). The torso asymmetry is visualized and quantified through a deviation map comparing the torso model with its reflection. The markerless ST method has been proposed to reduce the reliance on radiographs in the management of scoliosis. Classification trees

based on ST deviation maps have been developed to estimate curve severity and progression (Ghaneei et al., 2018; Hong et al., 2017; Komeili et al., 2014, 2015b).

1.2 Problem Statement

Standard screening tools have low specificity and high referral rates, which can expose unnecessary radiation to children. The accuracy of the standard tools also depends on the experience of the operator. In this work, we propose that the ST method quantifying torso asymmetry is an objective tool that can also be used for screening to detect the AIS accurately and address the low specificity of existing clinical screening tools. Aesthetics appearance is a primary concern for patients. The lack of consensus on methods for objectively evaluating aesthetic highlights the need to conduct further research in cosmetic outcomes evaluation techniques. Additionally, we propose that the ST method can assess posture and aesthetics outcomes associated with conservative treatment, specifically, posture effects from PSSE programs.

1.3 Thesis Objectives

The objective of this thesis is to develop and validate a markerless ST asymmetry-based model for AIS detection, and to assess the ST method as a tool for torso aesthetics outcomes in order to evaluate the impact of Schroth exercises on posture. The objective will be addressed with the following specific aims (SA):

SA1. To develop and propose a classification model aimed at AIS detection using the markerless ST asymmetry technique.

SA2. To validate the proposed screening model, determine its overall accuracy, and compare the performance with clinical screening tools such as the Adam's forward bend test and the scoliometer test.

SA3. To determine the changes in ST asymmetry measures and their relation to patient self-reported perception using the Global Rating of Change (GRC) in adolescents with idiopathic scoliosis

SA4. To determine the effect of Schroth exercises added to the standard of care on posture outcomes using ST asymmetry measures.

SA5. To determine whether the ST parameters are correlated with the underlying spinal curve to predict the spine shape in 3D.

1.4 Thesis Outline

The following is a presentation of the research conducted for this thesis:

Chapter 1 briefly introduces the topic and the background for this work. The motivation and objectives of this thesis are also presented in this chapter.

Chapter 2 presents a literature review of clinical screening tools, aesthetic and posture assessment methods, and surface topography techniques.

Chapter 3 presents the developed model that uses ST asymmetry maps for screening and early detection of AIS.

Chapter 4 validates the screening model with additional population data and compares the performance metrics with common clinical screening tools.

Chapter 5 presents research on the association of the ST measures of posture on the patient self-reported perception of their condition using the global rate of change of AIS patients following six months of treatment. The chapter also presents the minimally important change of ST asymmetry measurements for patients to perceive an improvement in their back condition.

Chapter 6 presents research on the effect of adding six months of Schroth exercises to the standard of care on ST outcomes.

Chapter 7 presents research on the correlation of ST indices and parameters the overall shape of the underlying spine of AIS patients in 3D.

Chapter 8 contains the discussion and concluding remarks of this thesis. Future research is also discussed in this chapter.

Chapter 2 : Literature Review

2.1 Scoliosis

Scoliosis is a condition characterized by an abnormal curvature and rotation of the spine in 3D (Weinstein et al., 2008). Instead of a straight alignment, the spine may curve to the side when viewed from behind. One common type of scoliosis is Adolescent Idiopathic Scoliosis (AIS), which typically appears during adolescence, between the ages of 10 and 18, and has no identifiable cause as suggested by the "idiopathic" term (Weinstein et al., 2008). AIS is the most prevalent form of scoliosis, affecting up to 5% of adolescents (Choudhry et al., 2016; Rogala et al., 1978).

Based on age, scoliosis can be categorized into several types: infantile, juvenile, adolescent, and adult scoliosis (Choudhry et al., 2016; Negrini et al., 2018). Infantile scoliosis occurs in infants, typically before the age of three. Infantile scoliosis may be congenital or acquired and is often detected early due to visible deformities or abnormalities in infancy (Negrini et al., 2018). Juvenile scoliosis manifests between the ages of three and ten (Negrini et al., 2018). Like infantile scoliosis, it may have congenital or acquired origins. Adolescent scoliosis is the most common form of scoliosis and is acquired between the ages of 10 and 18 (Choudhry et al., 2016). The curvature can progress during the adolescent growth spurt, potentially leading to more severe deformities if left untreated (Weinstein et al., 2008; Wong & Tan, 2010). Scoliosis that persists or develops in adulthood is categorized as adult scoliosis. As it is the most prevalent, adolescent idiopathic scoliosis will be the focus of this thesis (Weinstein et al., 2008).

AIS can lead to various visible asymmetries in the body due to the abnormal curvature of the spine. These asymmetries may become more pronounced as the curvature progresses. One noticeable asymmetry is a difference in shoulder height (Fig. 2.1) (Choudhry et al., 2016). One shoulder may appear higher than the other, particularly when the individual stands upright (Choudhry et al., 2016). Due to the rotation of the spine caused by AIS, one side of the ribcage may protrude more prominently than the other. This asymmetry often results in a visible rib hump or prominence on one side of the back, particularly when bending forward (Fig. 2.1) (Choudhry et al., 2016). The curvature of the spine can also lead to irregularities in the waistline. Furthermore, AIS can affect the alignment of the pelvis and hips, resulting in one hip appearing higher or more prominent than the other when standing or walking (Choudhry et al., 2016).

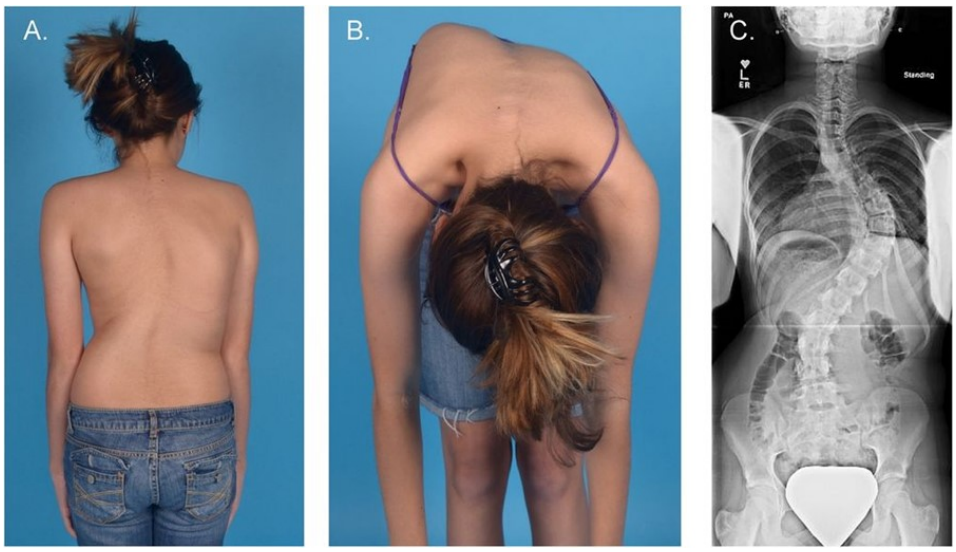


Fig. 2.1: Adolescent idiopathic scoliosis (Paria & Wise, 2015)

Several risk factors can contribute to the progression or worsening of the spinal curve of AIS patients. Firstly, the patient's age during diagnosis plays a crucial role, as growth spurts during adolescence can exacerbate the curvature (Bunnell, 1986). Children diagnosed with AIS between the ages of 10 and 12 have an 88% risk of progression (Bunnell, 1986). Girls are ten times more likely to experience progression compared to boys (Wong & Tan, 2010). The degree of the initial curvature is another significant factor, with larger curves being more prone to progression (Roach, 1999; Weinstein et al., 2008). Additionally, skeletal maturity affects the likelihood of progression (Choudhry et al., 2016; Wong & Tan, 2010). Skeletal maturity can be assessed using the Risser sign obtained from radiographs. The measure evaluates the degree of bone maturity of the iliac apophysis, a bony prominence at the top of the pelvis (Hacquebord & Leopold, 2012). The Risser sign is graded from 0 to 5, with 0 indicating no ossification and 5 indicating complete ossification and closure of the growth plate. Typically, during growth, the Risser sign progresses from 0 to 5, reflecting the closing of growth plates and the completion of skeletal development (Hacquebord & Leopold, 2012). Patients with Risser 0 at the time of diagnosis have a reported 68% risk of curve worsening (Bunnell, 1986). Patients with Risser 1 or 2 at the time of diagnosis have a 52% risk of curve progression (Bunnell, 1986). Finally, a lack of compliance with prescribed treatments, such as bracing or physical therapy, can increase the risk of progression in AIS patients (Weinstein et al., 2008).

2.2 Scoliosis Curve Types

Spinal curves are classified based on their shape, direction, and location. The Lenke classification system is widely used to categorize the various types of spinal curves seen in AIS (Lenke et al., 2003). Six curve types are defined in the classification system in addition to the lumbar spine modifier and sagittal thoracic modifier (Lenke et al., 2003). The Lenke system is primarily used for operative treatment and is unsuitable for non-invasive bracing and exercises (Negrini et al., 2018). Moreover, the Lenke classification is not recommended for mild curves (Negrini et al., 2018).

Alternatively, curve types can be classified using the Schroth classification algorithm specifically for exercise therapy. Classification systems are used to prescribe exercise treatment for patients with scoliosis that is tailored to each patient's curve type (Schreiber et al., 2012, 2023; Watkins et al., 2012). Four Schroth curve types (Fig. 2.2) can be used to categorize scoliosis spinal curve patterns, which include two thoracic (3c and 3cp) and two thoracolumbar/lumbar (4c and 4cp) patterns (Schreiber et al., 2023).

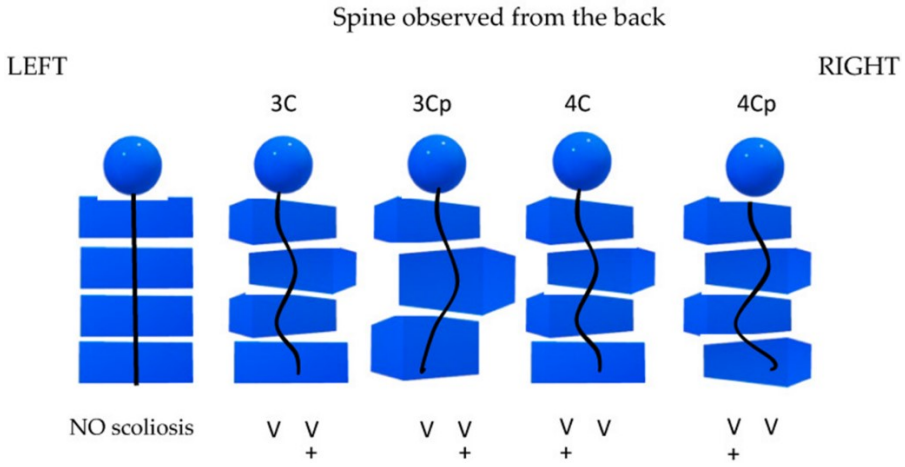


Fig. 2.2: Curve type classifications using the Schroth curve type algorithm (Schreiber et al., 2023)

The four blocks in the figure from top to bottom represent the following: shoulders and neck block, thoracic curve block, lumbar curve block, and hip pelvis block. A 3c curve type is a primary thoracic curve with a minor or no lumbar curve and a balanced pelvis. A 3cp curve is a single

thoracic curve with an unbalanced pelvis. A 4c curve is a significant lumbar curve with a balanced pelvis. A 4cp curve represents a primary thoracolumbar or lumbar curve, may have a thoracic curve, and an unbalanced pelvis (Schreiber et al., 2023).

2.3 Scoliosis Monitoring

2.3.1 Radiography

The gold standard method for assessing the scoliosis curve is the Cobb angle (Weinstein et al., 2008). Measured on a standing posterior-anterior radiograph, the Cobb angle quantifies the severity of the spinal curvature by determining the angle between the most tilted vertebrae at the upper and lower ends of the curve (Fig. 2.3) (Thulbourne & Gillespie, 1976). A Cobb angle of 10° or more is indicative of scoliosis (Weinstein et al., 2008). The Cobb angle provides a standardized way to assess the degree of curvature and helps guide treatment decisions. Generally, mild curves of Cobb less than 25° require observation, where patients are followed up every six or 12 months (Negrini et al., 2018; Roach, 1999). For moderate curves between 25° to 45° Cobb, bracing is prescribed (Roach, 1999). For severe curves (> 45°), surgical treatment to correct the spine and prevent further progression is recommended (Roach, 1999). In addition, the curve is said to have progressed if the Cobb angle increases by 5° or more between consecutive visits (Roach, 1999).

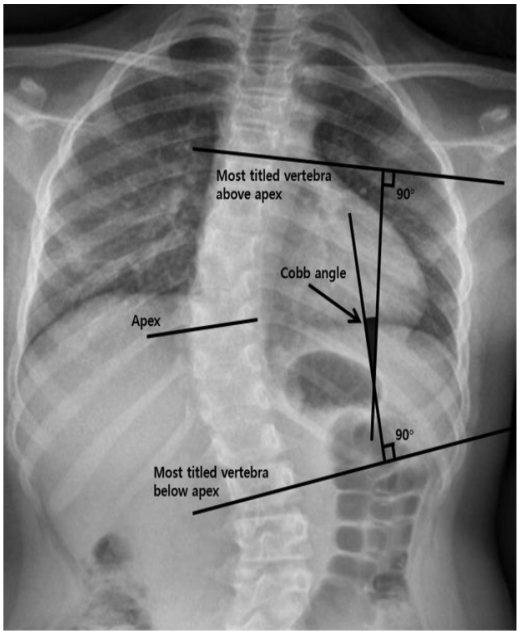


Fig. 2.3: Measuring the Cobb angle from posterior-anterior radiograph (S. Lee et al., 2018)

The Cobb angle measured from radiograph evaluation has several limitations. The measurement relies on 2D x-ray images, which may only partially capture the 3D complexity of the spinal deformities (Ramirez et al., 2006; Thulbourne & Gillespie, 1976). This can lead to inaccuracies, particularly in cases where the curvature occurs in multiple planes or involves spinal rotation. Frequent x-rays are taken to monitor the curve, which exposes the patients to ionizing radiation and may lead to adverse health effects such as an increased risk of cancer (Levine et al., 2016; Nash et al., 1979). Assessment of the scoliosis curve in 3D can be made possible through reconstruction of the spine from biplanar radiographs. However, obtaining additional projection is not recommended for children to reduce radiation exposure. The Cobb angle is also a poor indicator of assessing aesthetics and cosmetic outcomes (Parent, Wong, et al., 2010; Smith et al., 2006). Aesthetic appearance is a significant concern for adolescents with scoliosis (Bridwell et al., 2000). Poor self-image due to their torso and posture appearance is associated with a reduced quality of life (Bridwell et al., 2000; Cheshire et al., 2017; Ponseti et al., 1976). According to the International Society on Scoliosis Orthopaedic and Rehabilitation Treatment (SOSORT), aesthetics is a primary reason for treating AIS (Negrini et al., 2006). Aesthetics was ranked higher than the scoliosis Cobb angle for treatment priority (Negrini et al., 2006).

2.3.2 Surface Topography

Markerless surface topography (ST) has been proposed to monitor scoliosis (Komeili et al., 2014, 2015a, 2015b). The technique is non-invasive, requiring only visible light, and can be used as a secondary tool for monitoring AIS and reducing radiation exposure by reducing the reliance on radiograph assessment (Hong et al., 2017; Komeili et al., 2015b).

The ST technique aims to quantify surface trunk asymmetries associated with scoliosis. The best sagittal plane of symmetry of the torso is estimated and determined (Hill et al., 2014). The deviations between the original torso model and the reflected torso about the best plane of symmetry are obtained and illustrated in a deviation color map (DCM) (Komeili et al., 2014). Parameters are extracted from the areas of high asymmetry most associated with the underlying spinal curve and used to predict curve severity and progression (Fig. 2.4) (Ghaneei et al., 2018; Hong et al., 2017; Komeili et al., 2015b). The ST technique avoids placing anatomical markers on the torso, which introduces errors. Anatomical landmarks on the torso are used to calculate distances and angles to determine indices associated with scoliosis deformity, such as trunk

asymmetry scales like the posterior trunk symmetry index (POTSI) (Negrini et al., 2018). The midsagittal plane thought to represent the plane of symmetry, has landmarks on both the left and right sides. These distances are analyzed to evaluate trunk asymmetry. However, because of the spine's lateral and axial rotation, the optimal plane of symmetry for AIS patients does not align with the midsagittal plane (Ghaneei et al., 2019). Compared to radiographic assessments, the direction of the curve is predicted accurately 100% of the time using the ST technique (Komeili et al., 2015a).With a sample size of 100 AIS patients, accuracy of predicting the proximal thoracic curve location was 63% (Komeili et al., 2015a). The accuracy of predicting the thoracic-thoracolumbar curve location was 92% (Komeili et al., 2015a). For lumbar curves, the location prediction accuracy was 62% (Komeili et al., 2015a). In addition, regression analysis was used to predict the curve's apex using ST. The apex of thoracic-thoracolumbar curves was predicted with an R^2 of 0.89 (Komeili et al., 2015a). Likewise, the apex of lumbar curves was predicted with an R^2 of 0.58 (Komeili et al., 2015a).

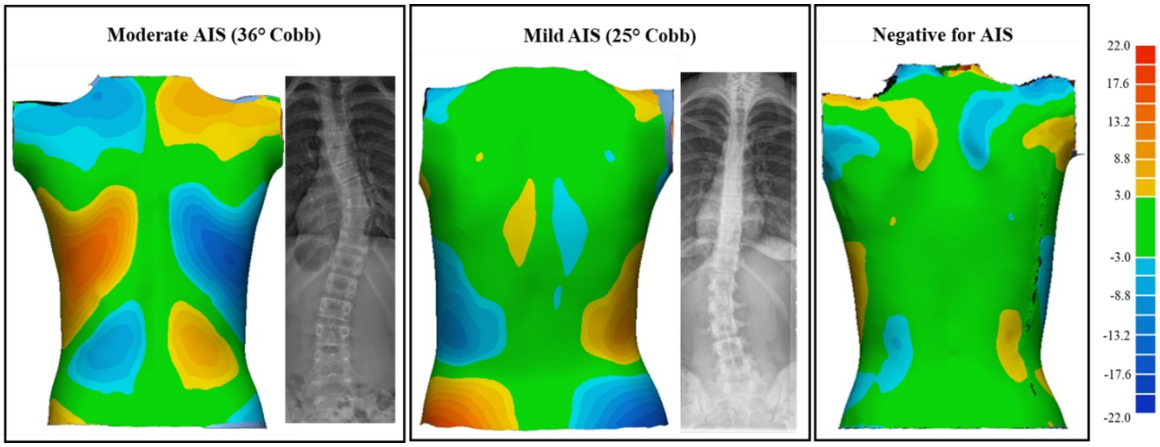


Fig. 2.4: Torso asymmetry assessments using the ST method. DCM shown for AIS participants with a Cobb of 36° in the thoracic region (left), a Cobb of 25° in the lumbar region (middle), and a non scoliotic participant with an ATR of 4° (right)

Classification trees based on extracted parameters from the asymmetry patches of root mean square (RMS) and maximum deviation (MaxDev) were developed to predict the severity of the curves. The decision trees classified ST maps into mild and moderate/severe groups. A classification accuracy of 73% for curves in the thoracolumbar region was obtained (Komeili et al., 2015a). Likewise, an accuracy of 59% was obtained for predicting the severity of lumbar curves (Komeili et al., 2015a). Intra-observer reliability of the decision tree was excellent (kappa coefficient 0.85)

(Komeili et al., 2014). The inter-observer reliability of the decision trees was good (kappa coefficient 0.62). Likewise, the test-retest reliability was excellent (Komeili et al., 2014). Progression classification trees based on RMS and MaxDev values were also developed. The decision tree to predict the progression, where the Cobb angle increased by 5 degrees between subsequent visits, obtained an accuracy of 87.7% (Komeili et al., 2015b). The curve severity and progression were further validated on a new set of data where the curve severity decision tree obtained a sensitivity 95% of and specificity of 35% (Hong et al., 2017). Likewise, the progression decision tree obtained a sensitivity of 73% and a specificity of 53% on the new validation set (Hong et al., 2017). *Ghaneei et al.* used a custom neighbourhood classifier to predict curve severity and progression, since the decision trees developed had high sensitivity but low specificity. With k-nearest neighbourhood classifier, the sensitivity and specificity were 81% and 79%, respectively, to predict curve severity (Ghaneei et al., 2019). Likewise, the sensitivity and specificity were 83% and 95%, respectively, to predict progression using k-nearest neighbourhood classifier (Ghaneei et al., 2019).

2.4 Scoliosis Screening

Radiographs diagnoses AIS with a Cobb angle greater than 10° indicating a positive diagnosis (Weinstein et al., 2008). Screening aims to detect scoliosis early to allow for non-invasive treatments and avoid serious treatments if the deformity goes unnoticed (Grivas et al., 2007; Torell et al., 1981). Screenings are typically conducted in clinical settings or school programs. As outlined in previous sections of this chapter, the risk of progression of the spinal curve can depends on age at diagnosis, menarche presence, and skeletal maturity (Bunnell, 1986). Screenings often detect mild curves, which do not require any conservative treatments, but they allow for monitoring of the spinal curve once the AIS diagnosis is confirmed to prevent curve progression.

The Adam's forward bending test is a standard clinical scoliosis screening test (Choudhry et al., 2016). During this test, the individual is asked to bend forward at the waist while standing with their feet together. As the individual bends forward, a trained observer examines the curvature of the spine, looking for any asymmetry in the rib hump. This test helps to reveal any potential spinal deformities, particularly lateral curvature, or rotation of the spine, which are characteristic of scoliosis. The Adam’s test alone has sensitivity ranging from 84.4%-92% and a specificity of 68% (Côté et al., 1998; Dunn et al., 2018). However, the test is very subjective and is not recommended

for use as a standalone screening tool (Negrini et al., 2018). The Adam's forward bend test alone has high referral rates and low specificity. The scoliometer is a standard device used for objective measurements for scoliosis screenings (Fig. 2.5). The device is designed to measure the angle of trunk rotation (ATR), indicative of spinal curvature (Bunnell, 1984). During the examination, the scoliometer is placed along the individual's back at the same time they are in a forward bend position, allowing for measurements of the rib prominence. The scoliometer measurement has been found to be correlated with radiographic indices (r = 0.685) (Bunnell, 1984; Negrini et al., 2018; Sapkas et al., 2003). In addition, the test showed good inter- and intra-observer reliability (Bunnell, 1984). When an ATR cutoff angle of 5° for positive prediction is used, it has a sensitivity of around 100% and a specificity of approximately 47% (De Wilde et al., 1998; Grosso et al., 2002; Huang, 1997). Specificity increases to 86% while sensitivity decreases to 83% at a cutoff ATR angle of 7° (De Wilde et al., 1998; Grosso et al., 2002; Huang, 1997). The Adam's test false positive and false negative rates are 4.8% and 15.6%, respectively (Dunn et al., 2018). The false positive and false negative rates from the scoliometer test are 19.3% and 9.4%, respectively (Dunn et al., 2018).

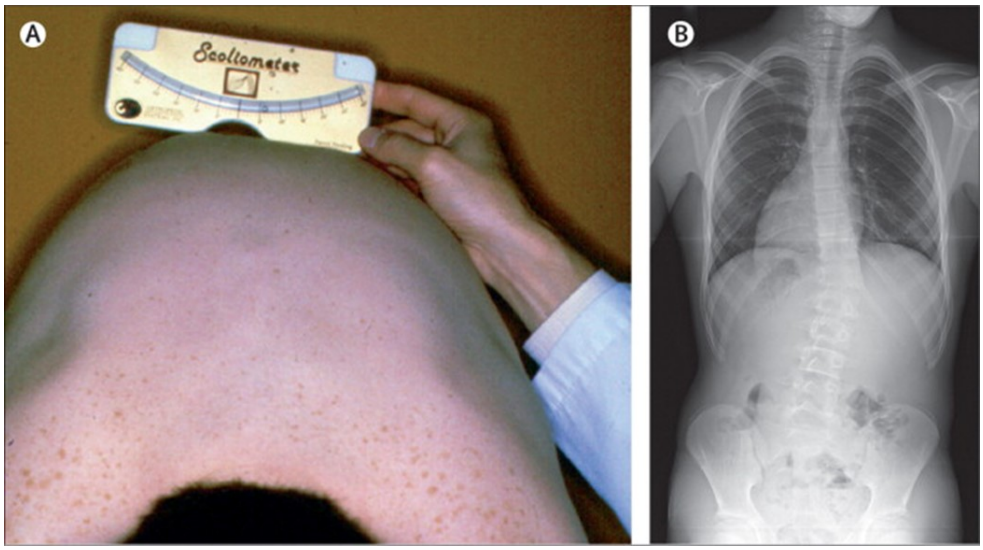


Fig. 2.5: Scoliosis screening using the scoliometer to measure rib hump inclination of a participant with 30° Cobb (Weinstein et al., 2008)

The limitations of the above-mentioned clinical tools for screening are the low specificity and high referral rates (Grivas et al., 2007; Negrini et al., 2018). A positive result from the test requires confirmatory radiography, which leads to unnecessary radiation exposure in participants who

ultimately have a negative diagnosis on the radiographic assessment (Negrini et al., 2018). In addition, the scoliometer tool measures the rib hump, which is only one index of posture, and does not consider the full 3D torso for scoliosis screening (Fortin et al., 2010).

The markerless ST technique can be an alternative tool for scoliosis screening. Deviation of the torso from normal symmetry can be analyzed to distinguish AIS from adolescents with typical spine development (Fig. 2.4). *Komeili et al.* compared ST asymmetry maps between healthy children and those with AIS (Komeili et al., 2014). The standard deviation (SD) of the deviation over the entire torso was 3.4 mm for healthy individuals; for patients with AIS, an SD of up to 8.3 mm was obtained. Thus, a threshold of ± 3 mm was proposed to indicate typical deviations not associated with scoliosis. The proposed deviation thresholds were estimated on a small sample size (n=5) (Komeili et al., 2014).

The markerless ST technique has been used to study asymmetry patterns in participants with typically developing spines. A study was conducted by *Ho et al.* to evaluate asymmetry patterns in non-scoliotic adolescents and to evaluate the relationship of torso asymmetries with factors such as age, gender, hand dominance, and physical activity (Ho et al., 2015). The surface torso of 83 participants was analyzed using the ST method. Based on the patterns observed in the DCM, participants are classified as having either twist or thickness asymmetry (Ho et al., 2015). In cases of twist asymmetry, the corresponding anterior and posterior regions of the torso have patches with opposing colors (Ho et al., 2015). In cases of thickness asymmetry, the corresponding anterior and posterior regions of the torso have patches with identical colors (Ho et al., 2015). Twist asymmetry reflects a lateral curvature or rotation in the spine associated with scoliosis (Komeili et al., 2014). The authors observed that all participants in the study had some degree of twist in their DCM (Ho et al., 2015). Twist asymmetries were also more prevalent in females, while thickness asymmetry was more evident in males (Ho et al., 2015). The difference in observed types of symmetries between genders reflects the higher prevalence of AIS in females. No association was found between the type of asymmetry and factors of age, hand dominance, and physical activity (Ho et al., 2015).

Convolutional Neural Networks (CNNs) have emerged as a powerful tool in medical image analysis, particularly in identifying conditions like AIS (K. Chen et al., 2021). Utilizing surface topography data to quantify torso surface asymmetries, CNNs can discern subtle differences

between healthy participants and those with AIS. CNN and other machine learning (ML) algorithms are increasingly used in spine research to predict suspected scoliosis, Cobb angle severity, and progression (K. Chen et al., 2021; Galbusera et al., 2019). *Jaremko et al.* were the first to predict rib rotation from spinal indices using ML algorithms (J. Jaremko et al., 2000). The authors noted the potential use of these algorithms to predict scoliosis spinal deformity from torso surface data (J. Jaremko et al., 2000; J. L. Jaremko et al., 2001).

Some studies have aimed to develop deep learning models to detect AIS from various surface topography systems. *Yang et al.* applied deep learning algorithms to detect scoliosis using images of unclothed individuals. Their method detected Cobb > 10° with a sensitivity of 87.5% and a specificity of 83.5% (Yang et al., 2019). Additionally, *Kobabu et al.* used CNN regression analysis to detect scoliosis from the degree of asymmetry on the back surface of individuals in a forward bending position. The model could predict Cobb>10° with a sensitivity of 79% and a specificity of 92% (Kokabu et al., 2021). *Watanabe et al.* applied CNN on Moire topography images to estimate the position of the 12 thoracic and five lumbar vertebrae to estimate the Cobb angle. The mean absolute error of their method to estimate typically developing spine (Cobb angles <10°) was 3.42 (Watanabe et al., 2019).

2.5 Conservative Treatments and Scoliosis Specific Exercises

Observation is the first step in managing a scoliosis patient, which consists of regular visits for clinical evaluation. The timing of the follow-ups often depends on the condition and the risk of progression of the scoliosis curve. For larger curves, bracing, a corrective orthosis, is prescribed to prevent progression. Observation and bracing are non-operative standard of care treatments of AIS as recommended by the Scoliosis Research Society (SRS) (*Diagnosis And Treatment | Scoliosis Research Society*, n.d.).

Physiotherapeutic Scoliosis-Specific Exercises (PSSE) represent a potential approach to managing idiopathic scoliosis (Negrini et al., 2018). These exercises can be tailored to each individual's unique spinal curvature and aim to improve posture, spinal alignment, and muscle balance. PSSE programs typically include a combination of stretching, strengthening, and breathing exercises, all designed to address the specific imbalances associated with scoliosis. Targeting the muscles surrounding the spine, PSSE can help stabilize the curvature, reduce pain, and potentially halt progression, particularly when started early in adolescence (Negrini et al., 2018).

Among the emerging PSSE programs in literature is the Schroth PSSE, first proposed by Katharina Schroth (Weiss, 2011). The purpose of the current development of the Schroth program is to de-rotate, elongate, and stabilize the spine. The therapy focuses on posture alignment and teaches individuals to be aware of their posture. Breathing exercises are also another focus of the therapy (Weiss, 2011). In a randomized controlled trial (RCT), *Kuru et al.* investigated the effect of Schroth PSSE on the change in Cobb angle. After six months of treatment, the average Cobb angle decreased by 2.5° (Kuru et al., 2016). Comparatively, in the control group, which was not prescribed PSSE, the Cobb angle deteriorated by 3.1° (Kuru et al., 2016). In an RCT by *Schreiber et al.,* significant improvement in the Cobb angle was observed in the Schroth group compared to the control group, decreasing curve severity by 1.2° over a six month period compared to the control group, where the Cobb angle increased by 2.3° (Schreiber et al., 2016). *Schreiber et al.* also showed that Schroth PSSE improved muscle endurance and SRS-22 pain and self-image domains (Schreiber et al., 2016).

Few studies, however, have reported the effect of scoliosis-specific exercises on external deformity and aesthetic outcomes. In a review of evidence on the efficacy of scoliosis-specific exercises, a lack of studies on evaluating cosmetic outcomes on PSSE was found (Romano et al., 2013). The RCT by *Kuru et al.,* is one of the few reports and found that trunk rotation and waist asymmetry were improved after a six month supervised Schroth program (Kuru et al., 2016).

2.6 Trunk Aesthetics Evaluation

According to SOSORT, aesthetic appearance is a primary concern and consideration in treating scoliosis (Negrini et al., 2006). Methods to measure and monitor aesthetics have been proposed, considering patients' self-assessment of aesthetic impact and objective methods of torso appearance assessments. Questionnaires are collected to determine the patient's subjective assessment of their own aesthetics. Suggested self-assessment of aesthetic appearance questionnaires are the SRS-22, the Walter Reed visual assessment scale, and the trunk appearance perception scale (TAPS) (Negrini et al., 2018).

A standard method of objectively evaluating aesthetics is through trunk asymmetry scales. Based on visual assessment of the shoulders, scapulae, and waist asymmetries, the trunk aesthetic clinical evaluation (TRACE) scale has been proposed as inexpensive and reproducible with fair intra-rater reliability (Zaina, Negrini, & Atanasio, 2009). However, the method has poor inter-rater reliability

(Negrini et al., 2018). The POTSI scale is an alternative method to assess the asymmetry of the trunk seen from the back, but may not be sensitive enough to use for scoliosis progression evaluation (Negrini et al., 2018). Similarly, the ATSI scale has been proposed to analyze the anterior trunk asymmetry (Stolinski et al., 2012). 2D digital photography has also been proposed to assess aesthetics. Photographs of the trunk with surface markers has been suggested to monitor posture, with fair to good correlations observed between 2D and radiographic indices (Furlanetto et al., 2016). The limitations of the asymmetry scales and 2D photographs are that the methods are a 2D representation of a 3D condition, and the placement of anatomical landmarks on the torso introduces human error (Furlanetto et al., 2016). Capturing the torso model to derive indices to describe shape asymmetries using surface topography instrumentation such as the Formetric and Quantec systems has also been suggested for aesthetics appearance assessment (Bidari et al., 2023). However, these systems are limited due to the high cost of the equipment.

2.7 Summary

The limitations of conventional approaches for scoliosis screening have been explored in the literature review. Patients are most concerned about their cosmetic appearance, which affects their self-image and quality of life. Moreover, further study is required to determine a valid method for evaluating trunk aesthetics, since current methods for assessing aesthetics have their limitations. Posture and aesthetic appearance may be improved with PSSE such as the Schroth method. The Schroth program has been shown to improve the Cobb angle and reduce progression. However, further research is needed to determine the effect of Schroth on posture outcomes. This research addresses these gaps by proposing and validating the ST technique for AIS detection and treatment efficacy evaluation, which may ultimately improve patient care and outcomes.

Chapter 3 : Adolescent Idiopathic Scoliosis Screening

A version of the following chapter will be submitted to Scientific Reports as:

Mohamed, N., Gonzalez Ruiz, J., Mostafa, H., Burke, T., Mei, Q., & Westover, L. Adolescent Idiopathic Scoliosis Screening: A 3D Markerless Surface Topography Approach with Convolutional Neural Networks

3.1 Summary

Introduction**:** Adolescent idiopathic scoliosis (AIS) is a three-dimensional lateral and torsional deformity of the spine, affecting up to 5% of the population. Traditional scoliosis screening methods exhibit limited accuracy, leading to unnecessary referrals and exposure to ionizing radiation through x-ray examinations. The 3D markerless surface topography (ST) technique, which measures the extent of trunk asymmetry, holds promise as a potential scoliosis screening tool. However, the variations in trunk asymmetry between individuals with scoliosis and those with a typically developing spine have yet to be thoroughly defined. This study seeks to distinguish adolescents with AIS from those with typically a developing spine using the ST method.

Methods: Participants aged 10 to 18 years, comprising 285 individuals with confirmed AIS and 273 with typically developing spines, were included in the study (total scans including follow-ups: 693 for the AIS group and 298 for the control group). The positive for AIS group was identified through radiographic exams, specifically with curves ranging from 10° to 45°, while the negative (control) group qualified if their scoliometer test measured less than 7° and they had no known scoliosis diagnosis. The dataset comprised surface torso scans captured either using stationary Minolta cameras or with the Structure sensor. ST analysis involved the reflection of the 3D geometry of the torso, aligning it with the original torso by minimizing the distance between corresponding points. Deviations between the original and reflected torso over the back surface and torso surface depth were mapped onto 102 x 102 grids. A Convolutional Neural Network (CNN) was developed using deviations and depth (distance between back surface and frontal plane) maps as inputs to classify the torso surface of typically developing adolescents and those with AIS. 10-fold cross-validation was applied during model development. 20% of the data was

used as a holdout for final testing. Classification results of the proposed model were compared to the ground truth.

Results: The average training and validation accuracy across the ten folds was 100% and 94%, respectively. The classifications from the testing sets using the best performing model from the 10-fold cross validation obtained accuracy, sensitivity, and specificity of 95%, 97%, and 90%, respectively. The positive likelihood ratio (PLR) of the testing set was 9.7. Likewise, A negative likelihood ratio (NLR) of 0.032 was also attained. The model sensitivity for detecting curves with Cobb greater than 25° was 99%. The sensitivity for detecting mild cases (Cobb < 25°) was 96%.

Conclusion: The proposed CNN predictive model to detect AIS using ST showed excellent classification results. Markerless surface topography can serve as a dependable and non-invasive method for screening AIS.

3.2 Introduction

Adolescent idiopathic scoliosis (AIS) is a three-dimensional spine deformity marked by a lateral deviation and often accompanied by vertebral rotation (Weinstein et al., 2008). The prevalence of AIS can be up to 5% and has a higher incidence in girls (Choudhry et al., 2016). AIS is diagnosed through radiographic examination, which involves taking a standing posterior-anterior x-ray to assess the curvature using the Cobb angle (Negrini et al., 2018). AIS is confirmed when the measured Cobb angle is at least 10° (Weinstein et al., 2008). Observation is recommended for curves less than 25°, and bracing is prescribed for curves between 25° and 45° (Bettany-Saltikov et al., 2017). Severe curves greater than 45° may require corrective spinal surgery (Bettany-Saltikov et al., 2017). Moreover, the likelihood of curve progression increases during puberty and the rapid growth phase (Negrini et al., 2018; Rogala et al., 1978).

Screening can play a crucial role in the early detection of scoliosis among adolescents, often performed during routine school screenings or in a clinical setting (Grivas et al., 2007; Torell et al., 1981). Detecting scoliosis at an early stage allows for timely intervention and management, which can prevent the spinal curve from worsening (Grivas et al., 2007; Weiss, Negrini, et al., 2006). However, controversies about screenings in schools exist. Screening programs have high false positive rates, which leads to excessive referrals for x-ray imaging that exposes children to unwarranted radiation (Fong et al., 2010; Morais et al., 1985; Negrini et al., 2018; US Preventive

Services Task Force, 2018). Further screening can involve stress from examinations and can lead to psychological side effects (US Preventive Services Task Force, 2018). However, the recommendations in favour of screening rely on evidence of moderate quality, whereas those opposing screening are founded on low quality evidence (Płaszewski & Bettany-Saltikov, 2014). The effect of discontinued screening programs have been evaluated in Canada, resulting in 32% of confirmed cases of AIS were considered late referrals requiring immediate treatment by bracing or surgery (Beauséjour et al., 2007).

The Adam's forward bend test is a standard and simple test where the examiner observes the back from the rear while the participant is bent forward at the waist, looking for any asymmetry, humps, or other irregularities in the contour of the spine, which could indicate a higher likelihood of the presence of scoliosis (Côté et al., 1998). The Adam's forward bend test is subjective and not recommended on its own (Negrini et al., 2018) because the results often depend on the experience of the examiner. The test has high sensitivity, ranging from 92% to 84.4%, but shows a high risk of false positives and a low specificity of 68% (Côté et al., 1998; Dunn et al., 2018).

The scoliometer is a handheld instrument used in the assessment of scoliosis to measure the degree of axial trunk rotation of the hump revealed by the Adam's test (Bunnell, 1984). The scoliometer measures the angle of trunk rotation (ATR) and adds an objective measure to the Adams test (Bunnell, 1984). When a cutoff ATR angle of 5° is used, it has a sensitivity of around 100% and a specificity of 47% (De Wilde et al., 1998; Grosso et al., 2002; Huang, 1997; Negrini et al., 2018). Specificity increases to 86% while sensitivity decreases to 83% with a cutoff ATR angle of 7° (De Wilde et al., 1998; Grosso et al., 2002; Huang, 1997; Negrini et al., 2018). Additionally, the use of multiple tests increases sensitivity and specificity (Dunn et al., 2018; Labelle et al., 2013; C. F. Lee et al., 2010).

Surface topography (ST) can be an alternative screening approach. The markerless ST asymmetry analysis method reflects the 3D geometry of the torso around the best plane of symmetry to identify the external areas of asymmetry in a deviation color map (DCM) image (Hill et al., 2014; Komeili et al., 2014, 2015a). ST has been previously shown to have a medium to strong agreement with estimating curve location and predicting the curve severity (Ghaneei et al., 2018; Hong et al., 2017). Markerless ST can be used as an innovative and non-invasive approach for scoliosis screening, providing a three-dimensional assessment of the back's surface without markers or

radiation that may improve upon traditional screening approaches. However, differences in trunk asymmetries between individuals with scoliosis and typically developing individuals are not fully understood.

Convolutional neural network (CNN) is a supervised machine learning algorithm commonly applied to computer vision tasks such as image classification, localization, detection, and segmentation (Goodfellow et al., 2016; Ker et al., 2018). The neural network algorithm learns and captures features from inputs while preserving spatial relationships (Ker et al., 2018). CNN processes grid-like topology data, such as image data, as a 2D or 3D grid of pixels (Goodfellow et al., 2016). In research on scoliosis, machine learning algorithms, specifically CNNs, have been implemented in radiological images to detect spinal curvature and to automatically measure the Cobb angle (Jamaludin et al., 2020; Pan et al., 2019). CNN algorithms have also been implemented to develop scoliosis screening methods relying on various input modalities linking surface torso indices and features to the spinal curve (Kokabu et al., 2021; Yang et al., 2019). Photographs of the back of the torso when participants are in a standing posture were used to predict spine deformity, however the model obtained lower specificity than the scoliometer, which increases the referral rate (Yang et al., 2019). Asymmetry analysis during the forward bend position was also considered to predict and detect scoliosis (Kokabu et al., 2021). However the sample size was small, and a lower sensitivity was obtained compared to the scoliometer test (Kokabu et al., 2021).

To improve the overall accuracy of identifying AIS and typical spine growth cases, the CNN approach may be applied for AIS detection based on the ST asymmetry technique. This study aimed to develop and evaluate the efficacy of a classification model aimed at AIS detection using the markerless ST asymmetry technique.

3.3 Methodology

3.3.1 Data Acquisition

The dataset in this study included torso surface scans of participants with radiographically confirmed AIS and volunteers with a typical spine development screened negative for AIS from the scoliometer test, yielding a total of 558 participants. The dataset contained 285 individuals diagnosed with AIS and 273 with typically developing spines. All participants in the study were between 10 to 18 years old. Participants with AIS had curves between 10° to 45° at baseline.

In addition to the baseline scan, a subset of participants with AIS underwent follow-up scans. The total number of surface scans for participants with AIS, including baseline and follow-up scans, amounted to 693. Similarly, some participants without AIS also had baseline and follow-up surface scans. The total number of surface scans for participants without AIS, including baseline and follow-up scans, amounted to 298. This study reports secondary analysis of data recruited and collected across multiple centers, specifically in Edmonton, Canada, where 782 scans were collected and Campo Grande, Brazil, where 209 scans were collected. Ethical approval, granted by the University of Alberta Health Research Ethics Board, was obtained with approval numbers Pro00118643 and Pro00117065. Participants provided consent in the original studies.

3.3.2 Surface Topography Analysis

The ST analysis technique, previously documented in the literature, involved obtaining surface scans of the entire torso to conduct markerless asymmetry analysis (Ghaneei et al., 2018; Hill et al., 2014; Hong et al., 2017; Komeili et al., 2014, 2015a). The data contained torso surface scans acquired using 4 Minolta cameras (Fig. 3.1a), capturing the torso's front, back, right, and left sides. The captured views of the torso were imported into Geomagic Control 2015 (3D Systems, North Carolina, USA) and merged. Additionally, the dataset contained surface scans captured using a structure sensor (Structure Sensor Pro) mounted on an iPad (Fig. 3.1a). With the handheld device, a circular motion was performed around the participants with the camera focused on the torso to obtain a 3D surface scan. The Minolta cameras were used to collect 782 scans, while the structure sensors were used to collect 209 surface scans. For data collected with the stationary cameras, participants were standing within a frame, with their arms positioned at 90-degree elevation. For scans that was collected with the Structure scanner, the frame was not used; participants were instructed to position the arms at 90-degree elevation by using sticks to maintain their arms in a consistently elevated position throughout the duration of obtaining the scan. All the dataset's scans were cropped to isolate and maintain a full 3D model of the torso.

For asymmetry analysis, the torso model (Fig. 3.1b) was duplicated and mirrored along the midsagittal plane. The reflected torso was aligned with the original to minimize the distance between the two models (Fig. 3.1c). The roto-inversion plane associated with this alignment was termed the plane of best symmetry (Hill et al., 2014).

Torso asymmetry was evaluated by obtaining the distances between each point on the original torso and its corresponding point on the reflected torso (Fig. 3.1d). With a specific focus on the back region of the torso, only 180 degrees of the back's span, with its deviation information, remained (Fig. 3.1e). The deviations and their corresponding position on the torso were mapped onto a grid of 102 columns by 102 rows (Fig. 3.1f). The depth, representing the distance between the back surface and the coronal plane, was also mapped onto a 102x102 grid (Fig. 3.1g). The deviation and depth grid maps were the input channels to a convolutional neural network (CNN).

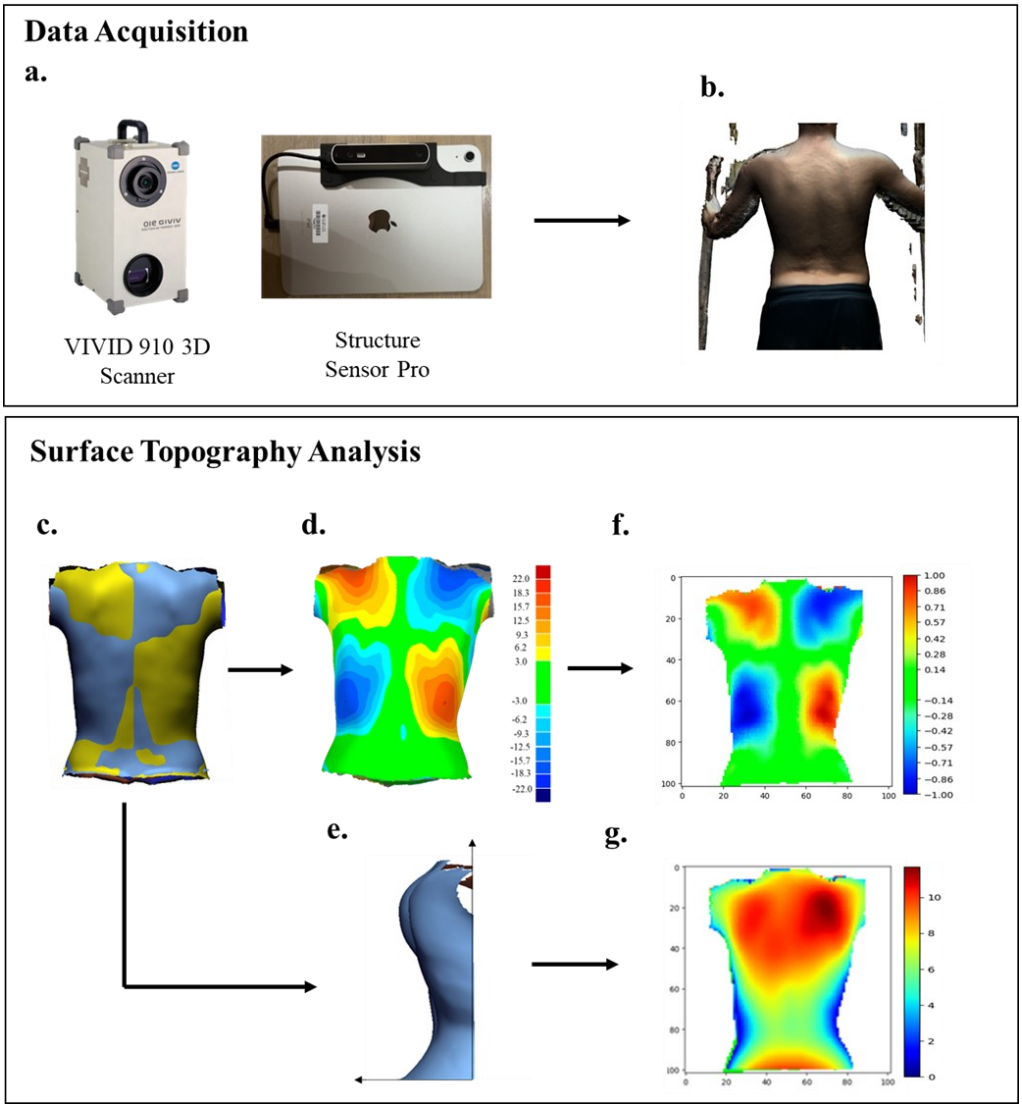


Fig. 3.1: Data collection procedure with (a) scanners to obtain (b) 3D trunk surface. ST analysis processing procedure showing (c) best fit alignment of the original and reflected

torsos, (d) deviation color map of the difference between original and reflected torsos, (e) back torso retained, representation of (f) back deviation and (g) surface torso depth maps

3.3.3 Convolutional Neural Network

The CNN architecture is represented in Fig. 3.2. The architecture consisted of multiple convolutional layers with max-pooling layers in between, then concluded in a fully connected layer for binary classification (Yamashita et al., 2018). A rectified linear activation function was applied at each convolutional layer. A sigmoid function was applied to the final output in the architecture. Network weights were updated using the Adam optimizer algorithm (Scherer et al., 2010; Yamashita et al., 2018). The outcomes of the CNN model resulted in an output prediction of 1 (positive for AIS) or 0 (negative for AIS). A hyperparameters optimization was conducted, and the procedure is outlined in Appendix A.

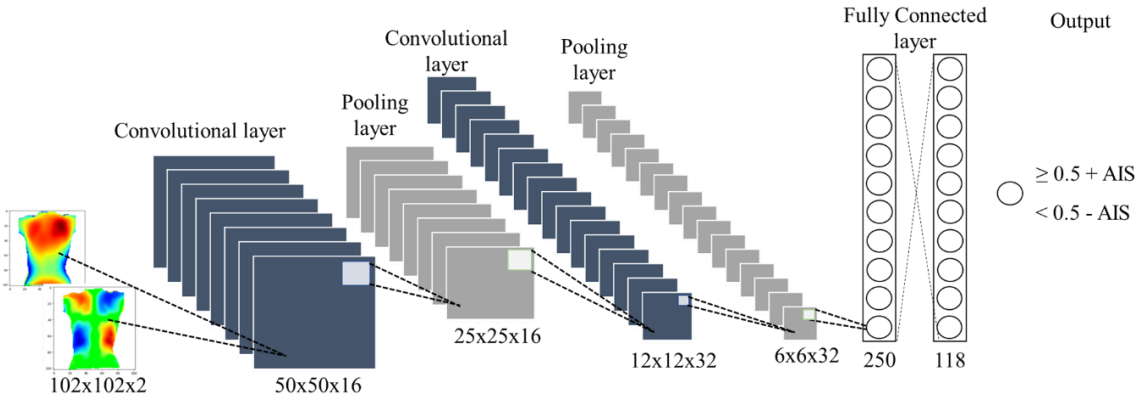


Fig. 3.2: Convolutional neural network (CNN) architecture

When developing the CNN model, a 10-fold cross-validation was performed. A testing set, which accounted for 20% of the data, was held out from all training for the final evaluation of the model. Data augmentation on the training and validation sets used during cross-validation was conducted to increase the dataset artificially. The data was increased by randomly rotating the input grid maps between -10 and 10 degrees in the coronal plane. The non-AIS group data was further increased by reflecting along the y-axis (medial-lateral). The number of epochs was limited to 500 during training. A batch size of 32 and a learning rate of 0.001 were also applied. Since the distribution between the AIS and control groups was unbalanced, the non-AIS group cases were oversampled during training to ensure a balanced number of cases in each batch. Specifically, random

oversampling was applied, where samples from the minority class (control group) were chosen at random to be duplicated and added to the training dataset.

The testing set containing controls and AIS data were randomly selected, with half of the chosen AIS scans exhibiting a Cobb angle of less than 25°, while the remaining half of the AIS group featured Cobb angles greater than 25°. This stratification allowed for a thorough examination of the algorithm's performance across a spectrum of scoliosis severity.

3.3.4 Statistical Analysis

Sensitivity, which is the ability of the model to correctly identify AIS, was determined and calculated from the prediction outcomes of true positives (TP) and false negatives (FN) (Table 3.1). Likewise, specificity, which is the ability to correctly identify negative cases, was determined. The true negatives (TN) and false positives (FP) outcomes are used to calculate specificity (Table 3.1). The model performance was also evaluated using the false positive rate and false negative rate. Additional measures to evaluate the model were the positive likelihood ratio (PLR) and the negative likelihood ratio (NLR) (Table 3.1). The testing set not used during training was used to evaluate the model's performance. Of the 10 models produced during cross-validation, the model with the best accuracy was selected and evaluated in this study.

Table 3.1 Confusion matrix outcome variables comparing prediction classifier with ground truth to determine model performance. Sensitivity, specificity, false positive rate, false negative rate, positive likelihood ratio, and negative likelihood ratio were determined from the confusion matrix to evaluate the overall performance of the model

CNN Classifier Using ST	**Ground Truth**	
	+ AIS (Confirmed with Radiograph)	**- AIS (Confirmed with Scoliometer <7°)**
+ AIS	True Positive (TP)	False Positive (FP)
- AIS	False Negative (FN)	True Negative (TN)

Sensitivity = TP/(TP + FN)
Specificity = TN/(TN + FP)
False positive rate = FP/(FP + TN)
False negative rate = FN/(FN + TP)
Positive likelihood ratio = Sensitivity/(1 - Specificity)

Negative likelihood ratio = (1 - Sensitivity)/ Specificity

A receiver-operating characteristic (ROC) analysis was conducted to determine the ability of the classification model to discern ST maps between positive AIS and negative AIS cases. The ROC curve is obtained by plotting the sensitivity over the false positive rate at various model output thresholds. From the ROC curve, the area under the curve (AUC) is determined, which measures the overall accuracy of the classifier compared to the ground truth. An ideal AUC would be 1.0 with excellent sensitivity and a negligible false positive rate (Nahm, 2022).

3.4 Results

Sample distribution of training and validation sets during 10-fold cross-validation and the testing set are shown in Table 3.2. Participants' age, height, weight, and Cobb angle of the largest curve were similar between training + validation and testing sets. Following data augmentation of the cross-validation dataset, the samples were increased from 796 samples to 1860.

Table 3.2: Training, Validation, and testing set characteristics (mean ± standard deviation)

	Training + Validation		**Testing**	
	+ AIS	**- AIS**	**+ AIS**	**-AIS**
Sample Size before data augmentation	558	238	139	60
Sample size following data augmentation	1146	714	139	60
Age (years)	13.3 ± 2.0	13.1 ± 2.6	13.7 ± 2.0	14.4 ± 1.8
Height (m)	1.58 ± 0.1	1.59 ± 1.3	1.60 ± 0.1	1.62 ± 0.1
Weight (kg)	48.3 ± 11.8	51.3 ± 13.1	49.0 ± 11.7	54.3 ± 15.7
Cobb largest curve (°)	27.3 ± 11.4	n/a	26.4 ± 10.3	n/a
Scoliometer (°)	n/a	3.35 ± 1.7	n/a	2.9 ± 1.5

Average diagnostic accuracy of 100% and 94% were obtained from the 10-fold training and validation sets, respectively. Additional results are shown for each cross-validation fold in Table

3.3. The average testing set across the 10-folds yielded an accuracy of 93%, a sensitivity of 96%, and a specificity of 87% (Table 3.4).

Table 3.3: Training and validation errors (dimensionless) and accuracy comparing CNN outputs to the ground truth for each cross-validation fold

Cross-Validation Fold	Training Set Error	Training Set Accuracy	Validation Set Error	Validation Set Accuracy
1	0.0003	99.9	0.032	92.3
2	0.0003	100.0	0.026	94.0
3	0.0006	99.3	0.020	95.6
4	0.0003	99.8	0.023	94.5
5	0.0004	99.6	0.027	92.9
6	0.0003	100.0	0.023	94.5
7	0.0003	99.9	0.036	91.8
8	0.0003	99.9	0.034	92.9
9	0.0002	100.0	0.022	95.1
10	0.0002	99.9	0.018	96.2
Average	**0.0003**	**99.8**	**0.026**	**94.0**

Table 3.4: Performance assessments from the testing set of the CNN model obtained after each cross-validation fold

Cross Validation Fold	True Positive (TP)	False positive (FP)	True Negative (TN)	False Negative (FN)	Accuracy	Sensitivity	Specificity
1	134	7	53	5	94.0	96.4	88.3
2	137	9	51	2	94.5	98.6	85.0
3	134	9	51	5	93.0	96.4	85.0
4	133	7	53	6	93.5	95.7	88.3
5	134	8	52	5	93.5	96.4	86.7
6	136	10	50	3	93.5	97.8	83.3
7	135	6	54	4	95.0	97.1	90.0
8	125	10	50	14	87.9	89.9	83.3
9	134	6	54	5	94.5	96.4	90.0
10	136	7	53	3	95.0	97.8	88.3
				Average	**93.4**	**96.3**	**86.8**

The model selection process involved choosing the fold with the highest training accuracy. The best-performing model during cross-validation was evaluated with the testing set presented in Fig. 3.3.

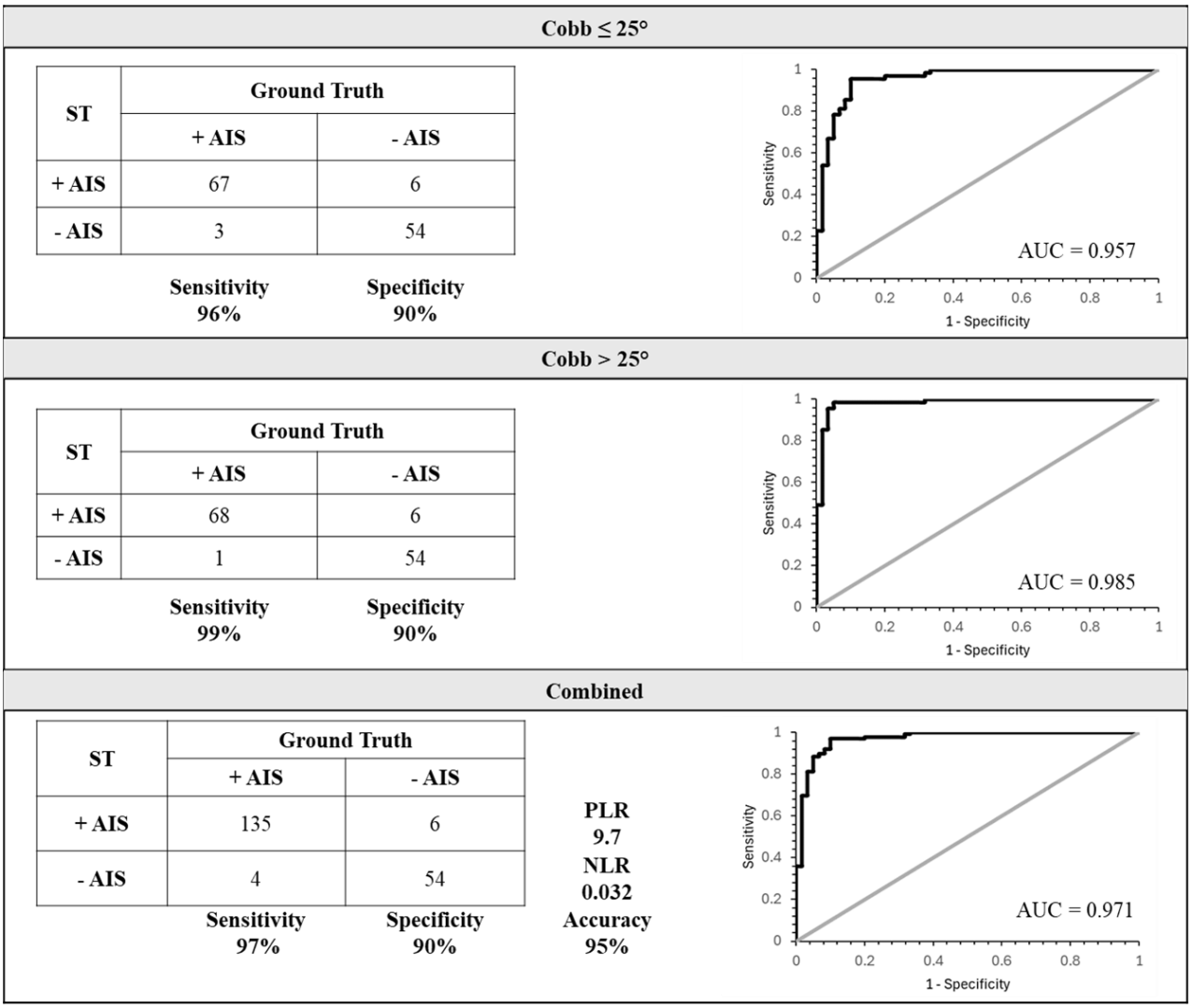


Fig. 3.3: Confusion matrices with corresponding ROC curve comparing CNN classification algorithm using ST with ground truth of the combined data, mild severity data (Cobb ≤ 25°) and moderate severity data (Cobb > 25°)

Model prediction of participants with curves less than 25° was compared to the ground truth (Figure 3). The sensitivity of detecting curves less than 25° was 96%. From the ROC curve, an AUC of 0.957 (95% CI 0.922;0.992) was obtained comparing the CNN classifier outputs to the ground truth. Participants with curves greater than 25° were also compared to the ground truth,

and a sensitivity of 99% was achieved. The AUC for the data with only moderate curves (Cobb > 25°) was 0.985 (95% CI 0.965;1.00). The sensitivity of the combined AIS data was 97%, with a false negative rate of 2.9%. Consequently, when comparing the model prediction of non-scoliosis participants against the ground truth, a specificity of 90% was obtained, with a false positive rate of 10%. The AUC of the combined data obtained from ROC analysis was 0.971 (95% CI 0.945;0.996). Diagnostic accuracy for the combined data in Figure 3.3 of 95% was obtained. A positive likelihood ratio (PLR) of 9.7 and a negative likelihood ratio (NLR) of 0.032 were also determined.

3.5 Discussion

The purpose of this study was to develop a classification model and to investigate the viability of markerless ST as a screening tool for AIS. A CNN-based binary classification model was developed using deviations and depth information of the back surface of the torso. Classification results from the CNN model indicate high overall accuracy in distinguishing between individuals with and without scoliosis based on torso asymmetry, yielding a classification accuracy of 95%. A sensitivity of 97% reflects the model's robust ability to identify adolescents with scoliosis correctly, minimizing false negatives. Additionally, a specificity of 90% highlights the model's proficiency in correctly recognizing non-scoliotic cases, minimizing false positives.

The achieved positive likelihood ratio of 9.7 indicates that the odds of obtaining a positive result in individuals with AIS are approximately 10 times higher than those without the condition, underscoring the diagnostic strength of the CNN model using the markerless ST technique in predicting the presence of AIS. The low negative likelihood ratio of 0.032 suggests that the odds of individuals with AIS obtaining a negative result are nearly negligible, emphasizing the high reliability of the method in correctly identifying AIS cases. These likelihood ratios further validate the robustness of the proposed approach, emphasizing its potential as a reliable screening tool with strong discriminative capabilities of AIS detection. Ultimately, The CNN algorithm shows promise in discriminating between the positive AIS class and the negative AIS class using the markerless ST technique.

Our results exhibit notable advantages compared to conventional screening modalities like the scoliometer and Adam’s forward bend test. The model based on ST maps demonstrated a lower false positive rate of 10% compared to the scoliometer test, with a false positive rate of 19.3%

(Dunn et al., 2018). The reduction in the false positive rate indicates a decrease in unnecessary x-ray referrals, which exposes children to ionizing radiation. While these findings suggest a promising approach for scoliosis screening, conducting a comprehensive validation study directly comparing the CNN model with the scoliometer is crucial. This further validation is essential to substantiate the observed improvements, ensuring the reliability and generalizability of the ST approach and would contribute valuable insights into the practical implications and potential adoption of our proposed method in clinical settings. Additionally, the proposed model using ST analysis is automatic and less prone to human error, while the accuracy of the standard tools depend on the experience of the operator.

Studies have evaluated the use of torso asymmetries to predict scoliosis. *Kokabu et al.* proposed an automatic system using a depth sensor to scan individuals back in Adam's forward bend position to detect AIS (Kokabu et al., 2021). The proposed method relies on determining the asymmetry index based on the deviations between the reflected and original point cloud of the back following the best alignment algorithm. The cutoff value of the asymmetry index to predict Cobb angle greater than 15° was predicted based on receiver operating characteristic (ROC) analysis. The authors reported that the cutoff value yielded an accuracy of 84% to predict cases with Cobb greater than 15°. The authors also obtained a sensitivity of 79% and a specificity of 92% (Kokabu et al., 2021). However, their study did not analyze the cutoff value for predicting Cobb greater than 10°, which is the minimum angle to be classified as scoliosis. The study included younger participants aged 7 to 18 years old. The sample size was much smaller in *Kokabu et al.'s* study compared to the present analysis, with 76 total participants suspected of having scoliosis based on an x-ray examination.

Yang et al. developed a deep-learning algorithm to screen for scoliosis using unclothed back images (Yang et al., 2019). The binary classification algorithm for detecting scoliotic images from non-scoliosis images exhibited an average area under the curve (AUC) of 0.946 from an external validation set (scoliotic images n=300, normal controls n=100). The authors reported a sensitivity of 87.5% and a specificity of 83.5% (Yang et al., 2019). This study focused exclusively on unclothed back photographs for classifying scoliosis and non-scoliosis cases. In contrast, our approach incorporated the ST technique to evaluate back asymmetry comprehensively. In addition,

our study also considered depth information of the back to capture variations in spatial dimensions of the torso.

Chowanska et al. used Moiré topography to compute the surface trunk rotation parameter as a scoliosis screening method (Chowanska et al., 2012). However, their study's findings reported unsatisfactory sensitivity and specificity (Chowanska et al., 2012). The primary challenge stemmed from the difficulty in determining an optimal surface trunk rotation parameter cutoff value, ultimately resulting in diminished screening accuracy (sensitivity of 77.4% and a specificity of 71.1%) (Chowanska et al., 2012). Notably, their study did not employ a deep learning algorithm to distinguish between scoliosis and non-scoliosis images. This absence of advanced computational techniques might have contributed to the limitations in classification accuracy, highlighting the potential advantages of integrating machine learning methods in enhancing the precision of scoliosis screening.

The results from the proposed screening model using ST were comparable to those evaluating torso asymmetries to detect and classify AIS. The overall accuracy of the CNN classifier (95%) was greater than that of the accuracy results using the asymmetry index of participants in the forward bend position (84%) proposed by *Kokabu et al*. Our accuracy performance (sensitivity of 97%, specificity of 90%) was also greater than that of the deep learning model developed by classifying photographs of participants' back torso surfaces (sensitivity of 87.5%, specificity of 83.5%), which were proposed by *Yang et al*. Likewise, the *Chowanska et al*. classification threshold to screen for scoliosis had lower accuracy (sensitivity of 77.4%, specificity of 71.1%) than the model proposed in this study.

Several limitations merit consideration in interpreting the findings of this study. First, the sample composition introduces potential bias, with a disproportionate representation favoring adolescents with AIS. This skewness may impact the generalizability of the results to a broader population. Another constraint is related to the torso scans, which were cropped around the posterior superior iliac spine. The exclusion of the pelvis introduces a limitation as the pelvis rotation, not accounted for during scanning due to clothing constraints, could contribute to the underlying curvature. Furthermore, the absence of pelvic data may overlook critical factors leading to uneven hips in individuals with scoliosis. Additionally, the study acknowledges a limitation in sample size, especially concerning the application of CNN algorithms and other neural networks, which

typically require larger datasets to mitigate the risk of overfitting. These limitations highlight the need for cautious interpretation and future refinement to enhance the robustness and applicability of the proposed model.

3.6 Conclusion

The study aimed to ascertain the viability of markerless ST as a screening tool for Adolescent Idiopathic Scoliosis (AIS). The results presented in this paper provide promising evidence supporting the effectiveness of this method in assessing torso asymmetry and identifying individuals with scoliosis using a CNN algorithm. Our findings demonstrate the potential of markerless surface topography as a reliable and non-invasive means of screening for AIS. The significance of using the ST method in clinics and community screening is that it can enhance the efficiency and accuracy of scoliosis detection, allowing for earlier diagnosis. Early detection is critical, as it can lead to timely treatment and potentially prevent the condition's progression, thereby improving patient outcomes.

Chapter 4 : Scoliosis Screening Validation

4.1 Introduction

In Chapter 3 of this thesis, the ST technique was proposed as a screening tool to detect AIS. The pattern and magnitude of the deviations of the torso obtained from ST analysis were used to predict positive AIS or negative AIS. A convolutional neural network model was developed to classify the torso asymmetries and depth into positive or negative AIS categories, achieving an accuracy of 95%. The model described in Chapter 3 obtained a sensitivity of 97%, which is the ability to detect scoliosis. Likewise, the model obtained a specificity of 90% on an external testing set, which is the ability to identify those who do not have scoliosis and exclude them from radiographic referrals.

The effectiveness of clinical screening tools has previously been published. The Adam's forward bend test had a sensitivity and specificity of 84.4% and 68%, respectively (Côté et al., 1998; Dunn et al., 2018). Likewise, the scoliometer had a sensitivity of 100% and specificity of 47%, using a cutoff of 5° ATR (De Wilde et al., 1998; Grosso et al., 2002; Huang, 1997). For a cutoff ATR of 7°, Sensitivity and specificity were 83% and 86%, respectively (De Wilde et al., 1998; Grosso et al., 2002; Huang, 1997). The proposed CNN model results were comparable to standard clinical tools, where the model based on ST has been shown to reduce false positive rates and increase specificity. The prevalence of AIS across different studies and populations does not affect the sensitivity and specificity of the screening tools (de Vet et al., 2011). However, the sensitivity and specificity can be influenced by the differences in the severity of the scoliosis curve in the population (de Vet et al., 2011). Thus, there is a need to validate and directly compare the proposed model using ST asymmetries of the torso with standard clinical tools.

The study's purpose was to validate the CNN model classifier's accuracy and performance using the ST technique to detect AIS and directly compare them with the Adam's forward bend and scoliometer clinical screening tests. We present preliminary findings in this study from an ongoing screening program designed to evaluate the effectiveness of the standard clinical tools and the developed ST approach to detect AIS.

4.2 Methods

Seventy participants were recruited in the local town of Torrelavega, Spain, specifically at sporting clubs. The inclusion criteria for the study were children aged between 9 and 18 years. Ethical approval was obtained from the University of Alberta Health Research Board (Pro00135870). Written informed consent from the parents and assigned consent from the participants were obtained. The participants and parents received a written explanation of the project before obtaining their consent. Results from the screening tests were communicated to the families and were recommended to consult their pediatrician if the screening tests yielded a positive test result.

4.2.1 Measurements

Demographic data are collected for each participant, which includes age, height, weight, and hours per week spent on sports activities. The Adam's forward bend test was then conducted. Participants were instructed to lean forward when standing, with hands in front of their knees, elbows straight, feet apart, shoulders relaxed, and knees stretched. In this bending position, the rib profile is examined. A positive result is indicative of observing a prominent rib hump. In the same bending pose, the scoliometer reading was collected. The device was placed on the thoracic and lumbar rib hump, and the ATR value stemming from the inclination of the scoliometer was recorded. An ATR greater than 5° indicates a positive test result. The Adam's test, along with the scoliometer reading, was conducted by an experienced physiotherapist.

A full surface scan of the torso was collected using a structure sensor (Structure Sensor Pro) mounted on an iPad. Participants were standing upright, feet apart, and arms extended slightly. Applying the proposed classification model using ST analysis required some preprocessing steps, detailed in Chapter 3 of this thesis. Briefly, the 3D model was cropped around the posterior superior iliac spine (PSIS) level; the arms were cropped around the armpits to retain only the torso (Hill et al., 2014; Komeili et al., 2015a). The ST analysis minimized the distance between the torso and the reflected copy about the sagittal plane (Hill et al., 2014; Komeili et al., 2014, 2015a). The deviations between the original and reflected torso were determined for only the back of the torso. The depth, the distance from the back to the frontal plane, was also collected. The deviations and depth cloud point data were decimated to map the information onto a 102 x 102 grid. The depth and deviation grid maps were used as inputs for the CNN model developed to classify and predict positive and negative AIS. Preprocessing of the torso scans to obtain the results from the model

was conducted by a research team member, blinded to the results from the Adam's forward bend test and scoliometer reading.

4.2.2 Statistical Analysis

Sensitivity and specificity were determined for each test based on their classification outcomes and the actual spine condition of the participants. Specificity will be determined from the ratio between true negative cases over the total number of participants who don't have scoliosis. Likewise, the sensitivity will be determined from that ratio between true positive over the total number of participants with scoliosis. The number of participants with AIS will be confirmed with x-rays. The likelihood ratios between the screening tests will also be compared. A comparison between the sensitivity and specificity of the different screening outcomes will be assessed using McNemar's test, often used in paired binary observations (Kim & Lee, 2017). The null hypothesis of McNemar's test is that the screening performance (sensitivity and specificity) is not better than the other. Likewise, the alternative hypothesis is that there is a difference between the different screening tests. Cohen's K coefficient will also assess the agreement between the Adam's forward bend, scoliometer, and ST method.

4.3 Results

Seventy participants were assessed in February 2024. The descriptive data of the participants are shown in Table 4.1. The average age of the participants was 12.32 ± 2.6 years old. The average BMI of the sample was 19.92 kg/m². The proportion of females was 26%, and among them, 8 reported menarches.

Table 4.1: Sample demographic characteristics with values represented as mean ± standard deviations for continuous variables and percentages for categorial variables

Variables	Sample (n=70)
Sex N(%)	
Female	19 (27)
Male	51 (73)
Age (yrs)	12.32 ± 2.6
Height (m)	1.55 ± 0.2
Weight (kg)	48.95 ± 16.3
Body Mass Index (BMI, kg/m²)	19.92 ± 3.8

Of the 73 assessed participants, parents of three of the participants reported previous diagnosis of AIS and were excluded from the study. Twenty-two participants were screened positive from the tests performed (Adam's forward bend test, scoliometer, ST). Among those screened positive, 18 (26%) were screened positive using the Adam's test, 5 (7%) were screened positive from the scoliometer reading, and 4 (7%) were screened positive from the ST approach. Participants with positive screening results from at least two tests were referred for x-ray examination. The referral rate was found to be 7/70 = 10%. Of the participants referred for radiographic assessment, one was confirmed to have scoliosis, where the participant was detected positive only through ST. One participant was found to have a normal spine from an x-ray examination and screened positive from both Adam's and scoliometer tests. Radiographic assessment results are unknown for the other four participants.

A 2x2 contingency table comparing the screening result using the ST approach with the Adam's and scoliometer test is shown in Table 4.2 for the complete sample, and Table 4.3 participants assumed not to have scoliosis, and individuals with confirmed scoliosis.

Table 4.2: Observed outcomes from ST, compared to Adams's forward bend, and scoliometer tests of the complete sample (n=70). Positive screening result indicated by (+) and negative screening results indicated by (-)

ST	Adam's Test		Scoliometer Test	
	-	+	-	+
-	48	17	60	5
+	3	2	4	1

The number of confirmed AIS cases needed higher to determine sensitivity across the different screening modalities. Therefore, only the specificities were compared between tests. The specificity of the Adam's, scoliometer and ST methods were estimated to be 76%, 96%, and 96%, respectively. Comparing the outcomes using the ST approach to the Adam's test, an agreement of 75% was obtained, and a kappa coefficient of 0.060 was obtained (slight agreement). McNemar's test determined a statistically significant difference in specificity between the ST approach and the Adams's test (p=0.006). Comparing the ST approach to the scoliometer test, a proportion agreement of 94% and a kappa coefficient of 0.113 (slight agreement) were obtained. McNemar's

test determined no statistically significant difference in specificity between the ST approach and the Adam's test (p=0.617).

Table 4.3: Observed outcomes from ST, compared to the Adams's forward bend, and scoliometer tests of the sample (n=64), excluding participants with confirmed scoliosis from x-ray exams and those awaiting results. Positive screening result indicated by (+) and negative screening results indicated by (-)

ST	Adam's Test		Scoliometer Test	
	-	+	-	+
-	48	14	60	2
+	2	0	2	0

The following Fig. 4.1, Fig. 4.2, and Fig. 4.3 show deviation color maps of participants screened positive using the ST method, Adam's test, and scoliometer, respectively. The figures note that some participants screened positive for more than one test (e.g., participant 037 screened positive for Adam's test, scoliometer test, and ST approach).

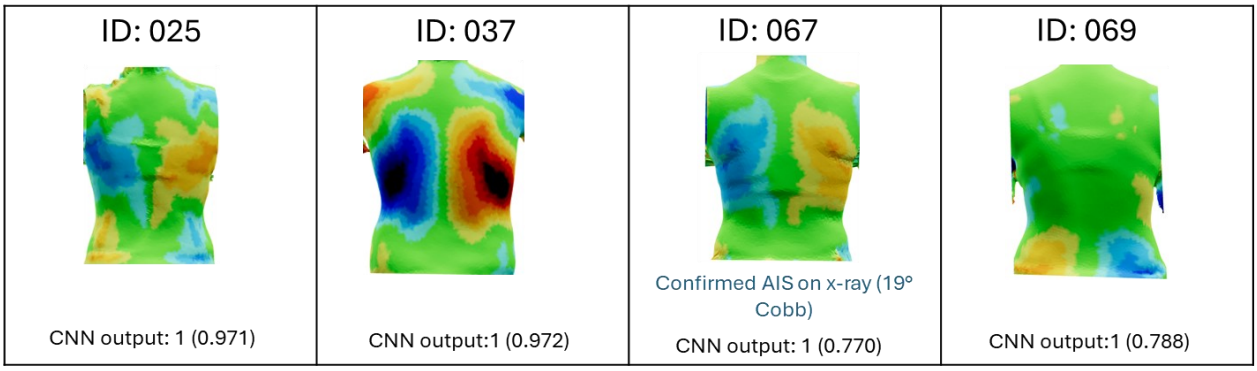


Fig. 4.1: Deviation color maps of participants screened positive with ST method. The CNN model output of each case with its sigmoid probability is reported

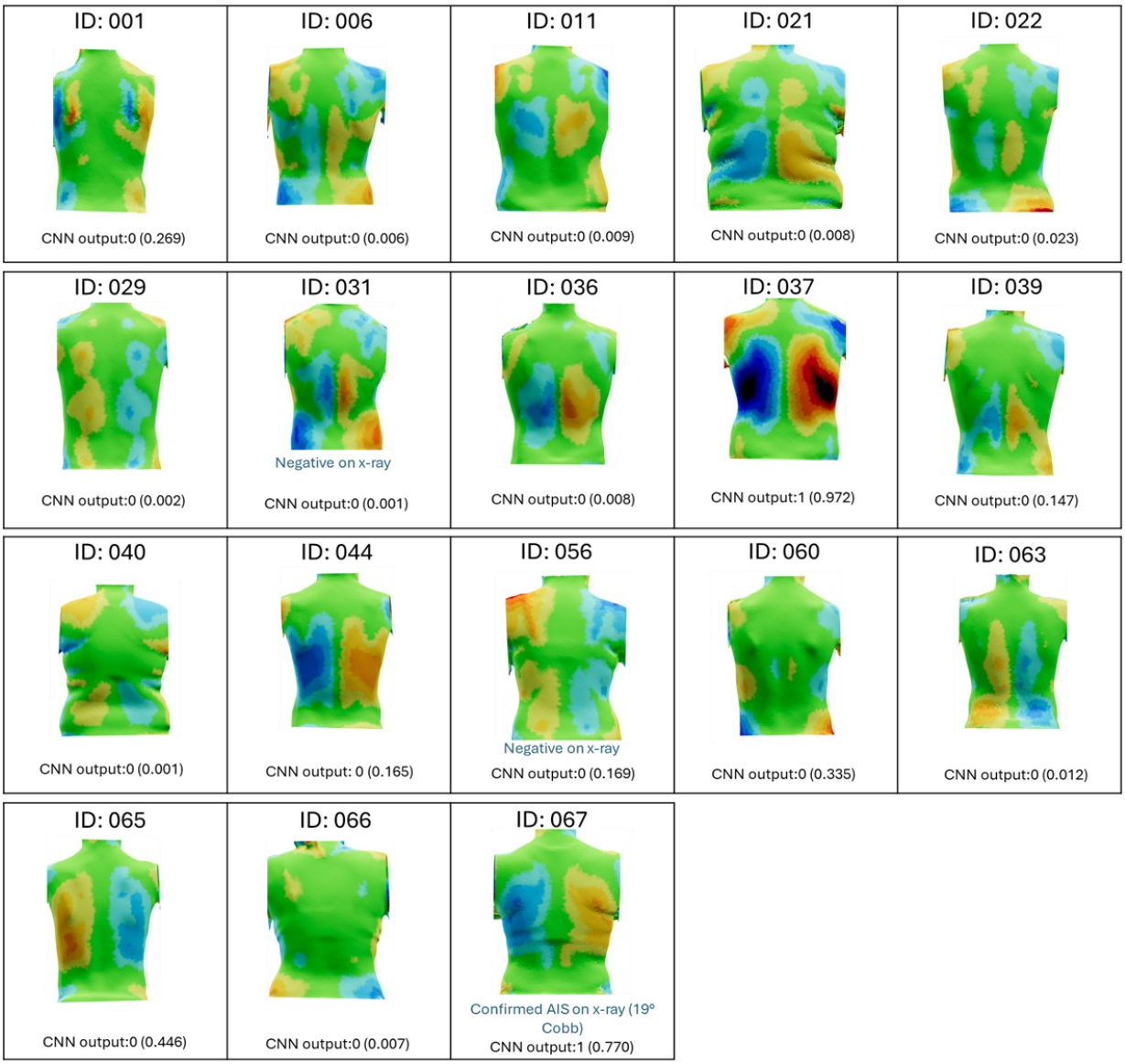


Fig. 4.2: Deviation color maps of participants screened positive with Adam's test. The CNN model output of each case with its sigmoid probability is reported

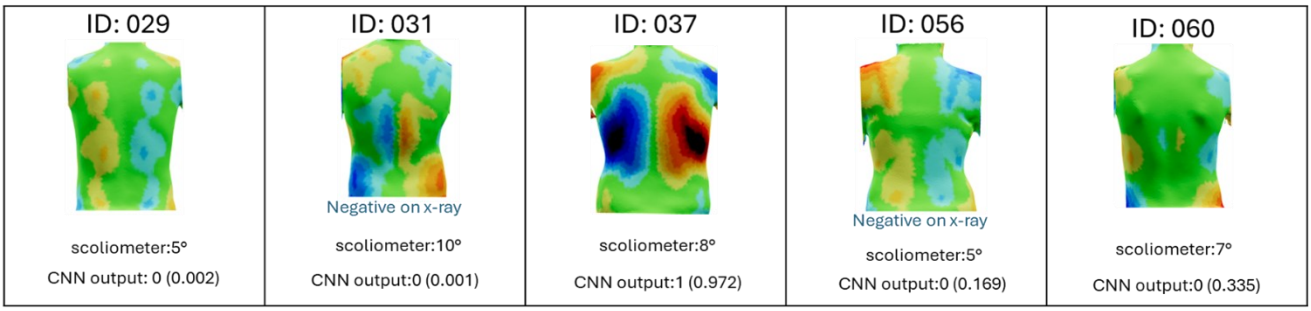


Fig. 4.3: Deviation color maps of participants screened positive with the scoliometer test. The CNN model output of each case with its sigmoid probability is reported

4.4 Discussion

This study aimed to compare screening outcomes using standard clinical tools such as the Adam's forward bend test and scoliometer tests with the ST approach. Preliminary results from screening 70 participants were reported.

Nineteen participants were screened from Adam's test. The prevalence of AIS in the Spanish population was estimated to be 7% (Cheng et al., 2015). From our sample size, the number of positive predictions from the Adams test suggests a clear over-referral rate. Adam's test is not recommended independently; our results reflect that assessment. As shown in Fig. 4.2, The ST maps of some participants screened positive from Adam's test do not suggest observed asymmetry maps of a participant with AIS. Deviations maps of participants with AIS show distinct color patches; blue patches represent negative deviation deviations, which reflect an outward deformation, and orange color patches represent positive deviations, which reflect an inward deformation. Distinct patches are often found in the thoracic, thoracolumbar, or lumbar area. More than one deviation patch is also expected and depends on the number and location of the spinal curve. Additionally, focusing on one side of the plane of sagittal symmetry, the patches alternate in their signed deviations if multiple patches are observed on the torso. Some of the participants screened positive from Adam's test did not show patterns usually associated with AIS. In the notable examples in Fig. 4.2, participants 001, 006, 029, 031, 039, 056, and 063 showed deviation patterns inconsistent with scoliosis deviations. Participants who were screened positive from the scoliometer test are shown in Fig. 4.3. The deviation maps of participants 031, 056, and 060 did not show the usual patterns associated with scoliosis. Our observations were validated since participants 031 and 056 did not have AIS, as confirmed by x-ray assessments. Fig. 4.1 showed

participants who were screened positive from the proposed ST approach. Participants 025, 037, and 067 did show distinct patches often associated with AIS, though only participant 067 was confirmed to have AIS from x-ray assessments.

Only specificity was compared across the different tests since there were insufficient children with confirmed scoliosis to evaluate the sensitivity of the screening tests. Likewise, the positive and negative likelihood ratios cannot be known in this early study sample. The prevalence of AIS was 2.7% from the current sample size. The estimated specificity (96%) using the ST method was found to be higher to the specificity of the testing set (90%) during the development stage of the model described in Chapter 3. This suggested the model classifying positive and negative cases of AIS generalized well to a new set of data from another population.

The CNN model based on the ST technique revealed higher specificity than the Adam's test alone. The specificity of the ST techniques with the scoliometer were equal. There was a significant difference in the specificity between ST and Adam's test, suggesting that the CNN model using ST was able to exclude non-scoliotic participants for radiographic assessment and reduce false positive rates. Additionally, no significant differences in specificity were observed between the ST method and the scoliometer test, suggesting that both tests perform similarly in excluding participants for x-ray exam referrals. The preliminary results showed that using the scoliometer increased specificity compared to the Adam's test and agreed with several reported literature on the effectiveness of the standard scoliosis screening tests.

The sensitivity and specificity of the Adam's and scoliometer tests for scoliosis screening depend on the experience and skill of the rater. In this study, an experienced physiotherapist conducted the screening tests. The ST approach, however, can be conducted by individuals with less experience in scoliosis screening since the method relies only on capturing the surface scan and cropping to retain the torso. The data preprocessing, followed by the ST analysis and the application of the CNN model to obtain the classification results, is automatic and less prone to human error.

A slight agreement between ST and the standard clinical tests was observed in this sample. The strength of agreement between the Adam's and scoliometer tests was categorized as fair based on the kappa values. The kappa agreement values were low despite the relatively high agreement between the ST method and the scoliometer (75% and 94%, respectively). The classification outcomes of the screening test are based on different aspects of the torso. The ST test is based on

patterns identified in the deviation map between the reflected and original torso in a standing posture. The scoliometer measures the inclination of the rib hump, and the Adam's test is a subjective test examining the asymmetry of the right and left sides of the back during the forward bend. The different approaches to examining asymmetries may affect the agreement between the tests using Cohen's kappa.

Participants recruited in this preliminary study were 73% male and 27% female. However, girls are more likely to have scoliosis and therefore further recruitment of participants in a screening study should reflect this fact. Additionally, the CNN classifier based on ST analysis was trained on data with the majority being from female participants (>90%). Further analysis should be conducted following the completion of the screening study. Follow-up of the referred cases for radiographic assessments should be conducted to determine the true positives, false positives, true negatives, and false negatives outcomes from the screening tests. For ethical reasons, not all recruited participants were asked to get an x-ray. Therefore, some assumptions were made: the participants not referred for x-rays were assumed not to have scoliosis. However, the referral rate in this sample was relatively high (9.6%), which increased confidence in not missing scoliosis cases during the screening program. In school screenings, referral rates have should be around 2% to 3% (Bunnell, 2005).

4.5 Conclusion

This study aimed to validate the surface topography (ST) approach detecting AIS and compare with standard clinical tools. Preliminary findings indicate that the ST method exhibits higher specificity compared to the Adam's forward bend test and comparable specificity to the scoliometer. The ST method's ability to reduce false positives and exclude non-scoliotic participants from further radiographic assessments suggests its potential utility in clinical settings, especially given its reduced reliance on the operator's experience. Further research is necessary to validate these findings in a larger sample size and to assess the sensitivity of each screening method, as the current study was limited by the small number of confirmed AIS cases. Classifying torsos based on asymmetry maps using ST analysis, facilitated by a CNN algorithm, holds potential as a reliable and efficient tool for scoliosis screening in school-based and clinical environments.

Chapter 5 : Aesthetic Appearance Assessment

A version of the following chapter will be submitted to The Spine Journal as:

Mohamed, N., Gonzalez Ruiz, J., Schreiber, S., Parent, E., Mei, Q., & Westover, L. Responsiveness and Minimally Important Changes in Surface Topography Asymmetry Parameters to Global Ratings of Change in Adolescents with Idiopathic Scoliosis in the Schroth Exercise Trial for Scoliosis

5.1 Summary

Introduction: Adolescent idiopathic scoliosis (AIS) affects up to 5% of adolescents, causing spinal curvature and functional limitations. Traditional assessment using the Cobb angle may only partially capture patient concerns about aesthetics. The Schroth physiotherapeutic scoliosis-specific exercises (PSSE), focusing on posture correction, has shown promise for reducing the Cobb angle and preventing progression, though research on its cosmetic impact is limited. Surface topography (ST) offers a radiation-free alternative to assess AIS, but determining its meaningful change for patients remains unknown. This study explores the association between perceived back improvement and ST measurements in AIS treatment.

Methods: This secondary analysis stems from a randomized controlled trial evaluating the impact of a six months Schroth PSSE added to standard care on AIS patients' curve severity, quality of life, and torso asymmetry outcomes. 124 participants were recruited from the Edmonton Scoliosis Clinic and randomized into standard care (observation and bracing) and Schroth intervention groups (one-hour weekly supervised session, 30-40 minutes daily home exercises). Global Rate of Change (GRC) was self-reported after six months, and asymmetry parameters of root mean square (RMS) and maximum deviation (MaxDev) were obtained through surface torso scans captured at baseline and six months. Pearson correlation and receiver-operating characteristic (ROC) curve analysis were performed to determine the Minimal Important Changes (MICs) and the accuracy of the asymmetry parameters with perceived improvement. Subgroup analyses based on curve type were also conducted to ascertain MICs for thoracic and lumbar curves.

Results: Significant correlations were observed between changes in RMS ($r=-0.510$, $p<0.001$) and MaxDev ($r=-0.409$, $p<0.001$) and perceived improvement in back condition. Participants who reported improved (GRC≥2) posture saw a 1.76 ± 2.9 mm and 3.29 ± 6.5 mm decrease in RMS and MaxDev, respectively. In contrast, RMS and MaxDev increased by 1.03 ± 3.0 mm and 1.26 ± 5.6 mm, respectively, among individuals who stated that their posture had deteriorated or had not changed (GRC < 2). Using ROC analysis, MICs for RMS and MaxDev were determined to be -0.27 mm (area under the curve (AUC) 0.746, sensitivity 67%, specificity 74%) and -0.49 mm (AUC 0.717, sensitivity 64%, specificity 68%), respectively, for overall improvement perception. Having met both thresholds reduced sensitivity to 62% and a specificity of 74%. Subgroup analysis based on curve types (thoracic vs. lumbar) revealed alternative MIC thresholds. MICs for thoracic curves were -0.58 mm (AUC 0.618, sensitivity 60%, specificity 53%) for RMS and 1.32 mm (AUC 0.632, sensitivity 73%, specificity 92%) for MaxDev. MICs for lumbar curves were -0.26 mm for RMS (AUC 0.881, sensitivity 73%, specificity 92%) and -0.61 mm for MaxDev (AUC 0.811, sensitivity 68%, specificity 83%) for MaxDev.

Conclusion**:** The study aimed to explore the correlation between surface topography (ST) parameters and patients' self-reported perception of their back condition, focusing on adolescent idiopathic scoliosis (AIS) conservative treatment with Schroth PSSE. Results showed significant correlations between changes in root mean square (RMS) and maximum deviation (MaxDev) with patients' Global Rate of Change (GRC) scores, indicating that changes in RMS and MaxDev align with perceived improvements in back condition. Stronger associations were observed between ST parameters and perceived improvement in lumbar curves. The study underscores the clinical relevance of ST parameters, particularly RMS and MaxDev, as objective measures for assessing treatment efficacy and patient-reported outcomes in AIS management.

5.2 Introduction

Adolescent idiopathic scoliosis (AIS) is a paediatric condition affecting up to 5% of the population, with a higher prevalence in females (Choudhry et al., 2016; Weinstein et al., 2008). The scoliosis spine exhibits a three-dimensional lateral deviation and rotation (Roach, 1999). Children with AIS can experience breathing difficulties, back pain, limited physical function, and poor self-image (J. Danielsson & L. Nachemson, 2003; Kan et al., 2023; Lau et al., 2024; Sperandio et al., 2014; Stone et al., 2023; Upasani et al., 2008). The Cobb angle indicator is used to quantify the severity of the

spinal curvature, measured from 2D posterior-anterior radiographic images, and is the gold standard method for AIS assessment (Thulbourne & Gillespie, 1976; Weinstein et al., 2008). In general, observation is conducted for curves with a Cobb angle less than 25°, bracing is prescribed for a Cobb angle between 25° to 45°, and corrective spine surgery is recommended for a Cobb angle greater than 45° (*Diagnosis And Treatment | Scoliosis Research Society*, n.d.; Komeili et al., 2015a; Weinstein et al., 2008). The Scoliosis Research Society (SRS) identifies the Cobb angle as the primary outcome for monitoring and treating AIS (Richards et al., 2005). The effectiveness of standard-of-care treatments is determined by limiting curve progression to 5° or less (Carman et al., 1990; Roach, 1999). However, aesthetic appearance is crucial for patients and parents (Bridwell et al., 2000; Cheshire et al., 2017; Ponseti et al., 1976). The International Society on Scoliosis Orthopaedic and Rehabilitation Treatment (SOSORT) suggests that the primary objective in treating patients with AIS is prioritizing aesthetics (Negrini et al., 2018). Monitoring the change in Cobb angle ranks behind aesthetics, quality of life, disability, back pain, physiological well-being, progression in adulthood, and breathing function (Negrini et al., 2018).

Exercise can improve aesthetic outcomes (Aulisa et al., 2014; Negrini et al., 2006, 2018). The Schroth method is a physiotherapeutic scoliosis-specific exercises (PSSE) aimed at correcting posture through stretching, strengthening, and breathing techniques (Fusco et al., 2011; Weiss, 2011). An important component of the Schroth method is auto-correction, which is the application of active postural realignment of the spine in 3D to reduce spinal deformity (Fusco et al., 2011). The Schroth PSSE has shown in published literature to reduce progression and Cobb angle and improve muscle endurance (Kuru et al., 2016; Otman et al., 2005; Schreiber et al., 2015, 2016). More evidence of the efficiency of the Schroth program in terms of aesthetic and cosmetic outcomes needs to be provided (Romano et al., 2013).

The Cobb angle obtained from radiograph assessments has several limitations. The outcome measured from a 2D posterior-anterior view does not fully capture the 3D nature of the spinal curvature of AIS (Ramirez et al., 2006; Thulbourne & Gillespie, 1976). Research indicates that radiological parameters correlate poorly with patients' subjective perception of body image (Parent, Wong, et al., 2010; Smith et al., 2006). Additionally, frequent radiographs are required to monitor the curve progression, which can lead to an increased risk of developing cancer (Levy et al., 1996; Nash et al., 1979).

Thorough research on aesthetics and cosmetic outcomes needs to be conducted, as less than 3.5% of the literature on AIS discusses this topic (Negrini et al., 2006). This gap is partly attributed to the limited methodologies for assessing aesthetics and posture outcomes (Negrini et al., 2006, 2018). Patient reported self-perception of their back condition during treatment has drawn more attention in recent years, and several evaluation instruments have been created in this regard. Questionnaires were designed to evaluate perceived aesthetic deformity experienced by individuals with scoliosis (Asher et al., 2003). Theses patient-reported outcome measures focus on assessing the subjective perception of aesthetics. Alternatively, a quantitative and objective aesthetics assessment can be achieved using tools like back photography, trunk asymmetry scales or surface topography (ST) systems (Furlanetto et al., 2016; Komeili et al., 2015a; Stolinski et al., 2012; Zaina, Negrini, & Atanasio, 2009).

A markerless ST has been developed to characterize the trunk asymmetry of patients with AIS without ionizing radiation (Hong et al., 2017; Komeili et al., 2014, 2015a, 2015b). The ST method is not intended to substitute the gold-standard radiograph measurements. Instead, in conjunction with radiographs, ST can enhance the management of scoliosis and reduce the risk associated with radiation exposure by limiting the frequency of x-ray scans (Hong et al., 2017). The ST markerless technique considers the entire 3D torso, quantifies the trunk asymmetry affected by AIS and presents a visual using a deviation colormap map (DCM) (Hill et al., 2014; Komeili et al., 2014). Areas of asymmetries are identified from the DCM, and extracted parameters from the cluster are used to predict severity and progression (Hong et al., 2017; Komeili et al., 2014). These parameters include the maximum deviation (MaxDev) within an asymmetry patch at the level of each curve and the root mean square (RMS) of the deviations measured in that patch (Ghaneei et al., 2018; Hong et al., 2017; Komeili et al., 2014, 2015a, 2015b).

It is unclear how much change in ST measures is needed for patients to perceive an improvement in their back condition. While treating AIS, patients can experience changes in their surface back asymmetries, varying in magnitude, with some being subtle and others more pronounced. However, not all changes are recognized or valued by patients. The minimal change in the measurement system that patients consider significant is represented by the minimally important change (MIC) (Jaeschke et al., 1989; Revicki et al., 2008). Clinicians rely on MIC to assess the effectiveness of an intervention and guide their decision-making (Revicki et al., 2008). There is a

lack of information concerning the MIC of ST parameters, limiting its use as a tool to evaluate aesthetic outcomes.

This study aimed to determine the association between perceived improvement in back status and changes in ST asymmetry measurements in patients treated for AIS and the ability of ST asymmetry measurements to detect patients perceiving improvements in response to treatment.

5.2 Methods

5.2.1 Data Acquisition

This is a secondary analysis from a randomized controlled trial (RCT). The RCT was conducted to evaluate the effect of a six month Schroth PSSE program added to the standard of care on curve severity, quality of life, and torso asymmetry outcomes of patients with AIS (Mohamed et al., 2024; Schreiber et al., 2015, 2016, 2019). 124 participants were recruited from the Edmonton Scoliosis Clinic. The cohort was divided into two groups: the standard of care group (n=60) and the Schroth group (n=64).

The Schroth group received six months supervised Schroth PSSE intervention, which included individual weekly sessions overseen by a trained Schroth therapist and daily 30- to 45-minute at-home exercises. The Schroth intervention was prescribed in addition to the standard of care (observation or bracing). A detailed description of the Schroth intervention was previously published (Schreiber et al., 2014; Watkins et al., 2012). The control group only received standard of care in the first six months but was offered six months of Schroth intervention in the latter half of a one-year follow-up.

The criteria for including potential participants were as follows: those between 10 and 18 years of age at the beginning of the study, Cobb angle between 10° and 45°, Risser grade between 0 to 5, and no prior corrective spine surgery. The study was approved by the University of Alberta Health Research Board (pilot study approval Pro00011552, extended study approval Pro00043397).

5.2.2 Measurements

5.2.2.1 Global Rate of Change

The Global Rate of Change (GRC) is a quick and effective evaluation to determine how patients perceive their health condition over time and whether they believe their condition has worsened,

had little change, or improved (Kamper et al., 2009). The GRC has been shown to have good test-retest reliability and validity (Kamper et al., 2009). This study used the 15-point GRC scale to ascertain the participants' impression of their back condition. The GRC scaled from -7 (a great deal worse) to +7 (a great deal better). Intermediate values of -1 represented "a tiny bit worse," 0 was "about the same," and +1 was "a tiny bit better." After six months, participants were asked using the GRC scale: "Please rate the overall condition of your back from the time you began the treatment until now." The radiograph findings were not disclosed to the participants beforehand, and the GRC was completed before a therapy session or without any treatment. Based on the GRC ratings, two groups were defined: improved (+2 little bit to +7 very great deal better) and not improved, which represents no change or a deteriorated back condition (+1 tiny bit better to -7 great deal worse) (Schreiber et al., 2019).

5.2.2.2 Asymmetry Parameters

Surface torso scans were obtained at baseline and six months using four stationary VIVID 910 3D laser Minolta scanners, capturing the torso's front, back, and side views. The patients' arms were maintained at 90 degrees of shoulder elevation as they stood in a frame. The four views were merged to obtain a 3D model. The region below the posterior superior iliac spine, the head, neck, and shoulders from the scapula were cropped out, followed by noise filtering and smoothing of the torso (Komeili et al., 2015a, 2015b). The ST technique involved mirroring the 3D cloud of the model over the sagittal plane. This reflected model was then aligned with the original model. Using the iterative least-squares method, the reflected torso coordinates are transformed to minimize the distance between the reflected and original torsos (Fig. 5.1) (Hill et al., 2014; Hong et al., 2017; Komeili et al., 2015a, 2015b).

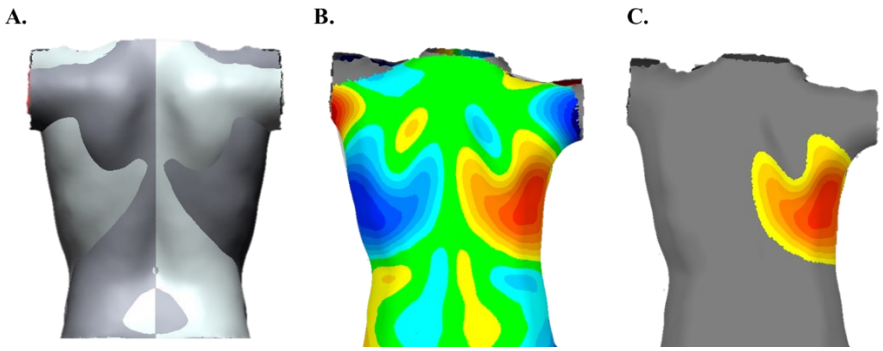


Fig. 5.1: ST analysis depicting A) best fit alignment of original (silver) and reflected torso (grey), B) deviation color map and C) isolated patch for obtaining RMS and MaxDev parameters

The asymmetries were represented using a deviation colour map (DCM) obtained from the distance between the original and reflected torsos. A deviation of ± 3 mm was deemed typical (green in Fig. 5.1b) and applied in the DCM to reveal patch areas of asymmetries caused by the spinal curvature (shades of yellow/red and blue in Fig. 5.1b and Fig. 5.1c) (Komeili et al., 2015a). In addition, a threshold of 9.33 mm was used in certain circumstances (i.e. when the maximum deviation in a patch was > 9.33) to improve the separation of asymmetry patches (Ghaneei et al., 2018). ST parameter measurements were obtained using the patch corresponding to the location of largest curve. Root mean square (RMS) and Maximum deviation (MaxDev) were determined from the isolated patch. ST measurements were computed from the same patch at baseline and six months scans. Finally, the change in ST parameters between scans was recorded.

Lastly, participants' curve patterns were classified using the Schroth classification algorithm (Schreiber et al., 2012; Watkins et al., 2012). The four Schroth curve types categorize scoliosis spinal curve patterns into two thoracic types (3c and 3cp) and two thoracolumbar/lumbar types (4c and 4cp). The 3c curve type is characterized by a primary thoracic curve, may have a minor or no lumbar curve, and a balanced pelvis. The 3cp curve features a single thoracic curve with an unbalanced pelvis. The 4c curve is defined by a main lumbar curve with a thoracic curve and a balanced pelvis. The 4cp curve type represents a primary thoracolumbar or lumbar curve with an unbalanced pelvis and may have thoracic curve (Schreiber et al., 2012).

5.2.3 Statistical analysis

Pearson correlation analysis was conducted to assess the strength of the correlation between the change in ST parameters and the GRC scores. The parameter change with a significant correlation to the GRC was identified and further analyzed to determine the MICs. The anchor method was applied to estimate the MICs of the ST parameters. The anchor method is a common and well-established approach that links an outcome measure to an external criterion (anchor) that distinguishes participants with important improvement or deterioration from participants with no important change (de Vet et al., 2011). In this study, the GRC was selected as the anchor to determine the minimum change in ST parameters that corresponds to perceived improvement of overall back condition. A receiver-operating characteristic (ROC) curve was used to compare the change in ST parameters and the GRC grouped (as improved and not improved) by plotting the sensitivity and false positive rate at various decision thresholds. The ROC curve was used to determine the MIC threshold for each parameter with the best balance of sensitivity and false positive rate. The ST parameter threshold change represents the cutoff value that best distinguishes between improved and not improved (no change or deteriorated). The threshold value is determined by selecting the point nearest to the top left corner of the curve on the plot of sensitivity over the false positive rate. The ST thresholds' diagnostic accuracy with the ground truth of perceived improvement based on GRC was also examined. Curve type-based subgroup analyses were also carried using the Schroth curve type classification algorithm (Schreiber et al., 2012). ROC analysis of the lumbar and thoracic curves was performed to ascertain the MICs for each kind of curve.

5.3 Results

80 participants made up the sample size from the recruited 124 participants in the study, after excluding 12 who did not finish the study and 32 who had missing data. Intention-to-treat analysis was not conducted since it was important to determine the MICs of the ST parameters of participants who followed the protocol as intended. Descriptives of the sample study are shown in Table 5.1. Participants had a mean age of 13.3 ± 1.7 years at baseline and ST parameters RMS and MaxDev of 10.6 ± 4.8 mm and 15.8 ± 8.2 mm, respectively. Treatment distribution in the sample size was as follows: 13 (16% of the sample) participants were under observation, 26 (33%) participants were wearing a brace, 16 (20%) participants underwent Schroth therapy, and 25 (31%)

participants were wearing a brace and had Schroth intervention. In addition, 8 (10%) participants had 3c curve types, 26 (33%) participants had 3cp curves, 10 (13%) participants had 4c curves, and 36 (45%) participants had 4cp curves.

Descriptives of the thoracic (3c or 3cp curves) and lumbar (4c or 4cp curves) subgroup are presented in Table 5.1.

Table 5.1: Baseline characteristics for the whole sample and split by curve type

Characteristic	Total Sample		Thoracic 3c/3cp Subgroup		Lumbar 4c/4cp Subgroup	
Sample Size	80		3c (8;10%)	3cp (26;33%)	4c (10;13%)	4cp (36;45%)
Age (years)	13.3 ±1.7		13.4 ±1.7		13.2 ±1.7	
Height (m)	156.8 ± 9.5		158.1 ± 7.3		155.7 ± 10.8	
Weight (kg)	45.2 ± 9.1		46.4 ± 8.0		44.3 ± 9.9	
Cobb angle (°)	27.9 ± 8.9		28.8 ± 8.7		27.3 ± 9.1	
RMS (mm)	10.6 ± 4.8		11.9 ± 5.0		9.7 ± 4.6	
MaxDev (mm)	15.8 ± 8.2		17.3 ± 7.8		14.7 ± 8.3	
Prescribed Treatment	Observation (13;16%)	Exercise alone (16;20%)	Observation (6;18%)	Exercise alone (7;21%)	Observation (7;15%)	Exercise alone (9;20%)
	Braced alone (26;33%)	Braced + Exercise (25,31%)	Braced alone (11;32%)	Braced + Exercise (10,29%)	Braced alone (15;33%)	Braced + Exercise (15,33%)

The average change in RMS and MaxDev was -0.43 ± 3.2 mm and -1.13 ± 6.4 mm, respectively. The mean GRC was 2.11 ± 3.1. Pearson correlation was obtained between GRC scores and the change in ST parameters (Table 5.2). The correlation between GRC and ΔRMS was significant ($r=-0.510$, $p<0.001$, Fig. 5.2). Likewise, the correlation between the GRC and ΔMaxDev was significant ($r=-0.409$, $p<0.001$, Fig. 5.2). Based on the self-reported GRC scores, 42 participants reported improved overall back condition and 38 participants reported deterioration or no change over the six months time period. RMS and MaxDev for participants who reported improved posture decreased by 1.76 ± 2.9 mm and 3.29 ± 6.5 mm, respectively. Alternatively, RMS and MaxDev for those who reported deteriorated or no change in posture increased by 1.03 ± 3.0 mm and 1.26 ± 5.6 mm, respectively.

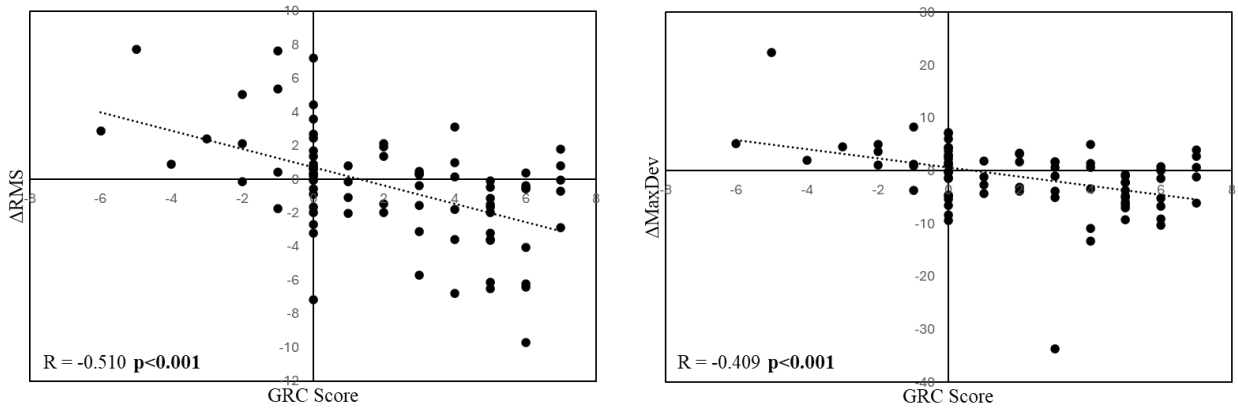


Fig. 5.2: GRC scores and the corresponding change in RMS and MaxDev with the correlation R value reported

Table 5.2: Correlation coefficients for the association between the change in ST parameters (ΔRMS, ΔMaxDev) with patient reported GRC scores

	Complete Sample (N=80)		Thoracic 3c/3cp Subgroup (N=34)		Lumbar 4c/4cp Subgroup (N=46)	
ST Parameter	**Pearson Correlation**	**P-value**	**Pearson Correlation**	**P-value**	**Pearson Correlation**	**P-value**
ΔRMS	-0.510	**<0.001**	-0.390	**0.023**	-0.495	**0.010**
ΔMaxDev	-0.409	**<0.001**	-0.453	**0.007**	-0.562	**0.003**

The anchor method using ROC analysis to estimate MIC was conducted for ΔRMS and ΔMaxDev. The ROC curve, displaying sensitivity and specificity for various changes in RMS and MaxDev are shown in Fig. 5.3. The area under the curve (AUC) was 0.746 (95%CI 0.638;0.853) for ΔRMS and 0.717 (95% CI 0.605;0.829) for ΔMaxDev (Table 5.3).

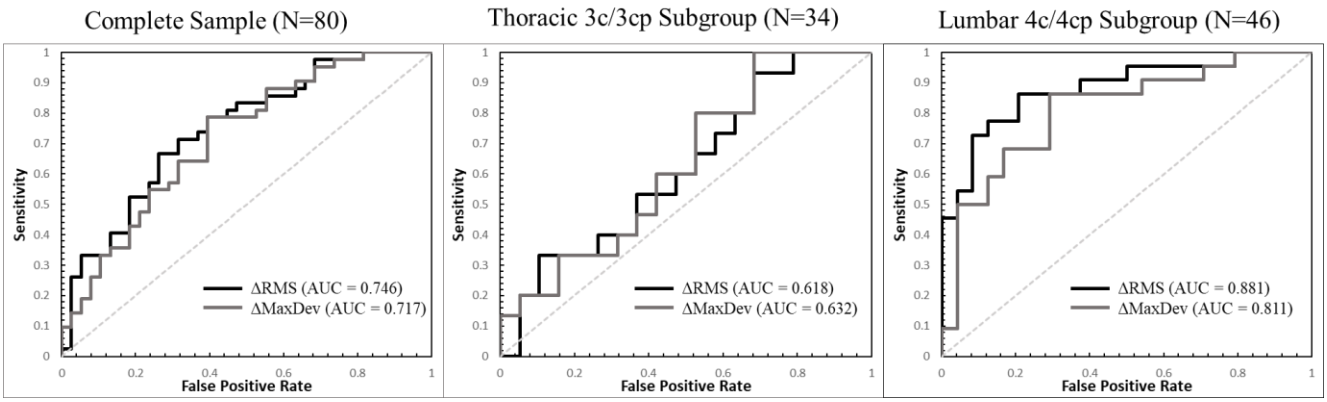


Fig. 5.3: Receiver operating characteristics (ROC) curves of patient perceived improvement (GRC ≥ 2) at various cutoff points of change in RMS and MaxDev

The estimated RMS MIC from the ROC curve had a sensitivity of 67% and a specificity of 74%. This corresponds to a change of RMS with the best ability to detect participants perceiving improvement of -0.27 mm. The estimated MaxDev MIC from the ROC curve had a sensitivity of 64% and a specificity of 68%, corresponding to a decrease in MaxDev by 0.49 mm to detect perceived improvement. Improvements in meeting both thresholds for surface asymmetry were also examined. A minimal decrease of 0.27 mm for RMS 0.49 mm for MaxDev showed an accuracy, sensitivity, and specificity of 68%, 62%, and 74%, respectively (Table 5.4).

Table 5.3: Responsiveness results and minimally important changes (MIC) of ST parameters from receiver operating characteristic (ROC) analysis from the complete group and curve type analyses

	Complete Sample (N=80)	**Thoracic 3c/3cp Subgroup (N=34)**	**Lumbar 4c/4cp Subgroup (N=46)**
Mean GRC	2.11 ± 3.1	1.88 ± 2.7	2.28 ± 3.3
RMS parameter			
ΔRMS (mm) Perceived Improvement (GRC ≥2)	-1.76 ± 2.9	-1.74 ± 2.4	-1.77 ± 3.1
ΔRMS (mm) Perceived No change/Deteriorated (GRC<2)	1.03 ± 3.0	-0.36 ± 3.0	2.42 ± 2.4

AUC	0.746 (0.638;0.853)	0.618 (0.427;0.808)	0.881 (0.777;0.973)
MIC (mm)	-0.27	-0.58	-0.26
MaxDev parameter			
ΔMaxDev (mm) Perceived Improvement (GRC ≥2)	-3.29 ± 6.5	-3.18 ± 4.5	-3.35 ± 7.4
ΔMaxDev (mm) Perceived No change/Deteriorated (GRC<2)	1.26 ± 5.6	0.17 ± 6.8	2.34 ± 3.9
AUC	0.717 (0.605;0.829)	0.632 (0.444;0.820)	0.811 (0.667;0.931)
MIC (mm)	-0.49	-1.32	-0.61

5.3.1 Analysis by Curve Types

Analysis of subgroups based on curve types was conducted. Participants with 3c and 3cp curve types, which are thoracic curves without and with pelvic imbalance, respectively, were combined into one subgroup. Participants with 4c and 4cp curve types, which are lumbar curves without and with pelvic imbalance, were combined into one subgroup. The ROC curve for participants with thoracic curves (3c or 3cp) and lumbar curves (4c and 4cp) are shown in Fig. 5.3.

For participants with thoracic curves, the correlation between GRC and ΔRMS was significant (r=-0.390, p=0.023). Likewise, the correlation between the GRC and ΔMaxDev was significant (r=-0.453, p=0.007) (Table 5.2). RMS and MaxDev for participants who reported improved posture decreased by 1.74 ± 2.4 mm and 3.18 ± 4.5 mm, respectively. Alternatively, the change in RMS and MaxDev for those who reported deteriorated or no change in posture was -0.36 ± 3.0 mm and 0.17 ± 6.8 mm, respectively. The AUC was 0.618 (95% CI 0.427;0.808) for ΔRMS and 0.632 (95% CI 0.444;0.820) for ΔMaxDev (Table 5.3). From the ROC analysis, a minimal decrease of 0.58 mm for RMS was selected, yielding a sensitivity of 60% and a specificity of 53%. In addition, a minimal decrease of 1.32 mm for MaxDev was selected, yielding a sensitivity of 60% and a specificity of 53%. Sensitivity and specificity of 58% and 59% are obtained when the thoracic patches decrease in both RMS and MaxDev below the thresholds (Table 5.4).

For participants with main lumbar curves, the correlation between GRC and ΔRMS was significant (r=-0.495, p=0.010). Likewise, the correlation between the GRC and ΔMaxDev was significant

(r=-0.562, p=0.003) (Table 5.2). RMS and MaxDev for participants who reported improved posture decreased by 1.77 ± 3.1 mm and 3.35 ± 7.4 mm, respectively. Alternatively, the change in RMS and MaxDev for those who reported deteriorated or no change in posture was 2.42 ± 2.4 mm and 2.34 ± 3.9 mm, respectively. The AUC was 0.881 (95% CI 0.777;0.973) for ΔRMS and 0.811 (95% CI 0.667;0.931) for ΔMaxDev (Table 5.3). The ROC analysis selected a minimal decrease of 0.26 mm for RMS, yielding a sensitivity of 73% and a specificity of 92%. In addition, a minimal decrease of 0.61 for MaxDev was selected, yielding a sensitivity of 68% and a specificity of 83%. Sensitivity and specificity of 68% and 92% are obtained when the thoracic patches decrease in both RMS and MaxDev below the thresholds (Table 5.4).

Table 5.4: Accuracy results of classifying perceived improvement and no change/deteriorated (GRC <2) using RMS and MaxDev MICs for the entire sample, thoracic curves group, and lumbar curves group

Complete Sample (N=80)			
	ΔRMS ≤ -0.27	**ΔMaxDev ≤ -0.49**	**ΔRMS≤ -0.27 and ΔMaxDev≤ -0.49**
Accuracy	0.70	0.66	0.68
Sensitivity	0.67	0.64	0.62
Specificity	0.74	0.68	0.74
Thoracic 3c/3cp Subgroup (N=34)			
	ΔRMS ≤ -0.58	**ΔMaxDev ≤ -1.32**	**ΔRMS≤ -0.58 and ΔMaxDev≤ -1.32**
Accuracy	0.56	0.56	0.59
Sensitivity	0.60	0.60	0.58
Specificity	0.53	0.53	0.59
Lumbar 4c/4cp Subgroup (N=46)			
	ΔRMS ≤ -0.26	**ΔMaxDev ≤ -0.61**	**ΔRMS≤ -0.26 and ΔMaxDev≤ -0.61**
Accuracy	0.83	0.76	0.80
Sensitivity	0.73	0.68	0.68
Specificity	0.92	0.83	0.92

5.4 Discussion

The study's primary objective was to determine the correlation of ST asymmetry parameters with patients' self-reported perception of their backs. The significant correlation obtained suggests that

patient perceptions of their backs reported as the GRC score are reflected and captured by the changes in ST parameters RMS and MaxDev. It was expected that there would be negative correlation as a positive GRC score (perceived improvement) is shown by negative ΔRMS and ΔMaxDev values (improved symmetry). In addition, the obtained correlation coefficients were greater than 0.30, which is the recommended minimal correlation value to estimate the MIC (Revicki et al., 2008). Few studies currently assess the relationships between surface topographic measures and back perception reported by patients. To our knowledge, no other studies have examined the correlation between the ST parameters from the markerless technique and the patient's perception of back status during treatment.

The secondary objective was to obtain the minimal asymmetry parameters required for patients to observe a positive change in their backs. An MIC of -0.27 mm for RMS was obtained from the ROC analysis, yielding a sensitivity of 67% and a specificity of 74%. An MIC of -0.49 mm for MaxDev was obtained from the ROC analysis, yielding a sensitivity of 64% and a specificity of 68%. The low MIC values of RMS and MaxDev indicate that participants who have little to no change in their ST parameters, perceive improvement in their back condition. Participants with a decrease in their ST parameters, beyond the MICs values, also perceive an improvement in their back condition. Similarly, if the change in ST get worse and increase over time, participants in this case will perceive a worsening of their back condition.

The accuracy in predicting patient perception between the RMS and MaxDev thresholds were compared. The sensitivity was higher using the RMS threshold compared to the MaxDev threshold. This suggests that compared to MaxDev, RMS can better detect a positive change in their condition according to the GRC. Likewise, the specificity was greater using the RMS MIC, indicating that it can more accurately identify patients' perceived lack of improvement (no change or deterioration). In general, when the criteria of both RMS and MaxDev MICs are satisfied to predict the perceived back condition, sensitivity decreases, while specificity increases but is equal to the specificity when only the RMS threshold is used. The MICs estimated for participants with thoracic curves reveal no difference in sensitivity and specificity between RMS and MaxDev.

There is a notable difference in the discriminative abilities of the surface topography parameters between participants with thoracic curves (3c or 3cp) and those with lumbar curves (4c or 4cp); participants with 4c or 4cp curves exhibited a stronger association between surface topography

and perceived improvement, as evidenced by the higher AUC. Participants with 3c or 3cp curves show a less clear relationship, with the surface topography parameters having a more moderate ability to predict perceived improvement. In contrast to the thoracic curve types, more severe exterior surface asymmetries may be found in the lumbar 4c and 4cp types. An AUC of at least 0.70 is recommended for an appropriate responsiveness study (de Vet et al., 2011). The AUCs of the complete sample and lumbar subgroups ROC curves were greater than 0.7. The AUC of the thoracic subgroup ROC curves was less than the recommended value (ΔRMS AUC 0.618, ΔMaxDev AUC 0.632). The observed difference between the thoracic and lumbar subgroups could be attributed to what patients perceive as meaningful change in these regions. It is possible that participants can notice changes for curves in the lumbar regions more easily than curves in the thoracic region. Mean ST parameters between participants classified as improved and not improved based on their GRC score revealed clear differences in all analyses (complete sample, thoracic subgroup, and lumbar subgroup). Despite the decreased responsiveness of the thoracic subgroup, participants with perceived improvement had a decreased RMS and MaxDev by 1.74 mm and 3.18 mm, respectively, while the not improved group had a slight decreased RMS by 0.36 mm and an increase in MaxDev by 0.17 mm.

The MICs in this study were estimated using the GRC as the anchor, a subjective measure reflecting the patient's point of view. The GRC score is not a continuous measurement and is recorded after six months of treatment. Participants make their impressions of their overall back condition over time and, therefore, might not accurately recall the state of their back condition at the beginning of the trial compared to the present time (Kamper et al., 2009). Participants may have overestimated or underestimated their GRC, which might be the reason for the low correlation between the GRC scores and the ST parameters' asymmetry.

Moreover, in response to the Schroth program intervention, participants who've perceived improvement in their back condition were greater in the Schroth group compared to the control group. The proportions of patients classified as torso asymmetry improvement based on the MICs were 59% in the Schroth group and 33% in the control group. The proportion of patients classified as not improved was 41% in the Schroth group and 68% in the control group. Differences in distribution among these categories between groups were significant (Chi-square=5.531, p=0.026).

Several limitations were identified in this study. First, the study faced constraints related to the sample size. The relatively small sample size restricted the ability to conduct subgroup analyses, particularly the differentiation between participants with 4c and 4cp curves and separating 3c and 3cp curves. Future research should prioritize larger sample sizes to allow for robust subgroup analyses. Another limitation is the need for more consensus regarding the methodology for estimating the MIC. Our study aimed to determine the MICs using ROC analyses, which revealed optimal thresholds for surface topography parameters associated with perceived improvement. Several methods are available for estimating the MIC; the anchor-based and distribution-based methods are the most common, though there has yet to be a consensus on the best method to determine the MIC (de Vet et al., 2011). The anchor-based method relies on external criteria, such as patient-reported outcomes, as a reference point for defining a meaningful change. On the other hand, distribution-based methods use statistical properties of the data, such as standard deviations, to establish a threshold for clinical significance. The open nature of the question, when participants are answering the GRC, may be another limitation. Participants may be evaluating shoulder asymmetries and waistline imbalance as factors affecting overall back condition. The ST parameters extracted from the thoracic or lumbar patches do not reflect the changes in the shoulders and waist asymmetries. The surface scans were cropped during preprocessing, and the ST patches on the shoulders and pelvis were excluded from the analysis. Finally, a validation study should be conducted to confirm the overall accuracy of the MIC estimates.

Despite these limitations, our study contributes valuable information regarding the relationship between ST parameters and perceived improvement in back condition, particularly in AIS treatment. The strong correlation between changes in root mean square (RMS) and maximum deviation (MaxDev) and patients' self-reported outcomes underscores the potential clinical relevance of surface topography measurements in assessing treatment efficacy.

5.5 Conclusion

The markerless ST technique has been proposed to assess torso asymmetries and posture outcomes. The extracted parameters from the ST technique correlated with self-reported GRC scores for evaluating back conditions. MICs were estimated using the anchor-based method. This study underscores the potential of surface topography parameters, particularly RMS and MaxDev, as objective measures capable of capturing changes in back conditions perceived by patients.

Chapter 6 : Effect of Scoliosis Treatment on Torso Asymmetry

This chapter was published in PLOS ONE journal:

Mohamed, N., Acharya, V., Schreiber, S., Parent, E. C., & Westover, L. (2024). Effect of adding Schroth physiotherapeutic scoliosis specific exercises to standard care in adolescents with idiopathic scoliosis on posture assessed using surface topography: A secondary analysis of a Randomized Controlled Trial (RCT). PLOS ONE, 19(4), e0302577. https://doi.org/10.1371/journal.pone.0302577

6.1 Abstract

Background**:** Adolescent idiopathic scoliosis (AIS) is a three-dimensional structural asymmetry of the spine and trunk affecting 2-4% of adolescents. Standard treatment is observation, bracing, and surgery for small, moderate, and large curves, respectively. Schroth exercises aim to correct posture and reduce curve progression.

Purpose: This study aimed to determine the effect of Schroth exercises added to the standard care compared to standard care alone on torso asymmetry in AIS.

Methods: In a randomized controlled trial (NCT01610908), 124 participants with AIS (age: 10-18, Cobb: 10°-45°, Risser: ≤3) were randomly assigned to the control (Standard care only) or Schroth (Standard care + Schroth treatment) group. Schroth treatment consisted of 1-hour weekly supervised sessions and 30-45 minutes of daily home exercises for six months. The control group received Schroth exercises in the last six months of the 1-year monitoring period. Markerless 3D surface topography assessed torso asymmetry measured by maximum deviation (MaxDev) and root mean square (RMS). Intention to treat linear mixed effects model analysis was compared to the per protocol analysis.

Results: In the intention to treat analysis, the Schroth group (n=63) had significantly larger decreased RMS (-1.2 mm, 95%CI [-1.5, -0.9]mm, p=0.012) and MaxDev (-1.9mm, 95%CI [-2.4, -1.5]mm, p=0.025) measurements compared to controls (n=57) after six months of intervention. In the per protocol analysis (Schroth n=39, control n=36), the Schroth group also had a significantly larger decrease compared to the control in both the RMS (-1.0mm, 95%CI [-1.9, -0.2]mm, p=0.013) and MaxDev measurements (-2.0mm, 95%CI [-3.3, -0.5]mm, p=0.037).

For the control group, both the intention to treat and per protocol analysis showed no difference in RMS and MaxDev in the last six months of Schroth intervention ($p>0.5$).

Conclusion: Schroth Exercise treatment added to standard care (observation or bracing) reduced asymmetry measurements in AIS. As expected, a greater effect was observed for participants who followed the prescribed exercise treatment per protocol.

6.2 Introduction

Adolescent idiopathic scoliosis (AIS) is paediatric condition, that leads to structural and morphological changes of the spine and trunk in all three planes of the body (Negrini et al., 2012). During adolescent growth, curves progress quickly, and frequent follow-ups are needed to monitor progression (Rogala et al., 1978). The Scoliosis Research Society (SRS) outlines treatment options which include observation (using frontal plane radiographs taken at regular intervals during growth) for patients with AIS having curves with Cobb angle <25°, bracing for curves 25°- 45°, and elective surgery for growing children with curves >45° (*Diagnosis And Treatment | Scoliosis Research Society*, n.d.). In addition to those modalities, the Society on Scoliosis Orthopedic Rehabilitation Treatment (SOSORT) guidelines recommends physiotherapeutic scoliosis specific exercises (PSSE) for smaller curves and as an add-on to bracing (Negrini et al., 2018).

The Schroth method is a PSSE approach aimed at recalibrating the postural alignment, capitalizing on the motor learning and control. It includes segmental realignment of the trunk, pelvis and legs, using specific eccentric and isometric muscle tension, and corrective breathing exercises to recalibrate the normal postural alignment (Fusco et al., 2011). Studies on AIS have shown positive results of Schroth exercises on Cobb angle, pain, self-image, vital capacity, and muscle endurance (Otman et al., 2005; Schreiber et al., 2015, 2016).

SOSORT Guidelines rank aesthetics as the most important goal when treating patients with AIS, which can be achieved through exercises (Aulisa et al., 2014; Negrini et al., 2006, 2018); however, very few studies on scoliosis address objective measurements of posture (Negrini et al., 2006). This is partly due to limited methods for measuring aesthetics and posture outcomes (Negrini et al., 2006). Most systems used clinically to measure aesthetics are based on photograph comparison between visits, with the implicit limitation that photographs are a 2D representation of a 3D

condition (Furlanetto et al., 2016). Aesthetics can be quantified using surface asymmetry tools such as trunk asymmetry scales or surface topography (ST) (Negrini et al., 2018).

In addition, the effect of Schroth exercises on external asymmetries of the trunk has not been adequately studied (Romano et al., 2013). One study reported waist asymmetry during a 24-week Schroth therapy program (Kuru et al., 2016). Cosmetic changes were also investigated to monitor the effects of other PSSE programs. The lateral deviation of the trunk after an intense rehabilitation program was evaluated in a pilot study by *Schumann et al* (Schumann et al., 2008). *Weiss et al* used surface topography to monitor the effect Activities of Daily Living (ADL) approach (Weiss, Hollaender, et al., 2006). The scarcity of studies reporting quantitative measurement of external trunk asymmetries during conservative treatment in addition to scoliosis-specific exercises highlights the need for studies to investigate the efficacy of PSSE on cosmetic outcomes. To address this critical gap, asymmetry measurement parameters from markerless surface topography analysis (ST) may be used to monitor posture outcomes during treatment. The proposed ST measurement parameters, which are developed using the whole torso surface information, have not previously been monitored during prescribed PSSE treatment programs.

Asymmetry analysis using ST can identify severity and monitor the effect of progression of scoliosis on torso posture (Ghaneei et al., 2018; Hong et al., 2017; Komeili et al., 2014, 2015a, 2015b). ST can quantify external torso asymmetry associated with scoliosis, and reduce the use of harmful radiographs for mild curves and those without curve progression (Ghaneei et al., 2018; Ronckers et al., 2008). A 3D markerless asymmetry analysis ST technique has been developed that assesses the 3D geometry of the torso by using the best sagittal plane of symmetry (Baclig et al., 2019; Hill et al., 2014; Komeili et al., 2014). This technique is reliable in identifying the location and direction of scoliosis curves by analyzing the severity of asymmetry (Komeili et al., 2015a). Extracted measurements from ST and decision trees help identify curve severity at a single visit and detect whether curves have progressed (>5° increase in Cobb angle) using scans from two temporally separated visits (Hong et al., 2017; Komeili et al., 2015b).

In this randomized controlled clinical trial (RCT), the objective was to determine the effect of the six months Schroth exercise intervention added to standard care compared to standard care used alone on asymmetry measures in AIS using ST analysis. Additionally, we determined the impact of offering Schroth exercises to participants from the control group during the last half of the 1-

year follow-up. We hypothesized that the Schroth exercise program added to standard care would be superior in reducing asymmetry compared to standard care alone.

6.3 Methods

6.3.1 Ethics Statement

This study has been approved by The University of Alberta Health Research Ethics Board (Pro00043397) and includes data collected from a pilot study (Pro00011552). The study reports a secondary analysis of data from a registered randomized controlled trial (NCT01610908). Its primary objective was to ascertain the impact of a six month Schroth PSSE intervention added to standard care versus standard care alone on the Cobb angle in participants with AIS (Schreiber et al., 2016). Participants provided signed assent and parents provided signed parental informed consent.

A total of 124 patients with AIS were recruited from the Edmonton Scoliosis Clinic. Participants were between 10 and 18 years old with curves between 10° and 45° (Schreiber et al., 2016). Participants were enrolled from April 2011 to December 2019. Recruitment of the first 50 participants during the pilot study included both sexes and Risser grades 0–5. The remaining 74 participants were restricted to only females with Risser grade ≤3. The selection criteria were modified following the pilot study to target patients with higher risk of progression. Recruitment of participants ceased when funding of the trial ended. The CONSORT flow chart is presented in Fig. 6.1.

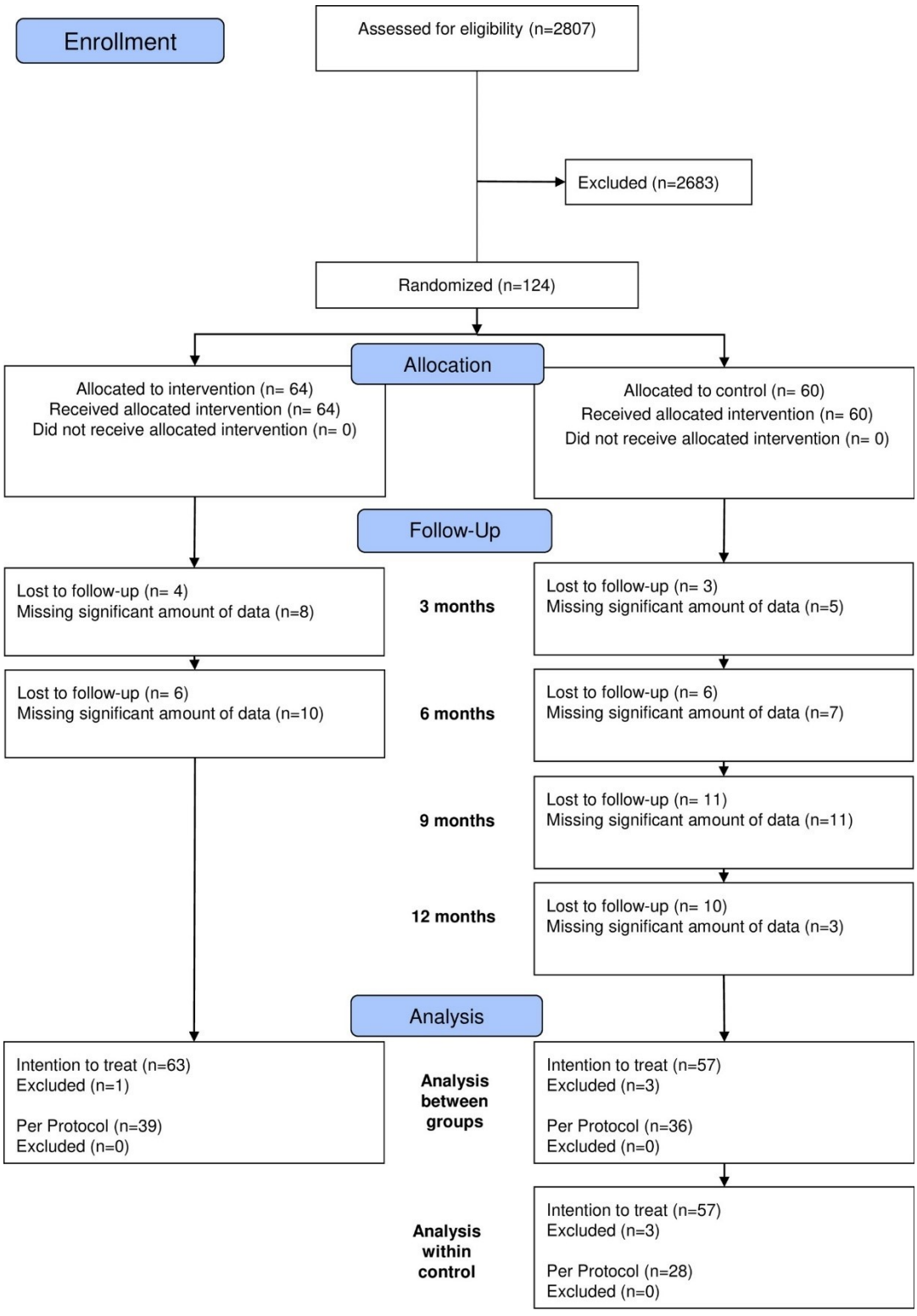


Fig. 6.1: CONSORT flow-chart showing Intention to treat and per protocol group sample sizes

In a parallel design, participants were randomly allocated to two groups: Control (n=60) or Schroth (n=64). The randomization sequence was stratified for four Schroth curve types and used random block sizes. Randomization was performed by the research coordinator using REDCAP to ensure the concealment of the randomization sequence (Schreiber et al., 2014).

The Schroth group was offered a six months supervised Schroth physiotherapeutic scoliosis specific exercises (PSSE) intervention combined with standard care. The treatment consisted of five one-hour private sessions during the first two weeks, followed by weekly one-hour private classes for six months and a 30-45 min daily home exercise program. After completing the six months supervised intervention, participants were advised to continue with the program, but the supervision was discontinued. An algorithm for determining scoliosis curve types and prescribing Schroth exercise based on curve type was used (Schreiber et al., 2012; Watkins et al., 2012). Exercise treatments were offered by four certified Schroth therapists over the trial duration aiming to pair individual participants with the same therapist.

Adherence was monitored using logbooks and session attendance. To maximize adherence, equipment and exercise description handouts were provided. Participants were asked to discuss their home exercise environment with their therapist. Parents were also asked to sign logbooks weekly, and adherence was discussed at every visit.

The control group received only standard care for the first six months, including observation or rigid bracing prescribed based on SRS recommendations (*SRS Bracing Manual | Scoliosis Research Society*, n.d.). During the last six months of the 1-year follow-up period, the control group underwent Schroth therapy.

6.3.2 Asymmetry Analysis

The ST analysis method was previously reported in the literature (Ghaneei et al., 2018; Hong et al., 2017; Komeili et al., 2014, 2015a, 2015b). Entire torso ST scans were collected at baseline, three, six, and 12 months by an examiner blinded to group allocation. Following the completion of data collection, all gathered information was anonymized to ensure the protection of participants' identities.

The four views, including both sides, front and back from the ST cameras were imported into Geomagic Control 2015 (3D Systems, North Carolina, USA) for merging and cropping to retain a

full 3D torso. For the asymmetry analysis, the torso model was duplicated and reflected about the midsagittal plane. The reflected torso was then aligned with the original (Fig. 6.2a) to minimize the distance between the models. The roto-inversion plane associated with this transformation was termed the plane of best symmetry (Baclig et al., 2019; Hill et al., 2014). A deviation color map (DCM) was created, illustrating the magnitude of the distances between each point on the original torso and its corresponding point on the reflected torso (Fig. 6.2b). The threshold to separate normal surface variations from relevant asymmetry was ±3mm (Komeili et al., 2014). However, in cases with maximum deviations greater than 9.33mm, a threshold value of ±9.33mm was also applied to ensure separation between asymmetry patches (Ghaneei et al., 2019). Color patches illustrating the asymmetry were isolated by removing points with deviations below the defined threshold values (Fig. 6.2c) using a custom algorithm (Wolfram Mathematica v. 12.1, Wolfram Research Inc., Illinois, USA) (Komeili et al., 2014).

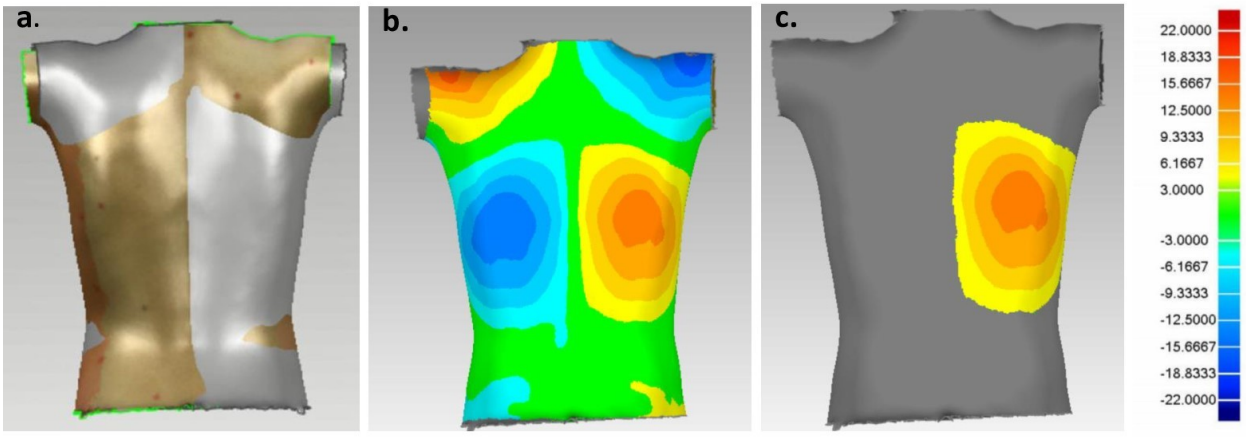


Fig. 6.2: Surface topography procedure(a) Alignment of the reflected (gold) and original torso (grey), (b) deviation color map where red reflects area of protrusion and blue to areas of depression relative to the other side, and (c) isolated patch of interest to calculate measurement parameters

For each isolated asymmetry patch, the maximum deviation (MaxDev) and root mean square (RMS) of the deviations were determined (Komeili et al., 2014). The RMS and MaxDev parameters for the largest patch by area were recorded for each patient at each time point. Patches on the shoulder were excluded since they do not show a corresponding curve on the radiograph and can manifest through uneven positioning of the shoulders (Komeili et al., 2015a). Patches

around the hip were also excluded as they might be affected by model cropping. ST analysis was completed by evaluators blinded to group allocation. The trunk asymmetry parameters of RMS and MaxDev are related to the cosmetic score stated in the protocol. The cosmetic score quantifies the visual appearance of trunk asymmetry, which is directly captured by the RMS value indicating the overall magnitude of asymmetry and the MaxDev value representing the most prominent deviation.

6.3.3 Longitudinal Asymmetry Evaluation

Participants were classified as experiencing progression, improvement, or no change during the first six months. A previously published classification tree was used to classify curves as either progression (increase in Cobb of ≥5°) or non-progression (Hong et al., 2017). To classify improvement, measurements were plotted within participants over time to evaluate trendlines. A patient with asymmetry worsening showed an increasing trendline and increasing intensity of their color patch (Fig. 6.3a). In contrast, the opposite is shown for a patient with asymmetry improvement (Fig. 6.3b). A constant trendline is observed for a patient without changes in asymmetry (Fig. 6.3c). Using a sub-sample of 50 participants (25 Schroth and 25 control) and by visual inspection, a threshold of -2.2mm for ΔRMS and ΔMaxDev was determined for asymmetry improvement classification. To avoid bias, the classification was applied blinded to group allocation.

6.3.4 Statistical Analysis

A covariate-adjusted linear mixed effects model was used to identify significant differences in RMS and MaxDev between Schroth and control groups. The initial analysis reported the significance of the group, time, and interaction effects. Final analyses were reported after dropping the effects that were not significant or dropping the main effects of time and group when interactions were significant. Separate linear mixed effects model analyses were carried out to compare the control group data from their last six months of the trial during which they had received Schroth exercises to the initial six months while they had not received exercises. Covariates included in the model were age, weight, height, scoliosis curve classification, and whether participants were wearing a brace. The covariance structure for each outcome were selected based on minimizing the Akaike information (AIC) value.

The proportions of participants classified as demonstrating progression, improvement, or no change in their asymmetry were compared between Schroth and Control groups using a Pearson Chi-square test. The significance level was set at $p \leq 0.05$.

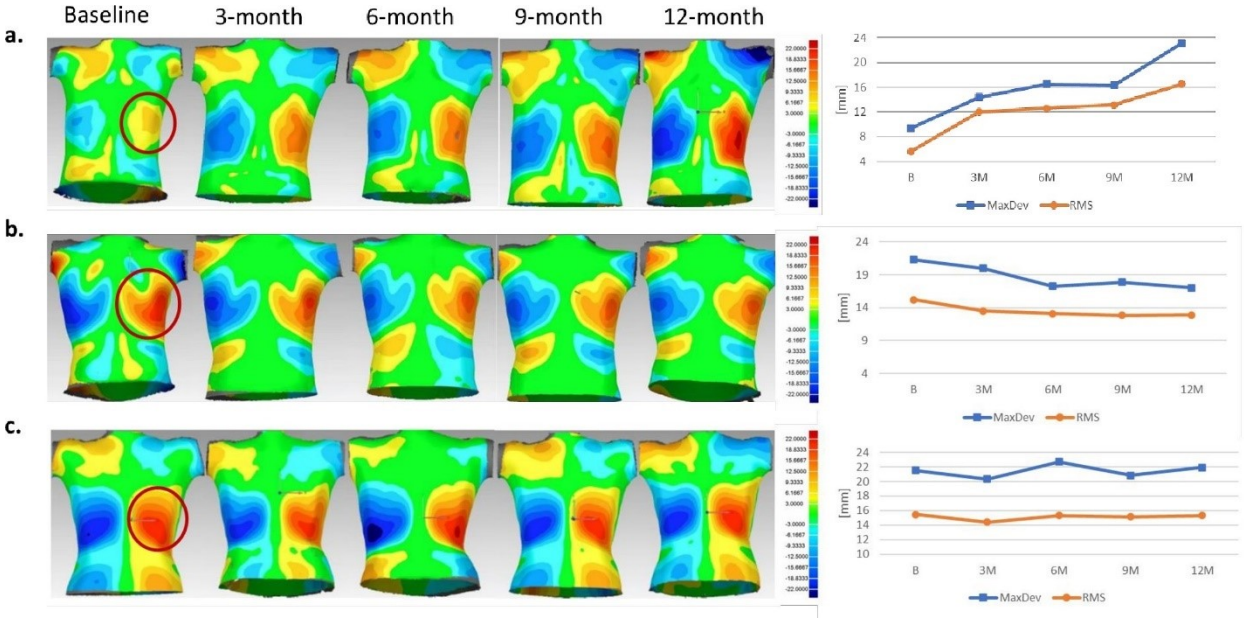


Fig. 6.3: Examples of deviation colour maps, RMS and MaxDev measurements of the largest patch by area over time. (a) Patient experiencing progression, (b) patient with asymmetry improvement, and (c) no change in asymmetry

Intention to treat and per protocol analyses were performed. The intention to treat analysis considered the entire sample, including dropouts and those with missing data. The per protocol analysis only included participants who completed the study per the prescribed protocol. To keep participants who dropped out or with missing data in the analysis, multiple imputation was used. The imputation model included all variables in the analysis model, and predictive mean matching (PMM) was used to impute the missing values (Fitzmaurice et al., 2011). Five complete datasets were obtained and analysed separately, and the resulting parameter estimates were then pooled across the imputed data sets using Rubin's rules (Rubin, 2004). Finally, Post hoc tests using Sidak adjustments to control for Type I error were conducted. In the intention to treat analysis, we opted for the implementation of multiple imputation as opposed to the last value carried forward technique stated in the protocol. The carry forward method may introduce bias by assuming that missing values remain constant over time (Fitzmaurice et al., 2011). The SAS software was used

for statistical analyses (SAS 9.4. SAS Institute Inc., Cary, NC). We report significant main effects to show the influence of individual independent variables, as well as significant interactions to show specific relationships among multiple variables and their collective impact on the outcome.

6.4 Results

Thirty-two of the 124 participants were lost to follow up or no longer wished to participate in the study. Seventeen of the 124 participants had scans that were of poor quality, or their 3D torso reconstruction was impossible because of missing data. Table 6.1 reports baseline characteristics of intention to treat and per protocol analyses between groups.

Table 6.1: Patient characteristics at baseline for the Schroth and Control groups for the intention to treat and per protocol analyses

	Intention to Treat		Per Protocol	
	Control	**Schroth**	**Control**	**Schroth**
	N=57	**N=63**	**N=36**	**N=39**
Age (yrs)	13.0 ± 1.7[a]	13.1 ± 1.8[a]	13.2 ± 1.81.8[a]	13.3 ± 1.8[a]
Height (cm)	156.4 ± 9.0[a]	155.6 ± 11.0[a]	156.9 ± 9.8[a]	156.0 ± 10.4[a]
Weight (kg)	46.6 ± 9.4[a]	44.1 ± 10.7[a]	46.9 ± 9.1[a]	43.3 ± 9.3[a]
Braced (%)	40 (70)[b]	41 (65)[b]	26 (72)[b]	24 (62)[b]
Number of girls (%)	56 (98)[b]	58 (92)[b]	36 (100)[b]	36 (92)[b]
Maximum Cobb Angle (°)	27.7 ± 9.1[a]	26.5 ± 8.2[a]	29.5 ± 8.3[a]	26.3 ± 8.5[a]
Risser	R0(26;46%)[b] R1(4;7%)[b] R2(7;12%)[b] R3(2;4%)[b] R4(4;7%)[b] R5(9;16%)[b]	R0(31;49%)[b] R1(2;3%)[b] R2(5;8%)[b] R3(8;13%)[b] R4(3;5%)[b] R5(5;8%)[b]	R0(16;44%)[b] R1(3;8%)[b] R2(5;14%)[b] R3(0;0%)[b] R4(2;6%)[b] R5(7;19%)[b]	R0(20;51%)[b] R1(1;3%)[b] R2(1;3%)[b] R3(7;18%)[b] R4(3;8%)[b] R5(2;5%)[b]
Curve Type	3c(5;9%)[b] 3cp(20;35%)[b] 4c(7;12%)[b] 4cp(25;44%)[b]	3c(9;14%)[b] 3cp(18;29%)[b] 4c(8;13%)[b] 4cp(28;44%)[b]	3c(3;8%)[b] 3cp(13;36%)[b] 4c(4;11%)[b] 4cp(16;44%)[b]	3c (5;13%)[b] 3cp (11;28%)[b] 4c (6;15%)[b] 4cp (17;44%)[b]
RMS (mm)	11.1 ± 4.1[a]	12.2 ± 4.6[a]	11.1 ± 4.0[a]	12.3 ± 5.2[a]
MaxDev (mm)	15.2 ± 6.4[a]	17.4 ± 8.0[a]	15.4 ± 6.2[a]	18.0 ± 8.9[a]

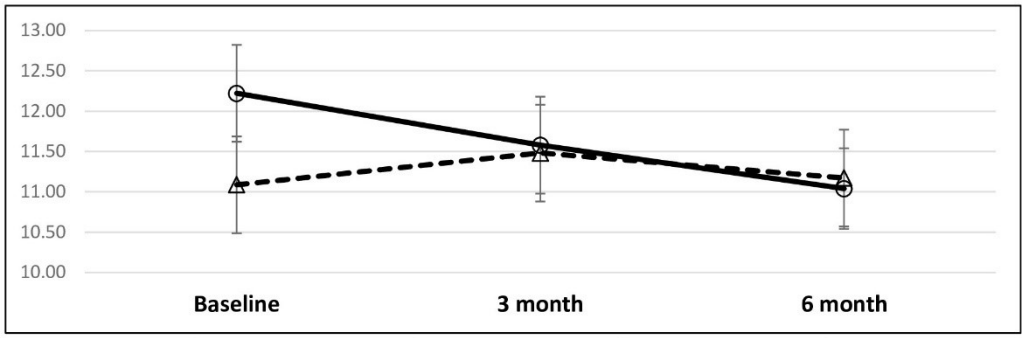
[a] Mean ± standard deviation [b] Frequency N (percent of whole sample)

6.4.1 Intention to Treat Comparisons of Changes in RMS and MaxDev

For the RMS outcome, group and time showed non-significant main effects (group effect p = 0.651, time effect p = 0.681), while their interaction was significant (time-by-group interaction p=0.03). The interaction was also significant in the reduced model without the main effects of group and time (time-by-group interaction p=0.019). (Fig. 6.4). Only age covariate was significant in the reduced model (p=0001). For the MaxDev outcome, group and time showed non-significant main effects (group effect p = 0.157, time effect p = 0.896), while their interaction was significant (time-by-group interaction p=0.040). The interaction was also significant in the reduced final model (time-by-group interaction p=0.003). (Fig. 6.4). Only age covariate was significant in the reduced model (p=0002). AIC was found to be the lowest using the Variance Components covariance structure in the linear mixed models for RMS and MaxDev outcomes. The final linear mixed effects model coefficient estimates are tabulated for RMS and MaxDev outcomes in Appendix B. From baseline to six months, post hoc comparison of the estimated marginal means revealed significant ST posture measurements of RMS and MaxDev improvement in the Schroth group from baseline to six months (p=0.001 and p=0.0002, respectively). No other significant improvement or worsening was observed in the control group. The complete post hoc analysis results are tabulated in Appendix B.

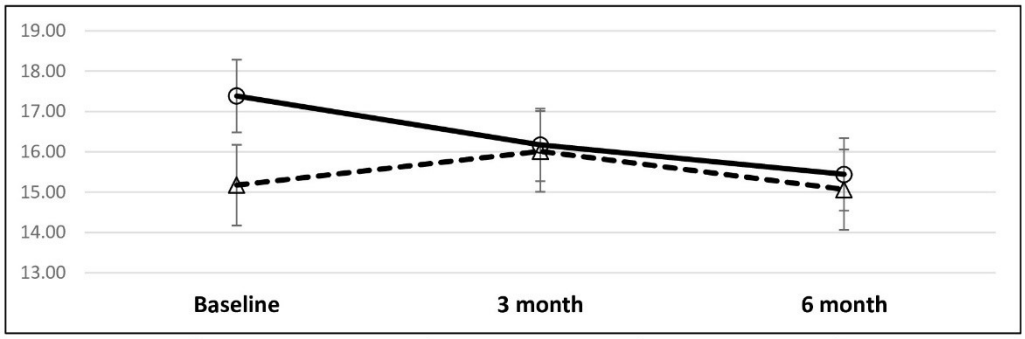
The Schroth group had a decrease in RMS by 4.9% (-0.6mm, 95%CI [-0.9,-0.4]mm) at three months, and 9.8% (-1.2 mm, 95%CI [-1.5,-0.9]mm) at six months compared to baseline. In the control group, RMS increased by 3.6% (0.4mm, 95%CI [0.08, 0.7]mm) at three months and increased by 0.9% (0.08mm, 95%CI [-0.3, 0.5]mm) at six months compared to baseline. Additionally, the Schroth group had a decrease in MaxDev by 6.9% (-1.2mm, 95%CI [-1.7,-0.7]mm) at three months, and 11.5% (-1.9mm, 95%CI [-2.4,-1.5]mm) at six months compared to baseline. In the control group, MaxDev increased by 5.3% (0.8mm, 95%CI [0.4, 1.3]mm) at three months and decreased by 0.7% (-0.1mm, 95%CI [-0.6, 0.4]mm) at six months compared to baseline (Fig. 6.4).

a.



Groups	Baseline	3 month	6 month	Group x Time
--△- Control	11.1 ± 0.6	11.5 ± 0.6	11.2 ± 0.6	**p=0.019**
⊖ Schroth	12.2 ± 0.6	11.6 ± 0.6	11.0 ± 0.5	

b.



Groups	Baseline	3 month	6 month	Group x Time
--△- Control	15.2 ± 1.0	16.0 ± 1.0	15.1 ± 1.0	**p=0.003**
⊖ Schroth	17.4 ± 0.9	16.2 ± 0.9	15.4 ± 0.9	

Fig. 6.4: Comparison between groups from intention to treat analysis (mean ± SE). Comparison of (a) RMS (mm) and (b) MaxDev (mm) outcomes. Linear mixed effects analysis produced the p-value presented

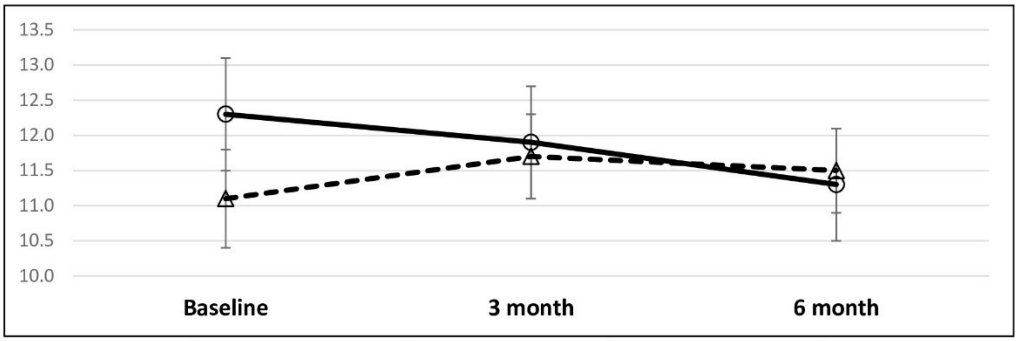
6.4.2 Per Protocol Comparisons of Changes in RMS and MaxDev

For the RMS outcome, group showed non-significant main effect (group effect p = 0.452), while time and time-by-group interaction were significant (time effect p = 0.028, time-by-group interaction p=0.041). The interaction was also significant in the reduced model without the main effects of group and time (time-by-group interaction p=0.013). Only age and curve classification covariates were significant (p=0001, p=0.022, respectively). (Fig. 6.5). For the MaxDev outcome,

group showed non-significant main effect (group effect p = 0.248), while time effect and time-by-group interaction were significant (time effect p = 0.002, time-by-group interaction p=0.050). The interaction was also significant in the reduced final model (time-by-group interaction p=0.001). (Fig. 6.5). Only age and curve classification covariates were significant (p=0002, p=0.040, respectively. The final linear mixed effects model coefficient estimates are tabulated for RMS and MaxDev outcomes in Appendix B. From Baseline to six months, post hoc comparison revealed significant ST posture measurements of RMS and MaxDev improvement in the Schroth group from baseline to six months (p=0.01 and p=0.001, respectively). No other significant improvement or worsening was observed in the control group. The complete post hoc analysis results are tabulated in Appendix B.

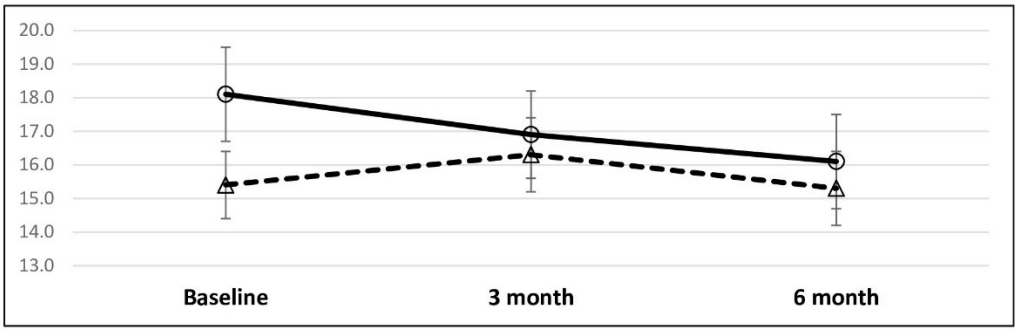
Over the six months follow-up, the Schroth group had a decrease in RMS by 3.2% (-0.4mm, 95%CI [-1.3,0.4]mm) at three months, and 8.1% (-1.0mm, 95%CI [-1.9, -0.2]mm) at six months compared to baseline. In the control group, RMS increased by 5.4% (0.6mm, 95%CI [-0.3, 1.5]mm) at three months and increased by 3.6% (0.4mm, 95%CI [-0.7, 1.5]mm) at six months compared to baseline. Additionally, the Schroth group had a decrease in MaxDev by 6.6% (-1.2 mm, 95%CI [-2.6,0.3]mm) at three months, and 11.0 % (-2.0mm, 95%CI [-3.3,-0.5]mm) at six months compared to baseline. In the control group, MaxDev increased by 5.8% (0.9 mm, 95%CI [-0.6,2.4]mm) at three months and decreased by 0.6% (-0.1mm, 95%CI[-1.7,1.5]mm) at six months compared to baseline (Fig. 6.5).

a.



Groups	Baseline	3 month	6 month	Group x Time
Control	11.1 ± 0.7	11.7 ± 0.6	11.5 ± 0.6	p=0.013
Schroth	12.3 ± 0.8	11.9 ± 0.8	11.3 ± 0.8	

b.



Groups	Baseline	3 month	6 month	Group x Time
Control	15.4 ± 1.0	16.3 ± 1.1	15.3 ± 1.1	p=0.001
Schroth	18.1 ± 1.4	16.9 ± 1.3	16.1 ± 1.4	

Fig. 6.5: Comparison between groups from the per protocol analysis (mean ± SE). Comparison of (a) RMS (mm) and (b) MaxDev (mm) outcomes. Linear mixed effects analysis produced the p-value presented

6.4.3 Intention to Treat Group Comparison Based on the Classification

Proportions of patients classified as presenting asymmetry improvement were 32% in the Schroth group and 18% in the control group after six months (Table 6.2). Participants whose asymmetry progressed were 35% and 19% in the control and Schroth group, respectively. Both groups had similar percentages with no change in asymmetry measurement (Schroth=49%; control=47%).

Differences in distribution among these categories between groups did not reach significance (Chi-square=2.396, *p*=0.092; Table 6.2). Combining improved and no change as a positive outcome, 81% of Schroth participants had a positive outcome compared to 65% of controls. This difference also did not reach statistical significance (Chi-square=2.278, *p*=0.137; Table 6.2).

Group	Asymmetry Improvement N (%)	No change N (%)	Progression N (%)	Pearson Chi-square 3 categories	Pearson Chi-square Improve or no change vs progression
Intention to Treat					
Control	10 (18)	27 (47)	20 (35)	p=0.092	p=0.137
Schroth	20 (32)	31 (49)	12 (19)		
Per Protocol					
Control	5 (14)	16 (44)	15 (42)	p=0.072	p=0.083
Schroth	12 (31)	19 (49)	8 (21)		

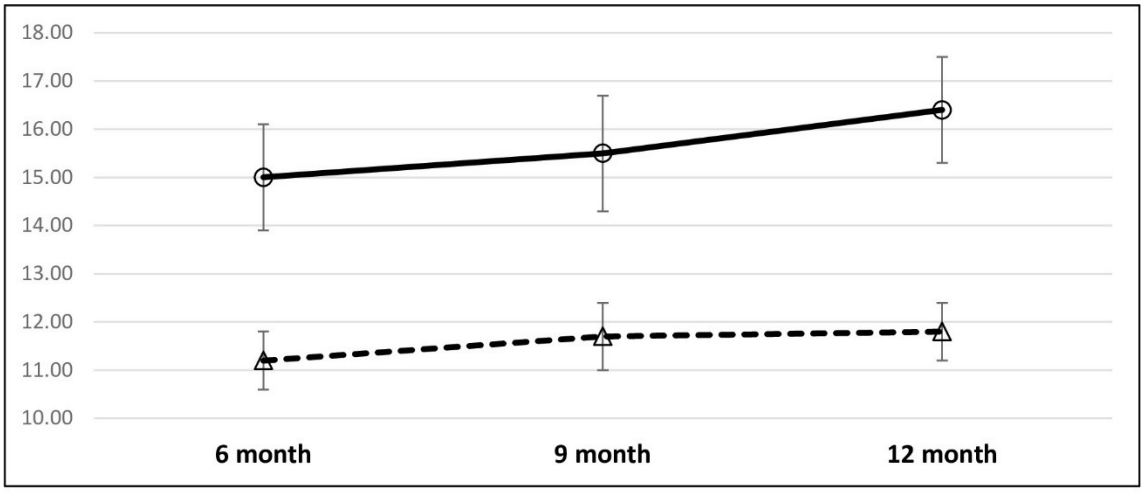
Table 6.2: Proportions of participants reported as N (percent of group sample) showing improvement, no change or progression in asymmetry after six months from the intention to treat and per protocol analyses

6.4.4 Per Protocol Group Comparison Based on the Classification

In the per protocol analysis, the percentage of participants classified as presenting with improved asymmetry was 31% in the Schroth group and 14% in the control group (Table 6.2). The control group had 42% of their participants whose asymmetry worsened. In contrast, 21% of the participants in the Schroth group exhibited asymmetry progression. Stable asymmetry measurements were 49% in the Schroth group and 44% in the control group. Differences between groups in distribution among the three categories did not reach statistical significance (Chi-square=5.158, *p*=0.072; Table 6.2). The percentage of participants classified as presenting positive outcomes was 79% in the Schroth group and 63% in the control group, but differences did not reach statistical significance (Chi-square=3.008, p=0.083; Table 6.2).

6.4.5 Intention to Treat Analysis of Control Group when Receiving Delayed Schroth Exercises

The time factor in the mixed model for the delayed Schroth exercises in the control group was not significant for the RMS measurements from nine months to 12 months (p=0.245). Likewise, difference in MaxDev over time was not significant from nine months to 12 months (p=0.255). Percent change of RMS was 4.5% (0.5mm, 95% CI [0.1,0.9]mm) from six to nine months, and 5.4% (0.6mm, 95% CI [0.2,1.0]mm) from six to 12 months (Fig. 6.6). Percent change of MaxDev was 3.3% (0.5mm, 95% CI [-0.04,1.0]mm) from six to nine months, and 9.3% (1.3mm, 95% CI [0.8,1.9]mm) from six to 12 months.

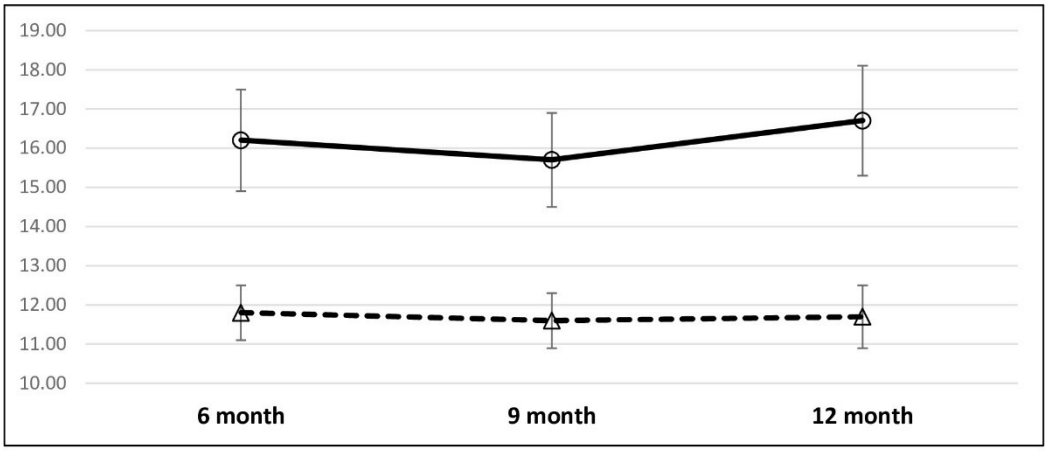


ST Outcomes	6 month	9 month	12 month	Time
--△- RMS (mm)	11.2 ± 0.6	11.7 ± 0.7	11.8 ± 0.6	p=0.245
—○— MaxDev (mm)	15.0 ± 1.1	15.5 ± 1.2	16.4 ± 1.1	p=0.255

Fig. 6.6: Average RMS (mm) and MaxDev (mm) in the group serving as control during the first six months and receiving exercises in the last six months presented over time from the intention to treat analysis (mean ± SE). Linear mixed effects analysis produced the p-value presented

6.4.6 Per Protocol Analysis of Control Group when Receiving Delayed Schroth Exercises

The time factor in the linear mixed effects model for the delayed Schroth exercises in the control group was not significant for the RMS measurements from nine months to 12 months (p=0.395). Likewise, difference in MaxDev over time was not significant from nine months to 12 months (p=0.373,). Percent change of RMS was -1.7% (-0.2mm, 95% CI [-1.3,1.0]mm) from six to nine months, and -0.8% (-0.1mm, 95% CI [-1.3,1.2]mm) from six to 12 months. Percent change of MaxDev was -3.1% (-0.5mm, 95% CI [-2.1,1.0]mm) from six to nine months, and 3.1% (0.5mm, 95% CI [-1.1,2.1]mm) from six to 12 months (Fig. 6.7).



ST Outcomes	Baseline	3 month	6 month	Time
--△- RMS (mm)	11.8 ± 0.7	11.6 ± 0.7	11.7 ± 0.8	p=0.395
—○— MaxDev (mm)	16.2 ± 1.3	15.7 ± 1.2	16.7 ± 1.4	p=0.373

Fig. 6.7: Average RMS (mm) and MaxDev (mm) in the group serving as control during the first six months and receiving exercises in the last six months presented over time from the per protocol analysis (mean ± SE). Linear mixed effects analysis produced the p-value presented

6.5 Discussion

For both RMS and MaxDev measurements, a significant interaction effect between group and time was determined in the intention to treat and per protocol analyses, suggesting that the longitudinal changes depend on the group (Schroth or control). During the six months intervention period, RMS and MaxDev significantly decreased in the Schroth group. In addition, MaxDev had a greater

reduction compared to RMS measurement. Both the intention to treat and per protocol analyses supported the same finding. The linear mixed model analyses revealed age and curve classification had significant effect on the outcomes. Age had a significant main effect for both MaxDev and RMS, such that older participants had higher RMS and MaxDev. Patients classified as 4C curve types had better RMS and MaxDev compared to those classified as 3C. Similarly, patients classified as 4CP, had better RMS and MaxDev compared to those classified as 3CP. These results could be due to more severe external surface deviations in 4C and 4Cp types that present with both lumbar and thoracic curves, compared to the 3C types presenting with a single thoracic curve. Although results show that the Schroth exercise program added to the standard care has potential clinical use, a clinically relevant threshold of improvement in asymmetry parameters is not yet known. Further analysis can include comparing perceived changes in appearance by participants and examining changes within specific curve types.

Applying different thresholds for calculating the root mean square (RMS) introduces a limitation that can impact the comparability of results. In order to ensure patch separation and prevent patch overlapping during ST analysis, a standard deviation threshold of 9.33 mm was applied to the majority of the data. However, for ST scans of certain patients where the maximum deviations were below 9.33 mm, it was not feasible to apply the same threshold. In these cases, a reduced normal deviation threshold of 3 mm was utilized instead (Ghaneei et al., 2018; Komeili et al., 2014, 2015a).

Classification results suggest that Schroth exercise treatment has the potential to stabilize or improve scoliosis-related torso asymmetry compared to standard care alone. However, the differences noted were not statistically significant. Results may suggest that a more extended treatment period or a higher exercise dose would have a more substantial effect. A review of early Schroth intervention studies outlines that high amounts of supervised exercises include programs where participants were exercising six days a week for 6-8 hours at a time (Negrini et al., 2008).

Several factors could have resulted in the lack of statistical significance. The improvement threshold of -2.2mm for ΔRMS and ΔMaxDev is unique to this study. Therefore, further work is needed to determine if this threshold is clinical meaningful. Additionally, the progression decision tree was developed to maximize specificity and ensure no cases with progression were missed for use when determining which patient may require a new treatment rather than to detect effects of

treatments. Moreover, our sample included patients with less severe curves, which could present a challenge in observing a change over time. Therefore, we recognize that our ST model might not have been sensitive enough to detect meaningful changes in this group of patients with mild to moderate curve magnitudes. Participants classified as presenting progression from ST only have a 30% chance of actual progression (Hong et al., 2017), and it is possible that participants were incorrectly classified in the progression group.

The effect of offering Schroth exercise during the last six months of the 1-year follow-up in the control group showed no significant difference for both intention to treat and the per protocol analysis. In the control group, attendance of the prescribed supervised sessions was 73%, and the completed home exercises was 68%. The Schroth group attended 76% of the supervised sessions and completed 72% of the prescribed home sessions. The discrepancy in adherence to the treatment between the Schroth group and the control group may have played a role in the observed outcomes. Additionally, participants in the control group with decreased RMS and MaxDev measures (ΔRMS ≤ 0 and ΔMaxDev ≤ 0, n=12) between the six-months and 1-year follow-ups attended 92% of the supervised sessions and completed 80% of the prescribed home exercises. Participants in the control group with increased RMS and MaxDev measures (ΔRMS > 0 and ΔMaxDev > 0, n=16) between the six-months and 1-year follow-ups attended 72% of the supervised sessions and completed 71% of the prescribed home exercises. This suggest that adherence to treatment is essential to yield significant results.

Studies that evaluated asymmetry were compared with our RCT. *Kuru et al.* evaluated waist asymmetry of participants randomized into control, home exercises, and supervised exercises groups (Kuru et al., 2016). Results showed that the supervised group had the most significant decrease between baseline and 24th-week measurement. *Schumann et al.'s* pilot study measured the average lateral deviation and average surface rotation of participants in different postural positions after 3-4 weeks of intensive rehabilitation (Schumann et al., 2008). The study showed an improvement of both parameters in the conscious and corrected posture. *Weiss et al.* focused on the effect of the Activities of Daily Living (ADL) rehabilitation approach using ST (Weiss, Hollaender, et al., 2006). Lateral deviation had a greater decrease (non-significant) in the ADL group compared to the control group. Although asymmetry analysis procedures reported in these

studies differed from this RCT, the findings support that exercise treatment can have a positive effect on aesthetic measures.

This RCT studied the effect of Schroth exercises added to the standard care, but the effect of Schroth alone cannot be determined. This study could not assess the exclusive impact of Schroth exercises due to their combination with standard care including observation and bracing. To achieve that, patients would need to be randomized into exercise-only and brace-only groups, which is ethically challenging. Our focus was on evaluating the supplementary effect of adding Schroth PSSE to standard care, rather than using it as a standalone treatment. In line with North American standards, involving observation and bracing for patients with curves ≤45°, our sample included an experimental group receiving Schroth in combination with observation or bracing and a control group receiving observation or bracing exclusively. We ensured balanced distribution of observation and bracing in both groups and controlled for brace-related factors in our analyses. The proportions of participants prescribed bracing between the Schroth and control group was relatively balanced, with 65% of participants having received Schroth and wearing a brace, compared to 70% of participants wearing brace only. Groups did not differ significantly for baseline Cobb angle. The average baseline Cobb angle for participants with mild curves (< 25°) in the Schroth group was 18.8° ± 3.8° and 18.1° ± 3.2° in the control group. Likewise, the baseline Cobb angle for participants with moderate curves (≥ 25°) in the Schroth group was 33.3° ± 5.6° and 32.8° ± 5.5° in the control group. Thus, there was a balanced distribution of mild and moderate Cobb angle degrees between the groups.

This RCT did not monitor participants' brace wear compliance, since brace monitors were not available to us at that time making it difficult to assess differences in bracing compliance between the Schroth and control groups. Ultimately, this RCT aimed to determine the effect of Schroth PSSE added to standard care, not as a standalone treatment. Another limitation was the treatment period of the Schroth exercises, where a minimum short-term treatment of 12 months is recommended to produce reliable results (Negrini et al., 2015). However, adherence would likely be affected during a longer treatment, and our results show that better outcomes are achieved when participants follow the prescribed protocol. Additionally, this study did not look at whether the change in asymmetry during the six months of the Schroth exercise can be maintained after the supervised treatment is ceased.

Another limitation is the absence of a minimum clinically important difference (MCID) for the ST parameters in this RCT. While, establishing a MCID for the ST parameters was not a focus of this study, future research should establish a MCID to improve the interpretation of clinical relevance of the findings in ST research.

Nevertheless, there are important strengths in this study to highlight. We conducted an RCT, reported both intention to treat and per protocol analyses, and reported an analysis with control group participants serving as their own control after receiving six months of exercises. We also aimed to reduce bias in the results by conducting a blinded analysis of the ST data and classification. We also showed that objective markerless ST asymmetry measurements are sensitive to change over a 6-month Schroth intervention.

6.6 Conclusion

Schroth exercise added to standard care for participants with AIS improved objective surface topography measures of back asymmetry. Both intention to treat and per protocol analyses yielded similar results, but those who completed the trial had a more significant decrease in quantitative asymmetry measurements.

Chapter 7 : Association of Toso Asymmetry with 3D Vertebral Position

7.1 Introduction

AIS is characterized as a 3D curvature of the spine and trunk, and the Cobb angle obtained from the posterior-anterior (PA) radiographs is a simplified indicator of the curve severity in 2D (Ramirez et al., 2006; Weinstein et al., 2008). The complex 3D deformity of the spine manifests into visible asymmetries of the trunk, leading to self-image concerns (Choudhry et al., 2016). The gold standard method for diagnosing and monitoring scoliosis using the Cobb has an increased risk of cancer from ionizing radiation (Nash et al., 1979; Weinstein et al., 2008).

In Chapter 3, a CNN model was developed to detect scoliosis using the surface topography (ST) method. The model's output is a binary classification, which outputs 1 for positive AIS prediction and 0 for negative AIS prediction. The classification model showed excellent results with a specificity of 90% and a sensitivity of 97%. Since the model was developed only to predict positive or negative AIS based on ST, the underlying curvature of the spine was not predicted. Using ST to predict the shape of the spinal curvature from trunk asymmetry analysis has been previously proposed in the literature (Kokabu et al., 2021; Watanabe et al., 2019). From the predicted curve, the Cobb angle may be determined. A CNN model can be developed for regression tasks, which outputs a continuous variable representing the profile of the spine columns. The regression model can be used to monitor the spinal curve of individuals with confirmed AIS and as a screening tool since the Cobb angle estimated from the spinal profile greater than 10° indicates possible scoliosis (Weinstein et al., 2008). For ethical reasons, the radiographs were not obtained for participants with typical spine development. Participants with normal spine development were included in the study using an inclusion criterion of ATR from the scoliometer of less than 5°. Thus, a regression model trained on ST data to predict vertebrae positions of the spine was limited since radiographs were not available for the negative AIS group.

The nature of the 3D characteristics of AIS highlights the need for scoliosis assessments using the 3D shape of the vertebral column. Capturing the lateral (LAT) x-ray view of the spine and the standard PA view increases the quantity of information and offers the opportunity to extract 3D features of the spine. In the lateral view, thoracic kyphosis and lumbar lordosis can be determined,

which defines the global balance of the spine (Negrini et al., 2018). 3D reconstructions of the spine and pelvis can be created from PA and LAT projections to better understand the scoliosis deformity. The 3D representation of the spine may be helpful in clinical diagnosis, quantitative evaluation, treatment planning, and progression monitoring, which are usually assessed using conventional 2D measurements (Negrini et al., 2018). However, the lateral projection is only captured during routine follow-up visits unless necessary to reduce radiation exposure (Negrini et al., 2018). Predicting the spine profile using the markerless ST method is an alternative to capturing multiple views during the radiographic assessment to reconstruct the 3D alignment of the spine. However, it is important to evaluate the association of torso asymmetry parameters from the ST method with the 3D positions of the vertebrae to determine the potential use of the technique to predict the underlying profile of the spine. The objective is to determine if the underlying spinal curve and the ST characteristics are correlated in order to predict the 3D form of the spine.

7.2 Methods

As part of the pilot study, full torso surface scans and biplanar radiographs were obtained from seven participants with AIS. Ethical approval, granted by the University of Alberta Health Research Ethics Board, was obtained with approval number Pro00117065. The low-dose EOS Imaging systems was used to obtain the PA and LAT radiographs (EOS Imaging, Paris, France). Surface scans were obtained from 4 VIVID Minolta cameras (Konica Minolta Sensing Inc., Ramsey, NJ, USA) capturing each side of the torso and merging the scans to obtain a complete 3D surface torso model. The 3D model of the spine was reconstructed from the bi-planar radiographs. The reconstructed model was obtained from a CNN model trained on a dataset of scoliosis patients, with inputs being the digitized x-ray images identifying the vertebras (12 thoracic, five lumbar) in PA and LAT views. The CNN model was developed as part of a parallel study with the objective of developing an automated fast reconstruction of the spine model from biplanar radiographs (W. Chen et al., 2024).

The registration process was conducted to align the 3D reconstructed spine to the surface scan. Anatomical landmarks were identified in the biplanar X-rays and their relative distance to the 3D vertebrae. The T9 vertebra centroid was the reference. The anatomical landmarks of the midpoint of posterior superior iliac spine (PSIS), and the left and right sides of the anterior superior iliac spine (ASIS) were identified on the biplanar x-rays. Surface markers were also placed on

participants' torso prior to capturing the scan. Rigid registration was implemented to align the 3D spine representation to the surface scan using the position of the landmarks identified in both models. Specifically, the iterative closest point algorithm was used to align and minimize the distance between the landmarks of the spine model to the corresponding point on the surface scan.

The ST technique, described in Chapter 3 and Chapter 5, was then applied to the combined torso-spine model. Briefly, the surface torso was duplicated and mirrored around the sagittal plane. The original and reflected torso were aligned in such a way as to minimize the distance between the torso models using the iterative closest point algorithm. The deviations between the aligned torsos are then obtained. A threshold representing typical deviations of ± 3 mm was applied to the raw deviations. With the threshold applied, areas of torso asymmetries related to the underlying curvatures are revealed in a deviation colour map (DCM). Indices describing the deformity from the ST technique were extracted, which included the root mean square (RMS), maximum deviation (MaxDev), and the centroid location in 3D of the asymmetry patch (Fig. 7.1).

The centroids of the 12 thoracic and five lumbar vertebrae of the spine model were also extracted after the spine-surface scan registration procedure, and the apex of the spinal curves were identified. The corresponding ST patch was used for each curve identified in the spinal column to obtain the RMS, MaxDev, and patch location parameters. In addition, the ST patch on the posterior and anterior surface of the trunk nearest to the curve location were evaluated.

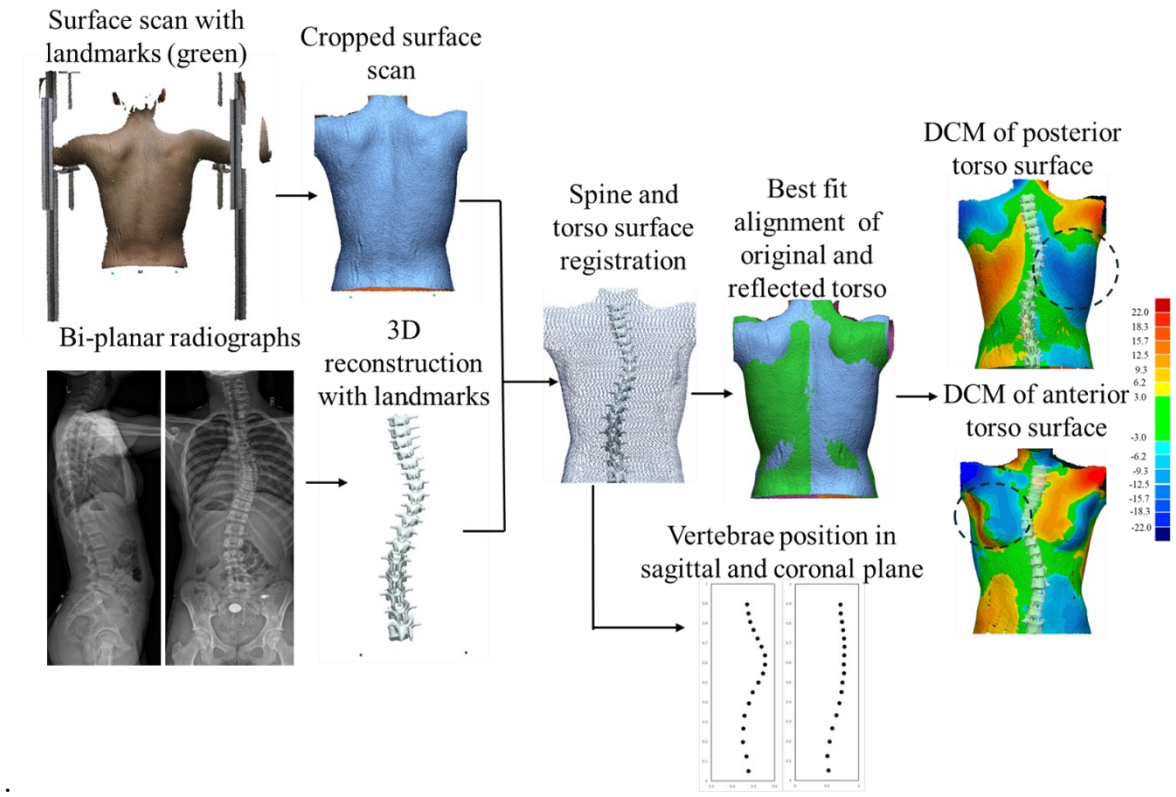


Fig. 7.1: Data processing and acquisition of ST and 3D spine features for analysis

Pearson correlations between the ST parameters indices and the vertebral positions of the spine were evaluated. Specifically, the ST parameters of RMS, MaxDev, and patch location in 3D space were compared to the centroid of the apex vertebrae.

7.3 Results

RMS and MaxDev obtained from the patch information of the posterior torso surface were plotted against the Cobb angle and the position of the curve apex to determine the association between the underlying curvature and the ST parameters from external torso asymmetries. RMS and MaxDev were significantly correlated with the Cobb angle ($p=0.008$ and $p=0.008$, respectively). However, RMS and MaxDev were not significantly correlated with the position of the apex of the spinal curve in all three planes (Fig. 7.2 and Fig. 7.3, respectively).

RMS and MaxDev from patch information of the anterior torso surface were compared to the Cobb angle and the apex vertebrae 3D position. The Pearson correlation between RMS and MaxDev and the Cobb angle was insignificant ($p=0.076$ and $p=0.195$, respectively). Additionally, the correlation

between RMS and MaxDev and the apex position in all three planes was not significant (Fig. 7.4 and Fig. 7.5).

The patch centroids were compared to the apical vertebra position, as shown in Fig. 7.6. The apical height was significantly correlated with the height of the patch extracted from the back of the torso surface (p=0.006). The apical position in the posterior-anterior direction was correlated with the patch centroid obtained from the posterior surface of the torso (p=0.021). Additionally, the height of the patch extracted from the anterior surface of the torso was significantly correlated with the height of the apical vertebra (p=0.018) (Fig. 7.7).

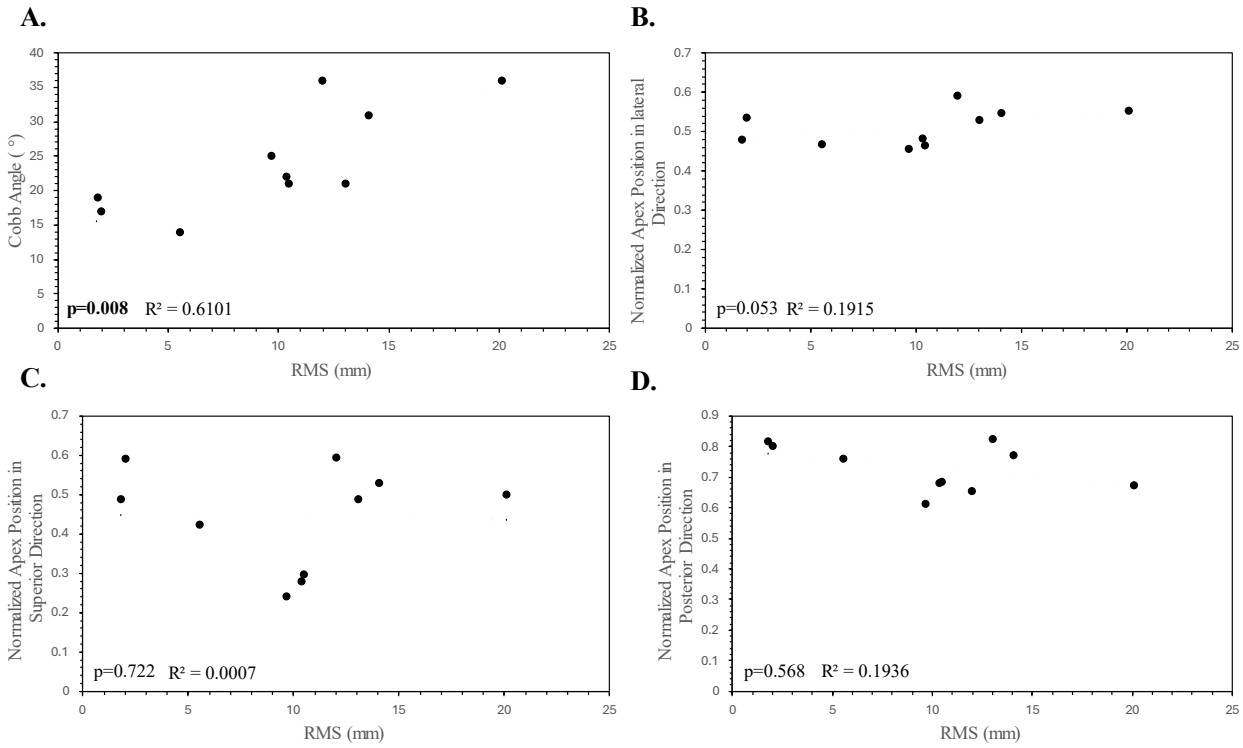


Fig. 7.2: Correlation of RMS of posterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D)

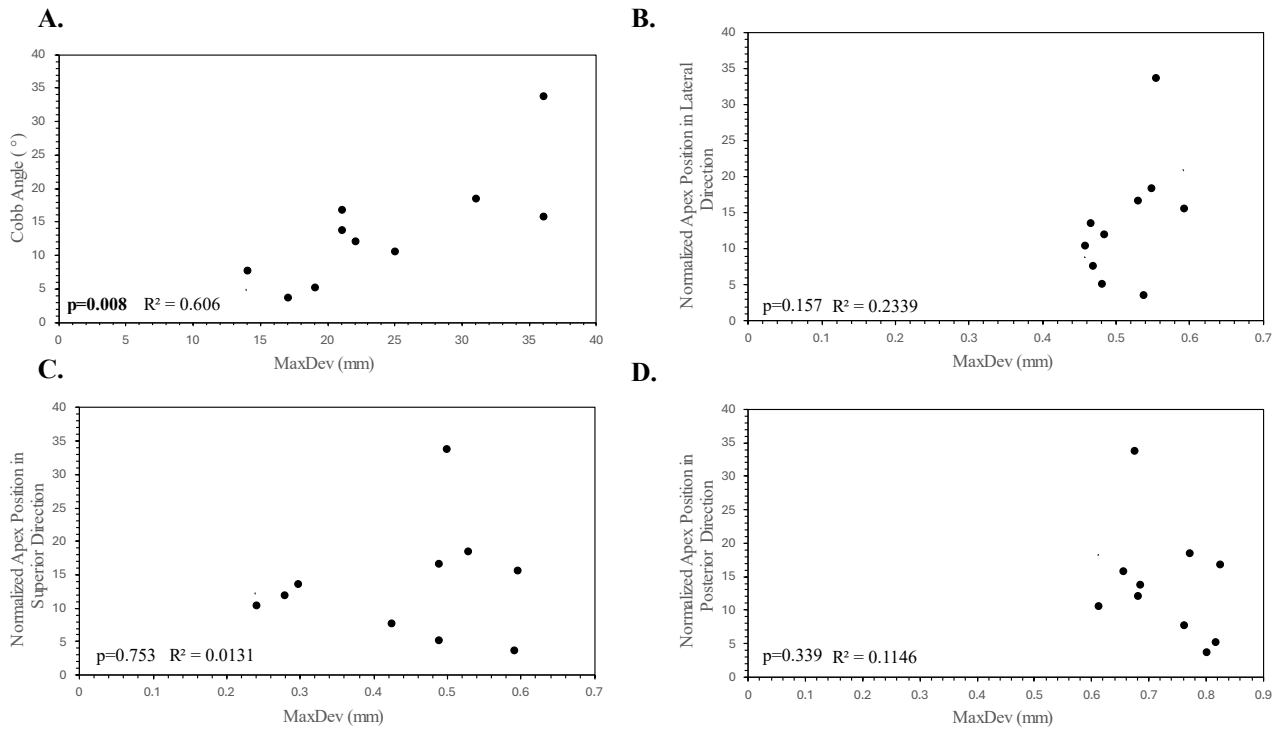


Fig. 7.3: Correlation of MaxDev of posterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D)

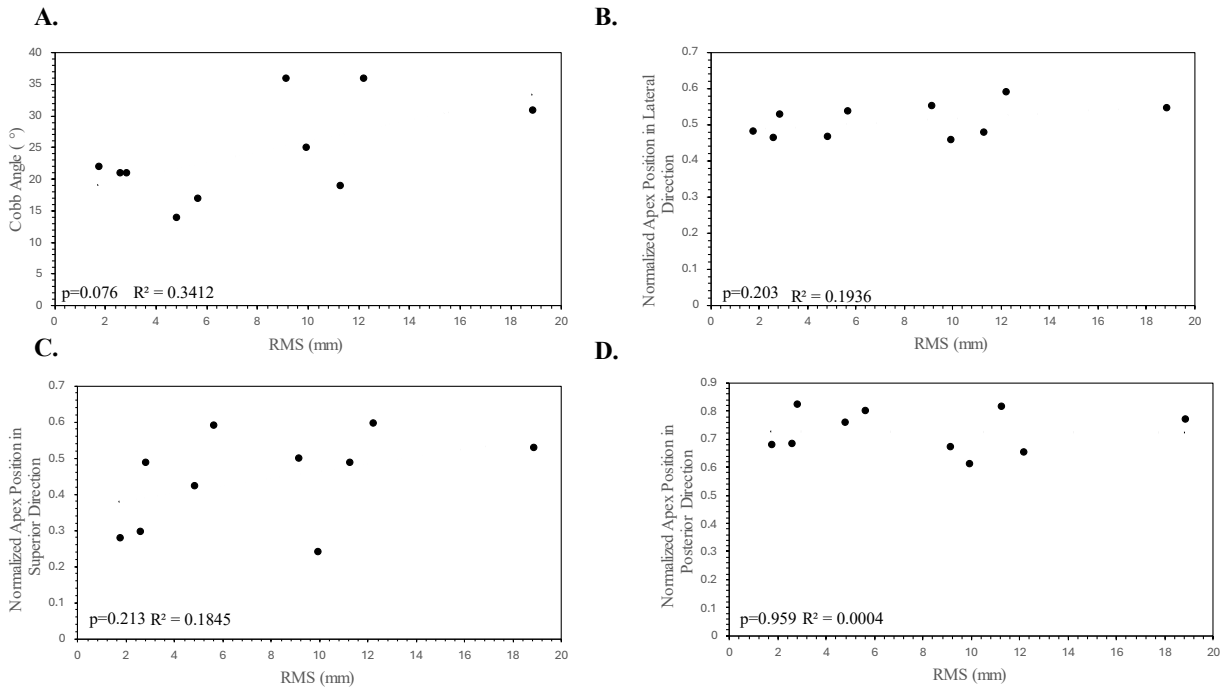


Fig. 7.4: Correlation of RMS of anterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D)

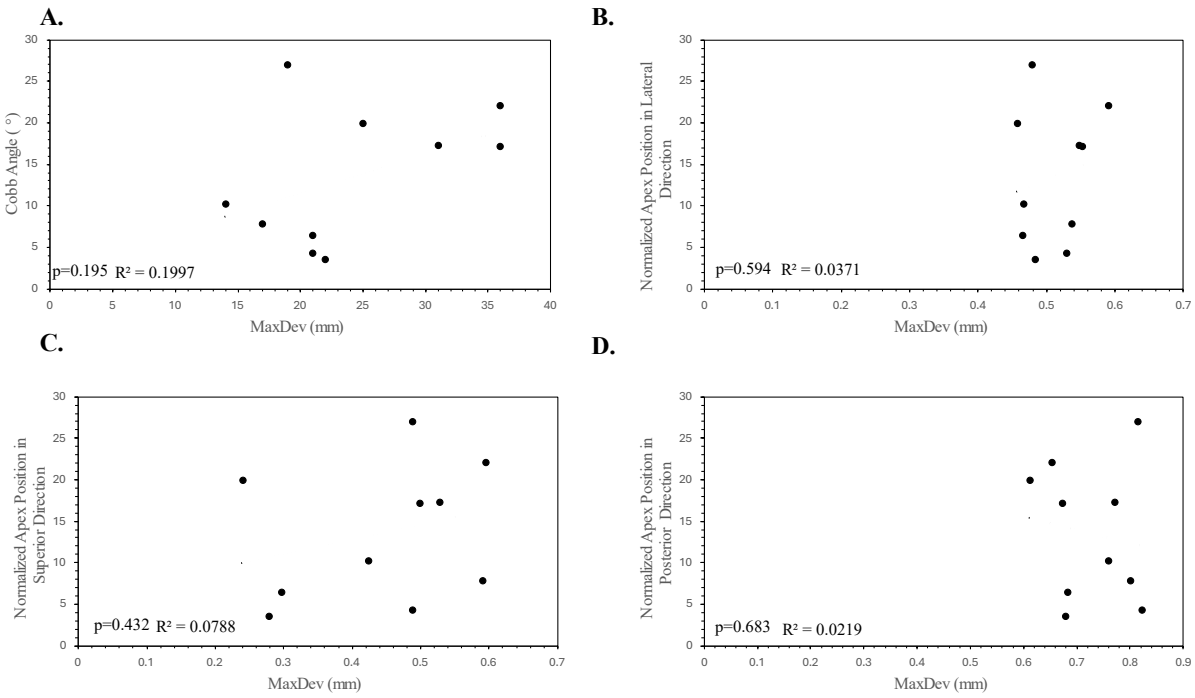


Fig. 7.5: Correlation of MaxDev of anterior torso surface patch with the Cobb angle (A) and the apical vertebra position in lateral direction (B), superior direction (C), and posterior direction (D)

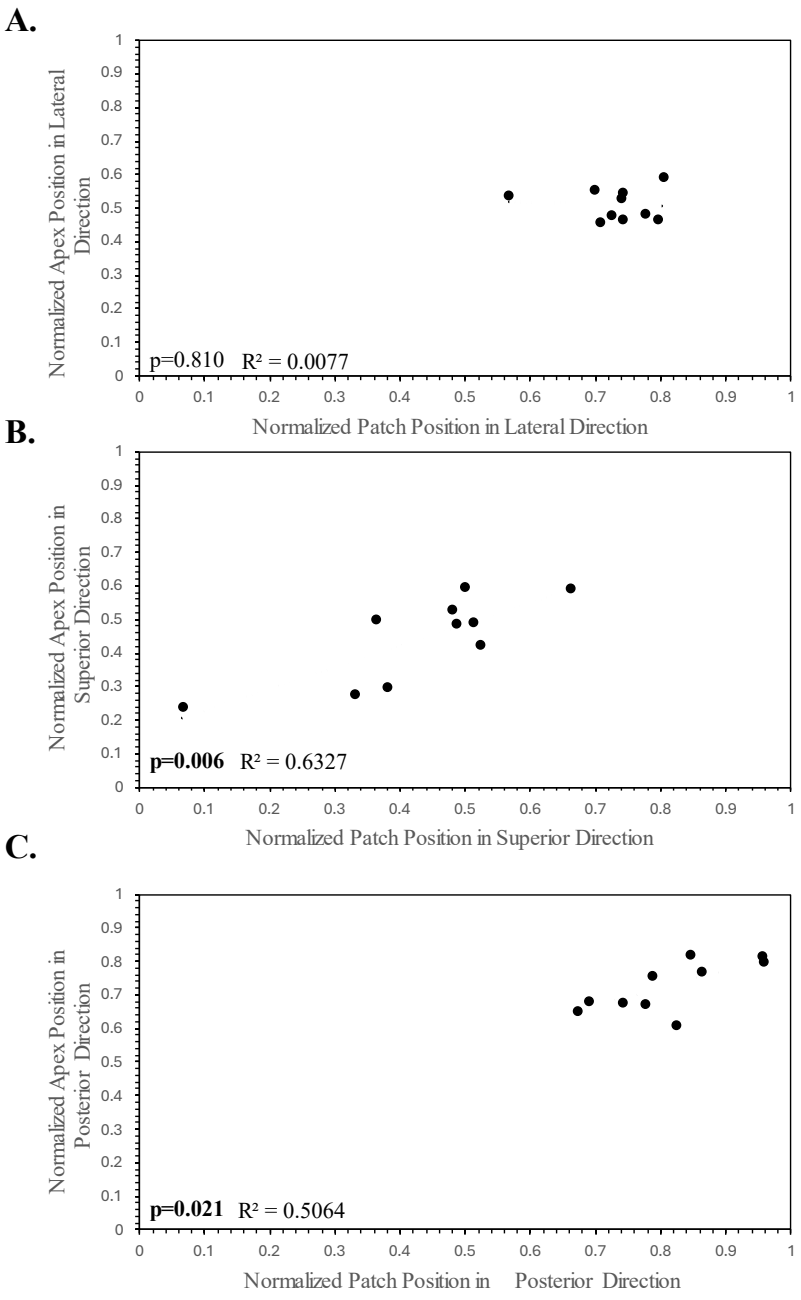


Fig. 7.6: Correlation of posterior torso surface patch centroid with the apical vertebra position in the lateral direction (A), superior direction (B), and posterior direction (C)

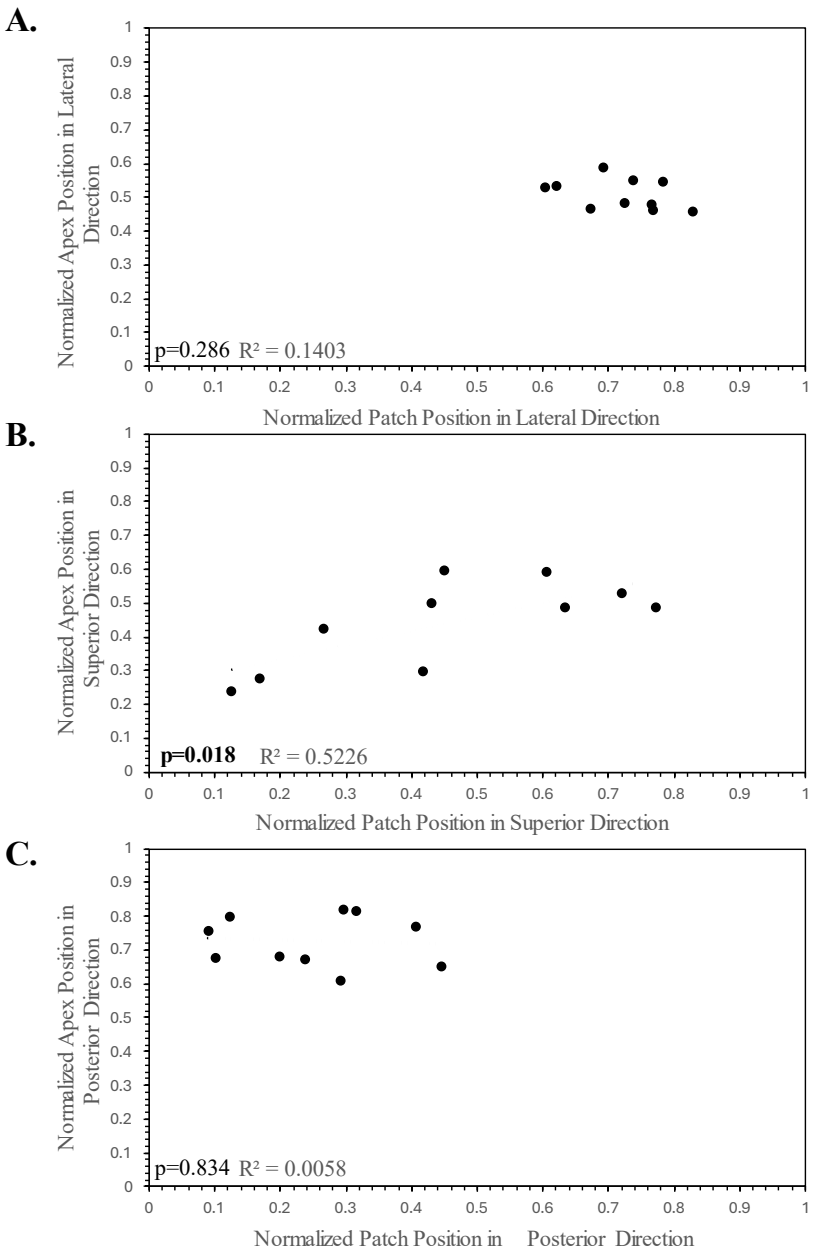


Fig. 7.7:Correlation of anterior torso surface patch centroid with the apical vertebra position in the lateral direction (A), superior direction (B), and posterior direction (C)

7.4 Discussion

The objective of this chapter was to determine the association of torso asymmetry using the ST technique with the 3D vertebral position. Specifically, the apex of the curvature was compared to the corresponding RMS, MaxDev and patch centroid locations. Additionally, the ST parameters were extracted from both the posterior and anterior of the torso surface to compare with the 3D apical curve position. Significant correlations were observed between the Cobb angle and RMS and MaxDev parameters. However, the RMS and MaxDev extracted from the posterior and anterior torso surfaces were not significantly correlated with the apical position of the vertebra in 3D. The position of the patch centroids from both the anterior and posterior torso surfaces was correlated with the position of the apical vertebra. Overall, the extracted parameters from torso asymmetry have shown potential in predicting the apical curve position in 3D space.

Some limitations have been identified in this analysis. The most notable limitation is the small data size, partly due to the limited number of participants with PA and LAT radiographs for reconstruction. The radiographs and surface topography were not taken simultaneously or on the same day. In our analysis, we assume that the curvature of the spine and the geometry of the torso remained consistent between the dates of the x-ray and the surface scan, which were taken within 30 days of each other. This assumption is based on the premise that significant changes in spinal curvature and torso shape are unlikely to occur within such a short period under normal conditions. This work serves as a preliminary pilot study and future work can include increasing the sample size and collecting biplanar radiographs and surface scans on the same day.

The analysis conducted assumed a linear relationship between the parameters extracted through the ST asymmetry technique and the 3D position of the curve apex. Further work can include applying pattern recognition algorithms such as CNNs to predict the position of the curve's apex. These algorithms can relate the input to the output data even when the relation is non-linear. Moreover, with algorithms like CNNs, the position of each vertebra of the spinal column may be predicted, provided the dataset is large enough.

7.5 Conclusion

The association ST parameters and the underlying spine 3D curvature was evaluated in a pilot study. Significant correlations were observed between extracted parameters and the apical vertebra location. Results indicate the potential of predicting the spine shape using the markerless ST method, reducing the reliance on radiographs.

Chapter 8 : Conclusion

8.1 Summary

The markerless surface topography (ST) technique has been previously developed to predict spinal curve severity and curve progression of two consecutive surface scans. This thesis details the further development of the ST method for idiopathic scoliosis screening. Additionally, the ST method was evaluated as a potential objective tool for torso aesthetic appearance assessment. The effect of the Schroth PSSE method added to standard care on external trunk surface asymmetries was also evaluated using posture indicators from ST technique.

8.1.1 Development of Screening Model based on ST Analysis Using a CNN Algorithm

The first objective of the thesis was to propose a model for detecting AIS from torso depth and asymmetries. A binary classification model using convolutional neural networks for detecting AIS was developed. Inputs to the model were deviations and depth of the back mapped onto grids of equal size. The proposed input features provided expanded information on surface torso asymmetries than the previous method of extracting parameters from localized areas of high deviations. The selected input features also allowed for the analysis of overall patterns of asymmetries between scoliotic and non-scoliotic individuals. In the development stage of the model, the parameters of the CNN were optimized to maximize validation accuracy.

The proposed approach detected scoliosis with a sensitivity of 97%. For ST maps associated with curves greater than 25°, the model detected 99% of the cases. Sensitivity decreased to 96% for mild curves ($\leq 25°$). The classification model identified typical asymmetry cases with a specificity of 90%. The overall accuracy of the model was 95%.

The results revealed that the proposed ST method can identify AIS based on torso depth and back deviation. In addition, the screening accuracy from the proposed model obtained a negative likelihood ratio of 0.023 and a positive likelihood ratio of 9.7. Likelihood ratios are not affected by the prevalence of AIS in the population data. The proposed model's derived likelihood ratios findings can be compared to similar population studies with different prevalence ratios. Literature on screening programs using a scoliometer and Adam’s forward bend test showed a higher positive likelihood than the ST screening model developed in this thesis. This suggests that individuals with normal asymmetries are less likely to receive a positive prediction from the CNN model than the

scoliometer and Adam's tests. Thus, the model was able reduce false positives that lead to unnecessary referrals for radiographic exams.

8.1.2 Validation and Comparison of ST Classification Model to Standard Clinical Screening Tools

Preliminary results from an ongoing screening program were also reported in this work. The proposed classification algorithm based on ST analysis of torso asymmetries was compared to the Adam's forward bend and scoliometer tests. the estimated specificity of screening adolescent participants using ST, Adam's forward bend test, and scoliometer screening modalities were 96%, 96%, and 76%, respectively. Significant difference in the specificity was observed between ST and the Adam's forward bend test, indicating that the screening test using ST was able to exclude cases from x-ray referrals and reduce the false positive rate. The specificity of the ST classification model compared to the scoliometer test were found to be comparable. However, the agreement using Cohen's kappa of the screening outcomes of the ST method compared to Adam's and scoliometer tests 0.060 and 0.113, respectively, indicating slight agreements. Due to unconfirmed outcomes from the referred participants for x-rays, the sensitivity of the screening modalities was not assessed. Overall, the proposed classification algorithm using ST method showed promise in reducing false positives and unnecessary exposure to radiation for children.

8.1.3 Evaluation of the ST Method for Assessing Torso Appearance Aesthetics

AIS patients have concerns about their self-perceived body image. The third objective of the thesis is to evaluate the ST asymmetry analysis method as a tool to assess aesthetics. This study determined the association of torso of asymmetry indices from ST analysis on patients' self-perceived condition of their back and posture. In response to conservative scoliosis treatment, we aimed to determine whether ST can detect self-perceived improvement of their torso. The minimally important difference in ST parameters, representing the smallest improvement considered worthwhile by patients, was determined. This thesis conducted a secondary analysis of an RCT that included participants randomized into standard of care or Schroth intervention added standard of care groups. The GRC, self-reported by participants after six months of treatment, can independently measure self-perceived improvement and is used as the anchor to determine the MID of ST indices outcome measures. After assessing several ST parameters, only ΔRMS and ΔMaxDev were significantly correlated with the GRC scores. The ROC analysis determined a

minimum decrease of RMS and MaxDev to detect perceived improvement of 0.27 mm and 0.49 mm. The RMS MID was able to detect perceived improvement with a sensitivity of 67% and non-improvement (no change/deterioration) with a specificity of 74%. Likewise, the MaxDev MID was able to detect perceived improvement with a sensitivity of 64% and non-improvement (no change/deterioration) with a specificity of 68%.

Additionally, the MIDs for RMS and MaxDev were each estimated based on curve types. A MID of -0.58 mm and -1.32 mm for the leading thoracic curve group for RMS and MaxDev were obtained, respectively. The RMS and MaxDev MID were each related to a sensitivity of 60% and a specificity of 53%. For the main lumbar curve group, an MID of -0.26 mm for RMS and -0.61 mm for MaxDev were obtained, respectively. The RMS MID was associated with a sensitivity and specificity of 73% and 92%, respectively. The MaxDev MID was related to a sensitivity of 68% and a specificity of 83%. Results showed that ST indices of RMS and MaxDev could discriminate perceived back condition better for AIS participants with lumbar curves, as evident by the higher sensitivity and specificity in the lumbar group compared to the thoracic curve group. The severity of the curves can account for the observed difference in MID estimates and the decrease in sensitivity and specificity for the thoracic group. The lumbar curves in this dataset were much more severe than the thoracic curves and thus were more noticeable to patients when reporting their self-assessment of their backs. Overall, the ST parameters of RMS and MaxDev were shown to have some association with perceived improvements of patients undergoing treatment.

8.1.4 Effect of Adding Schroth PSSE to Standard of Care on Posture Outcomes Using ST Method

The fourth objective, addressed in Chapter 6 of this thesis, was to assess the effectiveness of Schroth therapy added to the standard of care using ST torso asymmetry analysis. The secondary analysis of an RCT focused on the effect of the Schroth program on ST indices using mixed-effects model analysis. Participants were randomized into a control group (standard of care) or Schroth group (six months Schroth PSSE in addition to standard of care). In the per-protocol analysis, the Schroth group decreased significantly more than the control in RMS and MaxDev measurements. Likewise, in the intention-to-treat analysis, the Schroth group decreased significantly more than the control in the RMS and MaxDev parameters.

The control group was offered Schroth intervention after six months during a one-year follow-up. No significant difference in RMS and MaxDev parameters was observed. The decrease in compliance with the therapy in the control group compared to the Schroth group may have affected the changes in torso asymmetry parameters in the follow-ups.

An initial change in RMS and MaxDev measurement threshold for improvement was estimated. Based on the estimated threshold, the proportion of participants with improved ST parameters was greater in the Schroth group. However, the distribution of improved, stable, and progressed categories between groups was not significant.

Based on the MICs, 33% of patients in the control group and 59% in the Schroth group were identified as having improved torso asymmetry. 41% of the patients in the Schroth group and 68% in the control group were categorized as not improved. Significant differences existed in the distribution of these categories between the groups. Overall, the Schroth PSSE program added to standard care reduced asymmetry measurement in AIS.

8.1.5 Correlation of underlying spinal curve and the ST characteristics

In a pilot study conducted to determine the association of ST parameters with the 3D shape of the spine, seven pairs of ST and spine data from participants were analyzed. The pilot study investigated the relationship between torso surface asymmetry, measured using RMS and MaxDev, and the 3D vertebral position, precisely the Cobb angle and apex of the spinal curve. Significant correlations were found between RMS and MaxDev from the posterior torso surface and the Cobb angle, but not with the apex position in 3D. Comparisons from the anterior torso surface showed no significant correlations. However, patch centroid positions from both anterior and posterior surfaces correlated with the apical vertebra position. We have also highlighted limitations such as a small sample size and the absence of same-day x-ray and surface scan data collection.

8.2 Future Work

This thesis identified limitations in the study that stemmed from each objective, and future work can aim to address these limitations.

The development of the screening model required a significant amount of data. In this work, the number of scans of participants with AIS was greater than surface scans of negative cases, with a ratio of 2:1. Although this imbalance was mitigated during the training of the CNN model, having

a balanced dataset between groups reduces bias and overfitting. The dataset's size was insufficient to train a complex neural model with deep architectures despite implementing data augmentation. Larger datasets can lead to better accuracy results and can generalize well. Future work would include increasing the dataset and ensuring balanced proportions between groups. Transfer learning has been proposed to reduce model training time and improve generalization. Future work can include exploring various pre-trained models to develop a model that can detect idiopathic scoliosis using ST asymmetry analysis.

Significant correlations between ST torso asymmetry parameters and patients' perception of back condition were observed. The Global rate of change scale was used to gauge self-assessment of the aesthetical impacts of their spinal curve during treatment and the question asked in the GRC was broad and could be subject to interpretations. Rating their overall back condition, as prompted by the GRC question, may lead participants to examine shoulder or waist asymmetries and hip deviations. The ST parameters analyzed in Chapter 5 coincided with the curve locations and did not reflect the observable external asymmetries that patients notice the most. Future work may include investigating whether the changes in ST asymmetry parameters of the shoulder, waist, and hip are reflected in patients' self-perception of their aesthetic appearance.

The global rate of change is used to assess whether patients perceive improvement or deterioration of their health condition, which has broad applications and is not exclusive to AIS. The Walter Reed visual assessment scale and the trunk appearance perception scale (TAPS) have been designed to evaluate the physical deformity of scoliosis as perceived by patients. Future work can include determining the association of the ST technique to patient perception of their condition using validated scales specifically on AIS, such as TAPS and Walter Reed.

Short-term improvement of ST asymmetry parameters was observed when participants were prescribed the Schroth PSSE program in addition to their standard of care. However, some have reported that at least 12 months of Schroth intervention is required to improve their condition. Future work will determine the effect of a 12-month Schroth therapy on torso asymmetry and whether the improvement can be sustained after an additional 1-year follow-up.

References

Asher, M., Min Lai, S., Burton, D., & Manna, B. (2003). The reliability and concurrent validity of the scoliosis research society-22 patient questionnaire for idiopathic scoliosis. *Spine*, *28*(1), 63–69. https://doi.org/10.1097/00007632-200301010-00015

Aulisa, A. G., Guzzanti, V., Marzetti, E., Giordano, M., Falciglia, F., & Aulisa, L. (2014). Brace treatment in juvenile idiopathic scoliosis: A prospective study in accordance with the SRS criteria for bracing studies - SOSORT award 2013 winner. *Scoliosis*, *9*, 3. https://doi.org/10.1186/1748-7161-9-3

Baclig, M. M., Westover, L., & Adeeb, S. (2019). Categorizing Three-Dimensional Symmetry Using Reflection, Rotoinversion, and Translation Symmetry. *Symmetry*, *11*(9), Article 9. https://doi.org/10.3390/sym11091132

Beauséjour, M., Roy-Beaudry, M., Goulet, L., & Labelle, H. (2007). Patient characteristics at the initial visit to a scoliosis clinic: A cross-sectional study in a community without school screening. *Spine*, *32*(12), 1349–1354. https://doi.org/10.1097/BRS.0b013e318059b5f7

Belli, G., Toselli, S., Latessa, P. M., & Mauro, M. (2022). Evaluation of Self-Perceived Body Image in Adolescents with Mild Idiopathic Scoliosis. *European Journal of Investigation in Health, Psychology and Education*, *12*(3), 319–333. https://doi.org/10.3390/ejihpe12030023

Bettany-Saltikov, J., Turnbull, D., Ng, S. Y., & Webb, R. (2017). Management of Spinal Deformities and Evidence of Treatment Effectiveness. *The Open Orthopaedics Journal*, *11*, 1521–1547. https://doi.org/10.2174/1874325001711011521

Bidari, S., Kamyab, M., Kakavand, R., & Komeili, A. (2023). Performance of Surface Topography Systems in Scoliosis Management: A Narrative Review. *JPO: Journal of Prosthetics and Orthotics*, *35*(3), 208. https://doi.org/10.1097/JPO.0000000000000466

Bridwell, K. H., Shufflebarger, H. L., Lenke, L. G., Lowe, T. G., Betz, R. R., & Bassett, G. S. (2000). Parents' and patients' preferences and concerns in idiopathic adolescent scoliosis: A cross-sectional preoperative analysis. *Spine*, *25*(18), 2392–2399. https://doi.org/10.1097/00007632-200009150-00020

Bunnell, W. P. (1984). An objective criterion for scoliosis screening. *The Journal of Bone and Joint Surgery. American Volume*, *66*(9), 1381–1387.

Bunnell, W. P. (1986). The natural history of idiopathic scoliosis before skeletal maturity. *Spine*, *11*(8), 773–776. https://doi.org/10.1097/00007632-198610000-00003

Bunnell, W. P. (2005). Selective Screening for Scoliosis. *Clinical Orthopaedics and Related Research®*, *434*, 40. https://doi.org/10.1097/01.blo.0000163242.92733.66

Carman, D. L., Browne, R. H., & Birch, J. G. (1990). Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation. *JBJS*, *72*(3), 328.

Chen, K., Zhai, X., Sun, K., Wang, H., Yang, C., & Li, M. (2021). A narrative review of machine learning as promising revolution in clinical practice of scoliosis. *Annals of Translational Medicine*, *9*(1), 67. https://doi.org/10.21037/atm-20-5495

Chen, W., Khodaei, M., Reformat, M., & Lou, E. (2024). Validity of a fast automated 3d spine reconstruction measurements for biplanar radiographs: SOSORT 2024 award winner. *European Spine Journal*. https://doi.org/10.1007/s00586-024-08375-7

Cheng, J. C., Castelein, R. M., Chu, W. C., Danielsson, A. J., Dobbs, M. B., Grivas, T. B., Gurnett, C. A., Luk, K. D., Moreau, A., Newton, P. O., Stokes, I. A., Weinstein, S. L., & Burwell,

R. G. (2015). Adolescent idiopathic scoliosis. *Nature Reviews. Disease Primers*, *1*, 15030. https://doi.org/10.1038/nrdp.2015.30

Cheshire, J., Gardner, A., Berryman, F., & Pynsent, P. (2017). Do the SRS-22 self-image and mental health domain scores reflect the degree of asymmetry of the back in adolescent idiopathic scoliosis? *Scoliosis and Spinal Disorders*, *12*(1), 37. https://doi.org/10.1186/s13013-017-0144-9

Choudhry, M. N., Ahmad, Z., & Verma, R. (2016). Adolescent Idiopathic Scoliosis. *The Open Orthopaedics Journal*, *10*, 143–154. https://doi.org/10.2174/1874325001610010143

Chowanska, J., Kotwicki, T., Rosadzinski, K., & Sliwinski, Z. (2012). School screening for scoliosis: Can surface topography replace examination with scoliometer? *Scoliosis*, *7*(1), 9. https://doi.org/10.1186/1748-7161-7-9

Côté, P., Kreitz, B. G., Cassidy, J. D., Dzus, A. K., & Martel, J. (1998). A study of the diagnostic accuracy and reliability of the Scoliometer and Adam's forward bend test. *Spine*, *23*(7), 796–802; discussion 803. https://doi.org/10.1097/00007632-199804010-00011

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine: A Practical Guide*. Cambridge University Press. https://doi.org/10.1017/CBO9780511996214

De Wilde, L., Plasschaert, F., Cattoir, H., & Uyttendaele, D. (1998). Examination of the back using the Bunnell scoliometer in a Belgian school population around puberty. *Acta Orthopaedica Belgica*, *64*(2), 136–143.

Diagnosis And Treatment | Scoliosis Research Society. (n.d.). Retrieved October 26, 2023, from https://www.srs.org/Patients/Diagnosis-And-Treatment

Dunn, J., Henrikson, N. B., Morrison, C. C., Blasi, P. R., Nguyen, M., & Lin, J. S. (2018). Screening for Adolescent Idiopathic Scoliosis: Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA*, *319*(2), 173–187. https://doi.org/10.1001/jama.2017.11669

Essex, R., Bruce, G., Dibley, M., Newton, P., Thompson, T., Swaine, I., & Dibley, L. (2022). A systematic scoping review and textual narrative synthesis of the qualitative evidence related to adolescent idiopathic scoliosis. *International Journal of Orthopaedic and Trauma Nursing*, *45*, 100921. https://doi.org/10.1016/j.ijotn.2022.100921

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied Longitudinal Analysis* (2 edition). Wiley.

Fong, D. Y. T., Lee, C. F., Cheung, K. M. C., Cheng, J. C. Y., Ng, B. K. W., Lam, T. P., Mak, K. H., Yip, P. S. F., & Luk, K. D. K. (2010). A Meta-Analysis of the Clinical Effectiveness of School Scoliosis Screening. *Spine*, *35*(10), 1061–1071. https://doi.org/10.1097/BRS.0b013e3181bcc835

Fortin, C., Feldman, D. E., Cheriet, F., & Labelle, H. (2010). Validity of a quantitative clinical measurement tool of trunk posture in idiopathic scoliosis. *Spine*, *35*(19), E988-994. https://doi.org/10.1097/BRS.0b013e3181cd2cd2

Furlanetto, T. S., Sedrez, J. A., Candotti, C. T., & Loss, J. F. (2016). Photogrammetry as a tool for the postural evaluation of the spine: A systematic review. *World Journal of Orthopedics*, *7*(2), 136–148. https://doi.org/10.5312/wjo.v7.i2.136

Fusco, C., Zaina, F., Atanasio, S., Romano, M., Negrini, A., & Negrini, S. (2011). Physical exercises in the treatment of adolescent idiopathic scoliosis: An updated systematic review.

Physiotherapy Theory and Practice, *27*(1), 80–114. https://doi.org/10.3109/09593985.2010.533342

Galbusera, F., Casaroli, G., & Bassani, T. (2019). Artificial intelligence and machine learning in spine research. *JOR Spine*, *2*(1), e1044. https://doi.org/10.1002/jsp2.1044

Ghaneei, M., Ekyalimpa, R., Westover, L., Parent, E. C., & Adeeb, S. (2019). Customized k-nearest neighbourhood analysis in the management of adolescent idiopathic scoliosis using 3D markerless asymmetry analysis. *Computer Methods in Biomechanics and Biomedical Engineering*, *22*(7), 696–705. https://doi.org/10.1080/10255842.2019.1584795

Ghaneei, M., Komeili, A., Li, Y., Parent, E. C., & Adeeb, S. (2018). 3D Markerless asymmetry analysis in the management of adolescent idiopathic scoliosis. *BMC Musculoskeletal Disorders*, *19*(1), 385. https://doi.org/10.1186/s12891-018-2303-4

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Grivas, T. B., Wade, M. H., Negrini, S., O'Brien, J. P., Maruyama, T., Hawes, M. C., Rigo, M., Weiss, H. R., Kotwicki, T., Vasiliadis, E. S., Sulam, L. N., & Neuhous, T. (2007). SOSORT consensus paper: School screening for scoliosis. Where are we today? *Scoliosis*, *2*, 17. https://doi.org/10.1186/1748-7161-2-17

Grosso, C., Negrini, S., Boniolo, A., & Negrini, A. a. E. (2002). The validity of clinical examination in adolescent spinal deformities. *Studies in Health Technology and Informatics*, *91*, 123–125.

Hackenberg, L., Hierholzer, E., Pötzl, W., Götze, C., & Liljenqvist, U. (2003). Rasterstereographic back shape analysis in idiopathic scoliosis after anterior correction and fusion. *Clinical Biomechanics*, *18*(1), 1–8. https://doi.org/10.1016/S0268-0033(02)00165-1

Hacquebord, J. H., & Leopold, S. S. (2012). In Brief: The Risser Classification: A Classic Tool for the Clinician Treating Adolescent Idiopathic Scoliosis. *Clinical Orthopaedics and Related Research*, *470*(8), 2335–2338. https://doi.org/10.1007/s11999-012-2371-y

Hawes, M. C. (2003). The use of exercises in the treatment of scoliosis: An evidence-based critical review of the literature. *Pediatric Rehabilitation*, *6*(3–4), 171–182. https://doi.org/10.1080/0963828032000159202

Hill, S., Franco-Sepulveda, E., Komeili, A., Trovato, A., Parent, E., Hill, D., Lou, E., & Adeeb, S. (2014). Assessing asymmetry using reflection and rotoinversion in biomedical engineering applications. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, *228*(5), 523–529. https://doi.org/10.1177/0954411914531115

Ho, C., Parent, E., Watkins, E., Moreau, M., Hedden, D., El-Rich, M., & Adeeb, S. (2015). Asymmetry Assessment Using Surface Topography in Healthy Adolescents. *Symmetry*, *7*(3), 1436–1454. https://doi.org/10.3390/sym7031436

Hong, A., Jaswal, N., Westover, L., Parent, E. C., Moreau, M., Hedden, D., & Adeeb, S. (2017). Surface Topography Classification Trees for Assessing Severity and Monitoring Progression in Adolescent Idiopathic Scoliosis: *SPINE*, *42*(13), E781–E787. https://doi.org/10.1097/BRS.0000000000001971

Huang, S. C. (1997). Cut-off point of the Scoliometer in school scoliosis screening. *Spine*, *22*(17), 1985–1989. https://doi.org/10.1097/00007632-199709010-00007

J. Danielsson, A., & L. Nachemson, A. (2003). Back Pain and Function 22 Years After Brace Treatment for Adolescent Idiopathic Scoliosis: A Case-Control Study—Part I. *Spine*, *28*(18), 2078. https://doi.org/10.1097/01.BRS.0000084268.77805.6F

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status: Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, *10*(4), 407–415. https://doi.org/10.1016/0197-2456(89)90005-6

Jamaludin, A., Fairbank, J., Harding, I., Kadir, T., Peters, T. J., Zisserman, A., & Clark, E. M. (2020). Identifying Scoliosis in Population-Based Cohorts: Automation of a Validated Method Based on Total Body Dual Energy X-ray Absorptiometry Scans. *Calcified Tissue International*, *106*(4), 378–385. https://doi.org/10.1007/s00223-019-00651-9

Jaremko, J., Delorme, S., Dansereau, J., Labelle, H., Ronsky, J., Poncet, P., Harder, J., Dewar, R., & Zernicke, R. F. (2000). Use of Neural Networks to Correlate Spine and Rib Deformity in Scoliosis. *Computer Methods in Biomechanics and Biomedical Engineering*, *3*(3), 203–213. https://doi.org/10.1080/10255840008915265

Jaremko, J. L., Poncet, P., Ronsky, J., Harder, J., Dansereau, J., Labelle, H., & Zernicke, R. F. (2001). Estimation of Spinal Deformity in Scoliosis From Torso Surface Cross Sections. *Spine*, *26*(14), 1583–1591.

Kamper, S. J., Maher, C. G., & Mackay, G. (2009). Global rating of change scales: A review of strengths and weaknesses and considerations for design. *The Journal of Manual & Manipulative Therapy*, *17*(3), 163–170. https://doi.org/10.1179/jmt.2009.17.3.163

Kan, M. M. P., Negrini, S., Di Felice, F., Cheung, J. P. Y., Donzelli, S., Zaina, F., Samartzis, D., Cheung, E. T. C., & Wong, A. Y. L. (2023). Is impaired lung function related to spinal deformities in patients with adolescent idiopathic scoliosis? A systematic review and meta-analysis—SOSORT 2019 award paper. *European Spine Journal*, *32*(1), 118–139. https://doi.org/10.1007/s00586-022-07371-z

Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep Learning Applications in Medical Image Analysis. *IEEE Access*, *6*, 9375–9389. IEEE Access. https://doi.org/10.1109/ACCESS.2017.2788044

Kim, S., & Lee, W. (2017). Does McNemar's test compare the sensitivities and specificities of two diagnostic tests? *Statistical Methods in Medical Research*, *26*(1), 142–154. https://doi.org/10.1177/0962280214541852

Kokabu, T., Kanai, S., Kawakami, N., Uno, K., Kotani, T., Suzuki, T., Tachi, H., Abe, Y., Iwasaki, N., & Sudo, H. (2021). An algorithm for using deep learning convolutional neural networks with three dimensional depth sensor imaging in scoliosis detection. *The Spine Journal*, *21*(6), 980–987. https://doi.org/10.1016/j.spinee.2021.01.022

Komeili, A., Westover, L. M., Parent, E. C., Moreau, M., El-Rich, M., & Adeeb, S. (2014). Surface topography asymmetry maps categorizing external deformity in scoliosis. *The Spine Journal*, *14*(6), 973-983.e2. https://doi.org/10.1016/j.spinee.2013.09.032

Komeili, A., Westover, L., Parent, E. C., El-Rich, M., & Adeeb, S. (2015a). Correlation Between a Novel Surface Topography Asymmetry Analysis and Radiographic Data in Scoliosis. *Spine Deformity*, *3*(4), 303–311. https://doi.org/10.1016/j.jspd.2015.02.002

Komeili, A., Westover, L., Parent, E. C., El-Rich, M., & Adeeb, S. (2015b). Monitoring for idiopathic scoliosis curve progression using surface topography asymmetry analysis of the torso in adolescents. *The Spine Journal: Official Journal of the North American Spine Society*, *15*(4), 743–751. https://doi.org/10.1016/j.spinee.2015.01.018

Kuru, T., Yeldan, İ., Dereli, E. E., Özdinçler, A. R., Dikici, F., & Çolak, İ. (2016). The efficacy of three-dimensional Schroth exercises in adolescent idiopathic scoliosis: A randomised

controlled clinical trial. *Clinical Rehabilitation*, *30*(2), 181–190. https://doi.org/10.1177/0269215515575745

Labelle, H., Richards, S. B., De Kleuver, M., Grivas, T. B., Luk, K. D. K., Wong, H. K., Thometz, J., Beauséjour, M., Turgeon, I., & Fong, D. Y. T. (2013). Screening for adolescent idiopathic scoliosis: An information statement by the scoliosis research society international task force. *Scoliosis*, *8*, 17. https://doi.org/10.1186/1748-7161-8-17

Lau, K. K. L., Kwan, K. Y. H., Cheung, J. P. Y., Wong, J. S. H., Shea, G. K. H., Law, K. K. P., & Cheung, K. M. C. (2024). Incidence of back pain from initial presentation to 3 years of follow-up in subjects with untreated adolescent idiopathic scoliosis. *Spine Deformity*, *12*(2), 357–365. https://doi.org/10.1007/s43390-023-00794-8

Lee, C. F., Fong, D. Y. T., Cheung, K. M. C., Cheng, J. C. Y., Ng, B. K. W., Lam, T. P., Mak, K. H., Yip, P. S. F., & Luk, K. D. K. (2010). Referral criteria for school scoliosis screening: Assessment and recommendations based on a large longitudinally followed cohort. *Spine*, *35*(25), E1492-1498. https://doi.org/10.1097/BRS.0b013e3181ecf3fe

Lee, S., Lee, J. S., Kim, J. P., Kim, K., Hwang, C. H., & Koo, K. -i. (2018). Precise Cobb Angle Measurement System Based on Spinal Images Merging Function. *IRBM*, *39*(5), 343–352. https://doi.org/10.1016/j.irbm.2018.09.002

Lenke, L. G., Edwards, C. C., & Bridwell, K. H. (2003). The Lenke classification of adolescent idiopathic scoliosis: How it organizes curve patterns as a template to perform selective fusions of the spine. *Spine*, *28*(20), S199-207. https://doi.org/10.1097/01.BRS.0000092216.16155.33

Levine, G. N., Bates, E. R., Blankenship, J. C., Bailey, S. R., Bittl, J. A., Cercek, B., Chambers, C. E., Ellis, S. G., Guyton, R. A., Hollenberg, S. M., Khot, U. N., Lange, R. A., Mauri, L.,

Mehran, R., Moussa, I. D., Mukherjee, D., Ting, H. H., O'Gara, P. T., Kushner, F. G., … Zhao, D. X. (2016). 2015 ACC/AHA/SCAI Focused Update on Primary Percutaneous Coronary Intervention for Patients With ST-Elevation Myocardial Infarction: An Update of the 2011 ACCF/AHA/SCAI Guideline for Percutaneous Coronary Intervention and the 2013 ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Society for Cardiovascular Angiography and Interventions. *Circulation*, *133*(11), 1135–1147. https://doi.org/10.1161/CIR.0000000000000336

Levy, A. R., Goldberg, M. S., Mayo, N. E., Hanley, J. A., & Poitras, B. (1996). Reducing the lifetime risk of cancer from spinal radiographs among people with adolescent idiopathic scoliosis. *Spine*, *21*(13), 1540–1547; discussion 1548. https://doi.org/10.1097/00007632-199607010-00011

Mohamed, N., Acharya, V., Schreiber, S., Parent, E. C., & Westover, L. (2024). Effect of adding Schroth physiotherapeutic scoliosis specific exercises to standard care in adolescents with idiopathic scoliosis on posture assessed using surface topography: A secondary analysis of a Randomized Controlled Trial (RCT). *PLOS ONE*, *19*(4), e0302577. https://doi.org/10.1371/journal.pone.0302577

Morais, T., Bernier, M., & Turcotte, F. (1985). Age- and sex-specific prevalence of scoliosis and the value of school screening programs. *American Journal of Public Health*, *75*(12), 1377–1380.

Nahm, F. S. (2022). Receiver operating characteristic curve: Overview and practical use for clinicians. *Korean Journal of Anesthesiology*, *75*(1), 25–36. https://doi.org/10.4097/kja.21209

Nash, C. L., Gregg, E. C., Brown, R. H., & Pillai, K. (1979). Risks of exposure to X-rays in patients undergoing long-term treatment for scoliosis. *The Journal of Bone and Joint Surgery. American Volume*, *61*(3), 371–374.

Negrini, S., Aulisa, A. G., Aulisa, L., Circo, A. B., de Mauroy, J. C., Durmala, J., Grivas, T. B., Knott, P., Kotwicki, T., Maruyama, T., Minozzi, S., O'Brien, J. P., Papadopoulos, D., Rigo, M., Rivard, C. H., Romano, M., Wynne, J. H., Villagrasa, M., Weiss, H.-R., & Zaina, F. (2012). 2011 SOSORT guidelines: Orthopaedic and Rehabilitation treatment of idiopathic scoliosis during growth. *Scoliosis*, *7*(1), 3. https://doi.org/10.1186/1748-7161-7-3

Negrini, S., Donzelli, S., Aulisa, A. G., Czaprowski, D., Schreiber, S., de Mauroy, J. C., Diers, H., Grivas, T. B., Knott, P., Kotwicki, T., Lebel, A., Marti, C., Maruyama, T., O'Brien, J., Price, N., Parent, E., Rigo, M., Romano, M., Stikeleather, L., … Zaina, F. (2018). 2016 SOSORT guidelines: Orthopaedic and rehabilitation treatment of idiopathic scoliosis during growth. *Scoliosis and Spinal Disorders*, *13*, 3. https://doi.org/10.1186/s13013-017-0145-8

Negrini, S., Fusco, C., Minozzi, S., Atanasio, S., Zaina, F., & Romano, M. (2008). Exercises reduce the progression rate of adolescent idiopathic scoliosis: Results of a comprehensive systematic review of the literature. *Disability and Rehabilitation*, *30*(10), 772–785. https://doi.org/10.1080/09638280801889568

Negrini, S., Grivas, T. B., Kotwicki, T., Maruyama, T., Rigo, M., Weiss, H. R., & Members of the Scientific society On Scoliosis Orthopaedic and Rehabilitation Treatment (SOSORT). (2006). Why do we treat adolescent idiopathic scoliosis? What we want to obtain and to

avoid for our patients. SOSORT 2005 Consensus paper. *Scoliosis*, *1*, 4. https://doi.org/10.1186/1748-7161-1-4

Negrini, S., Hresko, T. M., O'Brien, J. P., Price, N., SOSORT Boards, & SRS Non-Operative Committee. (2015). Recommendations for research studies on treatment of idiopathic scoliosis: Consensus 2014 between SOSORT and SRS non–operative management committee. *Scoliosis*, *10*(1), 8. https://doi.org/10.1186/s13013-014-0025-4

Otman, S., Kose, N., & Yakut, Y. (2005). The efficacy of Schroth s 3-dimensional exercise therapy in the treatment of adolescent idiopathic scoliosis in Turkey. *Saudi Medical Journal*, *26*(9), 1429–1435.

Pan, Y., Chen, Q., Chen, T., Wang, H., Zhu, X., Fang, Z., & Lu, Y. (2019). Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays. *European Spine Journal: Official Publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, *28*(12), 3035–3043. https://doi.org/10.1007/s00586-019-06115-w

Parent, E. C., Damaraju, S., Hill, D. L., Lou, E., & Smetaniuk, D. (2010). Identifying the best surface topography parameters for detecting idiopathic scoliosis curve progression. *Studies in Health Technology and Informatics*, *158*, 78–82.

Parent, E. C., Wong, D., Hill, D., Mahood, J., Moreau, M., Raso, V. J., & Lou, E. (2010). The Association Between Scoliosis Research Society-22 Scores and Scoliosis Severity Changes at a Clinically Relevant Threshold. *Spine*, *35*(3), 315. https://doi.org/10.1097/BRS.0b013e3181cabe75

Paria, N., & Wise, C. A. (2015). Genetics of adolescent idiopathic scoliosis. *Seminars in Spine Surgery*, *27*(1), 9–15. https://doi.org/10.1053/j.semss.2015.01.004

Płaszewski, M., & Bettany-Saltikov, J. (2014). Are current scoliosis school screening recommendations evidence-based and up to date? A best evidence synthesis umbrella review. *European Spine Journal: Official Publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, *23*(12), 2572–2585. https://doi.org/10.1007/s00586-014-3307-x

Ponseti, I. V., Pedrini, V., Wynne-Davies, R., & Duval-Beaupere, G. (1976). Pathogenesis of scoliosis. *Clinical Orthopaedics and Related Research*, *120*, 268–280.

Ramirez, L., Durdle, N. G., Raso, V. J., & Hill, D. L. (2006). A support vector machines classifier to assess the severity of idiopathic scoliosis from surface topography. *IEEE Transactions on Information Technology in Biomedicine*, *10*(1), 84–91. IEEE Transactions on Information Technology in Biomedicine. https://doi.org/10.1109/TITB.2005.855526

Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, *61*(2), 102–109. https://doi.org/10.1016/j.jclinepi.2007.03.012

Richards, B. S., Bernstein, R. M., D'Amato, C. R., & Thompson, G. H. (2005). Standardization of criteria for adolescent idiopathic scoliosis brace studies: SRS Committee on Bracing and Nonoperative Management. *Spine*, *30*(18), 2068–2075; discussion 2076-2077. https://doi.org/10.1097/01.brs.0000178819.90239.d0

Roach, J. W. (1999). Adolescent idiopathic scoliosis. *The Orthopedic Clinics of North America*, *30*(3), 353–365, vii–viii. https://doi.org/10.1016/s0030-5898(05)70092-4

Rogala, E. J., Drummond, D. S., & Gurr, J. (1978). Scoliosis: Incidence and natural history. A prospective epidemiological study. *The Journal of Bone and Joint Surgery. American Volume*, *60*(2), 173–176.

Romano, M., Minozzi, S., Zaina, F., Saltikov, J. B., Chockalingam, N., Kotwicki, T., Hennes, A. M., & Negrini, S. (2013). Exercises for adolescent idiopathic scoliosis: A Cochrane systematic review. *Spine*, *38*(14), E883-893. https://doi.org/10.1097/BRS.0b013e31829459f8

Ronckers, C. M., Doody, M. M., Lonstein, J. E., Stovall, M., & Land, C. E. (2008). Multiple diagnostic X-rays for spine deformities and risk of breast cancer. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *17*(3), 605–613. https://doi.org/10.1158/1055-9965.EPI-07-2628

Rubin, D. B. (2004). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons.

Sanders, J. O., W. Polly, D. J., Cats-Baril, W., Jones, J., Lenke, L. G., O'Brien, M. F., Stephens Richards, B., Sucato, D. J., & Group, M. of the A. S. of the S. D. S. (2003). Analysis of Patient and Parent Assessment of Deformity in Idiopathic Scoliosis Using the Walter Reed Visual Assessment Scale. *Spine*, *28*(18), 2158. https://doi.org/10.1097/01.BRS.0000084629.97042.0B

Sapkas, G., Papagelopoulos, P. J., Kateros, K., Koundis, G. L., Boscainos, P. J., Koukou, U. I., & Katonis, P. (2003). Prediction of Cobb angle in idiopathic adolescent scoliosis. *Clinical Orthopaedics and Related Research*, *411*, 32–39. https://doi.org/10.1097/01.blo.0000068360.47147.30

Scherer, D., Müller, A., & Behnke, S. (2010). Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition. In K. Diamantaras, W. Duch, & L. S. Iliadis (Eds.), *Artificial Neural Networks – ICANN 2010* (pp. 92–101). Springer. https://doi.org/10.1007/978-3-642-15825-4_10

Schreiber, S., Parent, E. C., Hedden, D. M., Moreau, M., Hill, D., & Lou, E. (2014). Effect of Schroth exercises on curve characteristics and clinical outcomes in adolescent idiopathic scoliosis: Protocol for a multicentre randomised controlled trial. *Journal of Physiotherapy*, *60*(4), 234; discussion 234. https://doi.org/10.1016/j.jphys.2014.08.005

Schreiber, S., Parent, E. C., Hill, D. L., Hedden, D. M., Moreau, M. J., & Southon, S. C. (2019). Patients with adolescent idiopathic scoliosis perceive positive improvements regardless of change in the Cobb angle – Results from a randomized controlled trial comparing a 6-month Schroth intervention added to standard care and standard care alone. SOSORT 2018 Award winner. *BMC Musculoskeletal Disorders*, *20*(1), 319. https://doi.org/10.1186/s12891-019-2695-9

Schreiber, S., Parent, E. C., Kawchuk, G. N., & Hedden, D. M. (2023). Algorithm for Schroth-Curve-Type Classification of Adolescent Idiopathic Scoliosis: An Intra- and Inter-Rater Reliability Study. *Children*, *10*(3), Article 3. https://doi.org/10.3390/children10030523

Schreiber, S., Parent, E. C., Khodayari Moez, E., Hedden, D. M., Hill, D. L., Moreau, M., Lou, E., Watkins, E. M., & Southon, S. C. (2016). Schroth Physiotherapeutic Scoliosis-Specific Exercises Added to the Standard of Care Lead to Better Cobb Angle Outcomes in Adolescents with Idiopathic Scoliosis – an Assessor and Statistician Blinded Randomized Controlled Trial. *PLOS ONE*, *11*(12), e0168746. https://doi.org/10.1371/journal.pone.0168746

Schreiber, S., Parent, E. C., Moez, E. K., Hedden, D. M., Hill, D., Moreau, M. J., Lou, E., Watkins, E. M., & Southon, S. C. (2015). The effect of Schroth exercises added to the standard of care on the quality of life and muscle endurance in adolescents with idiopathic scoliosis—an assessor and statistician blinded randomized controlled trial: "SOSORT 2015 Award Winner." *Scoliosis*, *10*(1), 24. https://doi.org/10.1186/s13013-015-0048-5

Schreiber, S., Parent, E., Watkins, E., & Hedden, D. (2012). An algorithm for determining scoliosis curve type according to Schroth. *Scoliosis*, *7*(Suppl 1), O53. https://doi.org/10.1186/1748-7161-7-S1-O53

Schumann, K., Püschel, I., Maier-Hennes, A., & Weiss, H.-R. (2008). Postural changes in patients with scoliosis in different postural positions revealed by surface topography. *Studies in Health Technology and Informatics*, *140*, 140–143.

Smith, P. L., Donaldson, S., Hedden, D., Alman, B., Howard, A., Stephens, D., & Wright, J. G. (2006). Parents' and Patients' Perceptions of Postoperative Appearance in Adolescent Idiopathic Scoliosis. *Spine*, *31*(20), 2367. https://doi.org/10.1097/01.brs.0000240204.98960.dd

Soucacos, P. N., Zacharis, K., Gelalis, J., Soultanis, K., Kalos, N., Beris, A., Xenakis, T., & Johnson, E. O. (1998). Assessment of curve progression in idiopathic scoliosis. *European Spine Journal*, *7*(4), 270–277. https://doi.org/10.1007/s005860050074

Sperandio, E. F., Alexandre, A. S., Yi, L. C., Poletto, P. R., Gotfryd, A. O., Vidotto, M. C., & Dourado, V. Z. (2014). Functional aerobic exercise capacity limitation in adolescent idiopathic scoliosis. *The Spine Journal*, *14*(10), 2366–2372. https://doi.org/10.1016/j.spinee.2014.01.041

SRS Bracing Manual | Scoliosis Research Society. (n.d.). Retrieved October 26, 2023, from https://www.srs.org/Education/Manuals-and-Presentations/SRS-Bracing-Manual

Stolinski, L., Kotwicki, T., Czaprowski, D., Chowanska, J., & Suzuki, N. (2012). Analysis of the Anterior Trunk Symmetry Index (ATSI). Preliminary report. *Studies in Health Technology and Informatics*, *176*, 242–246.

Stone, L. E., Upasani, V. V., Pahys, J. M., Fletcher, N. D., George, S. G., Shah, S. A., Bastrom, T. P., Bartley, C. E., Lenke, L. G., Newton, P. O., Kelly, M. P., & Harms Study Group. (2023). SRS-22r Self-Image After Surgery for Adolescent Idiopathic Scoliosis at 10-year Follow-up. *Spine*, *48*(10), 683–687. https://doi.org/10.1097/BRS.0000000000004620

Su, X., Dong, R., Wen, Z., & Liu, Y. (2022). Reliability and Validity of Scoliosis Measurements Obtained with Surface Topography Techniques: A Systematic Review. *Journal of Clinical Medicine*, *11*(23), 6998. https://doi.org/10.3390/jcm11236998

Theologis, T. N., Jefferson, R. J., Simpson, A. H., Turner-Smith, A. R., & Fairbank, J. C. (1993). Quantifying the cosmetic defect of adolescent idiopathic scoliosis. *Spine*, *18*(7), 909–912. https://doi.org/10.1097/00007632-199306000-00016

Thulbourne, T., & Gillespie, R. (1976). The rib hump in idiopathic scoliosis. Measurement, analysis and response to treatment. *The Journal of Bone and Joint Surgery. British Volume*, *58*(1), 64–71. https://doi.org/10.1302/0301-620X.58B1.1270497

Torell, G., Nordwall, A., & Nachemson, A. (1981). The changing pattern of scoliosis treatment due to effective screening. *The Journal of Bone and Joint Surgery. American Volume*, *63*(3), 337–341.

Upasani, V. V., Caltoum, C., Petcharaporn, M., Bastrom, T. P., Pawelek, J. B., Betz, R. R., Clements, D. H., Lenke, L. G., Lowe, T. G., & Newton, P. O. (2008). Adolescent idiopathic

scoliosis patients report increased pain at five years compared with two years after surgical treatment. *Spine*, *33*(10), 1107–1112. https://doi.org/10.1097/BRS.0b013e31816f2849

US Preventive Services Task Force. (2018). Screening for Adolescent Idiopathic Scoliosis: US Preventive Services Task Force Recommendation Statement. *JAMA*, *319*(2), 165–172. https://doi.org/10.1001/jama.2017.19342

Watanabe, K., Aoki, Y., & Matsumoto, M. (2019). An Application of Artificial Intelligence to Diagnostic Imaging of Spine Disease: Estimating Spinal Alignment From Moiré Images. *Neurospine*, *16*(4), 697–702. https://doi.org/10.14245/ns.1938426.213

Watkins, E., Bosnjak, S., & Parent, E. (2012). Algorithms to prescribe Schroth exercises for each of four Schroth curve types. *Scoliosis*, *7*(Suppl 1), P22. https://doi.org/10.1186/1748-7161-7-S1-P22

Weinstein, S. L., Dolan, L. A., Cheng, J. C., Danielsson, A., & Morcuende, J. A. (2008). Adolescent idiopathic scoliosis. *The Lancet*, *371*(9623), 1527–1537. https://doi.org/10.1016/S0140-6736(08)60658-3

Weiss, H.-R. (2011). The method of Katharina Schroth—History, principles and current development. *Scoliosis*, *6*, 17. https://doi.org/10.1186/1748-7161-6-17

Weiss, H.-R., Hollaender, M., & Klein, R. (2006). ADL based scoliosis rehabilitation—The key to an improvement of time-efficiency? *Studies in Health Technology and Informatics*, *123*, 594–598.

Weiss, H.-R., Negrini, S., Hawes, M. C., Rigo, M., Kotwicki, T., Grivas, T. B., & Maruyama, T. (2006). Physical exercises in the treatment of idiopathic scoliosis at risk of brace treatment – SOSORT consensus paper 2005. *Scoliosis*, *1*, 6. https://doi.org/10.1186/1748-7161-1-6

Wong, H.-K., & Tan, K.-J. (2010). The natural history of adolescent idiopathic scoliosis. *Indian Journal of Orthopaedics*, *44*(1), 9–13. https://doi.org/10.4103/0019-5413.58601

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging*, *9*(4), Article 4. https://doi.org/10.1007/s13244-018-0639-9

Yang, J., Zhang, K., Fan, H., Huang, Z., Xiang, Y., Yang, J., He, L., Zhang, L., Yang, Y., Li, R., Zhu, Y., Chen, C., Liu, F., Yang, H., Deng, Y., Tan, W., Deng, N., Yu, X., Xuan, X., … Lin, H. (2019). Development and validation of deep learning algorithms for scoliosis screening using back images. *Communications Biology*, *2*, 390. https://doi.org/10.1038/s42003-019-0635-8

Zaina, F., Negrini, S., & Atanasio, S. (2009). TRACE (Trunk Aesthetic Clinical Evaluation), a routine clinical tool to evaluate aesthetics in scoliosis patients: Development from the Aesthetic Index (AI) and repeatability. *Scoliosis*, *4*(1), 3. https://doi.org/10.1186/1748-7161-4-3

Zaina, F., Negrini, S., Fusco, C., & Atanasio, S. (2009). How to improve aesthetics in patients with Adolescent Idiopathic Scoliosis (AIS): A SPoRT brace treatment according to SOSORT management criteria. *Scoliosis*, *4*, 18. https://doi.org/10.1186/1748-7161-4-18

Appendix A:

Convolutional Neural Network Model Parameter Optimization Procedure

Hyperparameter optimization analysis was performed to identify convolutional neural network (CNN) parameters to maximize the accuracy of the model for detecting AIS based on the ST asymmetry map. In Chapter 3 of this thesis, surface topography (ST) quantifying torso surface asymmetry has been introduced as a potential method for AIS screening. However, asymmetries between adolescents with scoliosis and those with a typical developing spine are unknown. Convolutional neural network algorithms (CNN) may be applied to detect and classify positive AIS and negative AIS from ST asymmetry maps. CNN has emerged as a powerful machine learning tool often implemented in image classification, object detection, and segmentation tasks. CNN uses a variety of building blocks, including convolution layers, pooling layers, and fully connected layers, to learn features from data using back-propagation. Despite their effectiveness, optimizing CNNs is required to maximize training and validation accuracy and avoid overfitting. The hyperparameters, including the learning rate, regularization parameters, and overall architecture, are varied and tested to obtain the optimal combination of hyperparameter values.

Dataset

Surface scans of participants with AIS (n=273) and those with typical spine development during growth (n=285) were used to optimize the CNN model accuracy. Follow-up surface scans of participants were also included, which increased the dataset to 697 surface scans of AIS cases and 298 scans of typical spine development. The distribution of the dataset is outlined in Chapter 3. In addition, 20% of the data were reserved for testing the model, while the remaining data was used for training and validation. Data augmentation was implemented on the training set to increase the data, as described in Chapter 3.

Data Preprocessing

Analysis of the surface scan was conducted to obtain ST maps used as inputs to the CNN model to detect AIS, as described in Chapter 3. The 3D torso model was duplicated for asymmetry analysis and reflected about the sagittal plane. The reflected torso was aligned with the original to minimize the distance between the models. The deviations and the magnitude of the distances between each point on the original torso and its corresponding point on the reflected torso were

determined. Only the deviations of the back of the torso (180-degree span) were retained. The depth was also obtained and retained for only the back of the original torso. Deviation and depth information were down-sampled and mapped on 102 x 102 matrices and were used as inputs to a CNN model to classify asymmetry patterns observed in adolescents with typical spine development and those with AIS.

CNN Parameters Optimization

Various hyperparameters were evaluated during the optimization process. In general, the architecture of a CNN model for binary classification tasks consists of convolutional layers with an activation function applied to weights, a pooling layer following each convolutional layer, fully connected layers with a specified number of units in the connected layers for feature extraction, and a final output value. Table A.1 presents a list of fixed parameters and hyperparameters for optimization.

Table A. 1: Fixed and optimization parameters with of the CNN classifier algorithm during training

Fixed Parameters		Value
	Activation Function	ReLU
	Pooling Operation	Max pooling
	Pooling Size	2x2
	Number of Epochs	300
	Dropout Rate	20%
Optimization Parameters		**Value**
	Learning Rate	$\{1 \cdot 10^{-x} \mid n \in Z, 1 \leq n \leq 4\}$
	Batch Size	$\{2^n \mid n \in Z, 4 \leq n \leq 7\}$
	Number of Convolutional Layers	$\{2, 3\}$
	Number of Filters (Optimized for each Convolutional Layer)	$\{2^n \mid n \in Z, 4 \leq n \leq 7\}$
	Convolutional Filter Size (Optimized for each Layer)	$\{(2n + 1) \times (2n + 1) \mid n \in Z, 1 \leq n \leq 3\}$
	Number of Fully Connected Layers	$\{2, 3\}$
	Number of Neurons in Fully Connected Layers	$\{10n \mid n \in Z, 1 \leq n \leq 40\}$
	L2 Regularization	$\{1 \cdot 10^{-x} \mid n \in Z, 1 \leq n \leq 4\}$

The number of convolution layer levels was $\{2, 3\}$. Higher convolutional layers in the architecture are not recommended for small datasets, like in this case, for developing the model for AIS screening. Filter size per convolutional layer were selected from $\{(2n + 1) \times (2n + 1) \mid n \in Z, 1 \leq n \leq 3\}$. The number of filters per convolution layer selected from $\{2^n \mid n \in Z, 4 \leq n \leq 7\}$. In general, the number of filters in the succeeding convolutional layer is greater than the number of filters in the previous convolution layer. The batch size during each epoch during training was selected from $\{2^n \mid n \in Z, 4 \leq n \leq 7\}$. In addition, the learning rate was selected from $\{1 \cdot 10^{-x} \mid n \in Z, 1 \leq n \leq 4\}$. The number of fully connected layers was selected from $\{2, 3\}$, and the number of neurons in the connected layers was drawn from $\{10n \mid n \in Z, 1 \leq n \leq 40\}$. The number of neurons decreases in

the next connected layer compared to the previous fully connected layer. Finally, the L2 regularization to reduce overfitting was selected from $\{1 \cdot 10^{-x} \mid n \in Z, 1 \leq n \leq 4\}$. Architecture representation of the parameters for optimization are shown in Figure 1

Table A.1 also presents the fixed parameters during optimization. A linear rectified unit (ReLu) was selected as the activation function for each convolutional layer. The Adam function was selected as the optimizer to minimize the loss during training. A pooling layer size of 2x2 was selected after each convolution layer. The maximum number of epochs was set to 300. Early stopping criteria were implemented to avoid overfitting, where model training stops if the validation loss does not improve after 15 epochs.

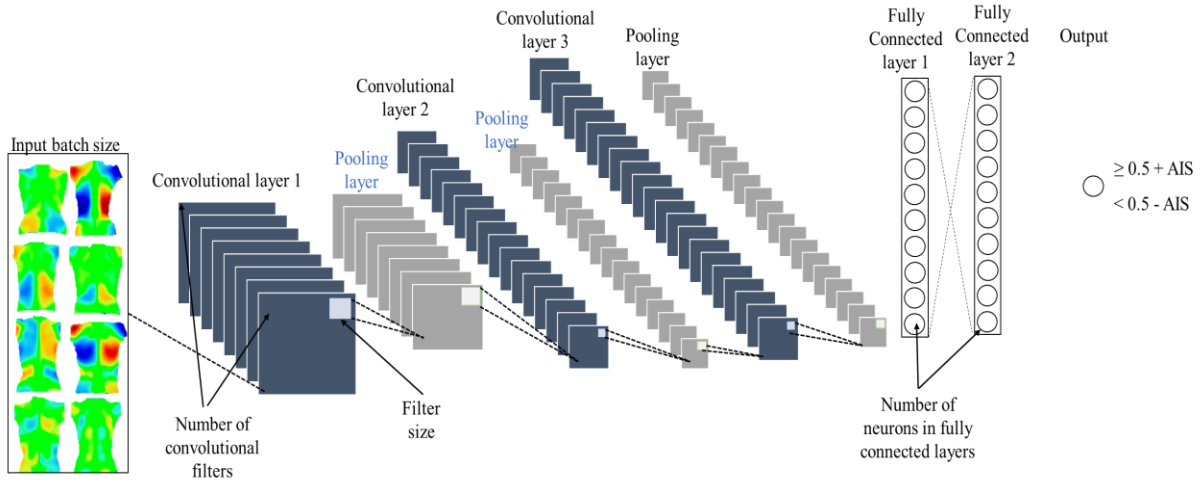


Fig A. 1: Representation of the CNN architecture and the selected parameters for optimization

The optimization procedure consists of training the CNN with various hyperparameters value combinations. For each parameter combination assessed, 10-fold cross-validation was applied, and the ten-fold' average training and validation accuracy were determined. A comparison of the average validation accuracy was made between the hyperparameter combination sets. The hyperparameter combination set drawn from the values outlined in Table A.1 with the highest validation accuracy was selected as the optimal parameter for the CNN model to detect AIS. Finally, the random search technique was applied which involves randomly selecting values for each hyperparameter from predefined ranges and evaluating the performance of the neural network with these values, allowing for a broad exploration of the hyperparameter space.

Optimized Parameters

The trial was selected with highest validation accuracy. Table A.2 shown below are the parameter values that obtained a validation accuracy of 98%.

Table A. 2: Optimized parameters of the CNN classifier value

Optimization Parameters	Value
Learning Rate	0.001
Batch Size	32
Number of Convolutional Layers	2
Number of Filters	
1ST layer	16
2nd layer	32
Convolutional Filter Size	
1ST layer	3x3
2nd layer	3x3
Number of Fully Connected Layers	2
Number of Neurons in Fully Connected Layers	
1ST layer	250
2nd layer	118
L2 Regularization	0.01

Appendix B:

Linear Mixed Effects Model Analysis to Determine the Effect of Schroth PPSE on Torso Asymmetry

Linear mixed effects model coefficients, standard error, and significance values for estimating RMS and MaxDev in the analysis between groups. Model outputs for intention to treat and per protocol analysis are reported. Covariables of the mixed effects model are the following: Age, height, weight, if participants are braced, and their curve classification.

Linear Mixed Effects Model Outputs for Estimating RMS						
	Intention to Treat			Per Protocol		
Term	**Estimate**	**Standard Error**	**p-value**	**Estimate**	**Standard Error**	**p-value**
Intercept	-1.26	6.62	0.85	2.19	8.63	0.80
[Group=0]*[Time=3]	-1.25	0.71	0.08	-1.66	0.98	0.09
[Group=0]*[Time=2]	-0.94	0.75	0.21	-1.212	0.97	0.21
[Group=0]*[Time=1]	-1.33	0.73	0.07	-1.53	0.98	0.12
[Group=1]*[Time=3]	-1.17	0.33	**0.001**	-1.63	0.44	**0.0004**
[Group=1]*[Time=2]	-0.62	0.33	0.06	-0.72	0.42	0.09
Age	1.08	0.28	**0.0001**	1.13	0.35	**0.001**
Height	-0.02	0.06	0.79	-0.04	0.08	0.57
Weight	0.00	0.05	0.97	0.003	0.08	0.97
Braced	1.26	0.71	0.08	1.38	0.95	0.15
3cp Curve	3.01	1.18	**0.01**	2.34	1.59	0.15
4c Curve	-0.73	1.37	0.59	-1.99	1.78	0.26
4cp Curve	1.11	1.16	0.34	1.49	1.51	0.32
Linear Mixed Effects Model Outputs for Estimating MaxDev						
	Intention to Treat			Per Protocol		
Term	**Estimate**	**Standard Error**	**p-value**	**Estimate**	**Standard Error**	**p-value**
Intercept	-5.17	11.11	0.64	0.80	14.69	0.96
[Group=0]*[Time=3]	-2.58	1.19	**0.03**	-4.17	1.65	**0.01**
[Group=0]*[Time=2]	-1.64	1.22	0.18	-2.73	1.64	0.10
[Group=0]*[Time=1]	-2.47	1.20	**0.04**	-3.14	1.65	0.06
[Group=1]*[Time=3]	-1.93	0.52	**0.0002**	-2.93	0.72	**0.0001**
[Group=1]*[Time=2]	-1.20	0.51	**0.02**	-1.71	0.68	**0.013**
Age	1.74	0.46	**0.0002**	1.83	0.59	**0.002**
Height	-0.01	0.10	0.90	-0.07	0.13	0.60
Weight	0.00	0.09	0.99	0.04	0.13	0.79
Braced	0.82	1.19	0.49	1.32	1.62	0.42
3cp Curve	3.86	1.95	**0.05**	3.17	2.72	0.25
4c Curve	-2.20	2.26	0.33	-3.47	3.03	0.25
4cp Curve	1.24	1.91	0.52	2.51	2.57	0.33

Post hoc comparison of estimated marginal mean from fitted linear mixed effects models of RMS and MaxDev outcomes. Comparisons for intention to treat and per protocol analyses reported. Control group coded a 0, Schroth group coded as 1. Baseline, 3 months, and 6 months variables were coded as Time =1, Time=2, and Time=3, respectively.

RMS Estimated Mean Difference Comparison				
	Intention to treat		**Per Protocol**	
Comparison	**Mean Estimate Difference**	**p-value**	**Mean Estimate Difference**	**p-value**
Group=0,Time=3 - Group=0,Time=2	-0.31	0.44	-0.45	1.00
Group=0,Time=3 - Group=0,Time=1	0.08	0.81	-0.13	1.00
Group=0,Time=3 - Group=1,Time=3	0.19	0.80	-0.04	1.00
Group=0,Time=3 - Group=1,Time=2	-0.37	0.62	-0.95	1.00
Group=0,Time=3 - Group=1,Time=1	-0.99	0.18	-1.66	0.76
Group=0,Time=2 - Group=0,Time=1	0.39	0.30	0.32	1.00
Group=0,Time=2 - Group=1,Time=3	0.50	0.53	0.42	1.00
Group=0,Time=2 - Group=1,Time=2	-0.06	0.94	-0.49	1.00
Group=0,Time=2 - Group=1,Time=1	-0.68	0.38	-1.21	0.97
Group=0,Time=1 - Group=1,Time=3	0.10	0.89	0.09	1.00
Group=0,Time=1 - Group=1,Time=2	-0.45	0.55	-0.82	1.00
Group=0,Time=1 - Group=1,Time=1	-1.07	0.16	-1.53	0.85
Group=1,Time=3 - Group=1,Time=2	-0.55	0.12	-0.91	0.40
Group=1,Time=3 - Group=1,Time=1	-1.17	**0.001**	-1.63	**0.01**
Group=1,Time=2 - Group=1,Time=1	-0.62	0.06	-0.72	0.77
MaxDev Estimated Mean Difference Comparison				
	Intention to Treat		**Per Protocol**	
Comparison	**Mean Estimate Difference**	**p-value**	**Mean Estimate Difference**	**p-value**
Group=0,Time=3 - Group=0,Time=2	-0.95	0.16	-1.44	0.48
Group=0,Time=3 - Group=0,Time=1	-0.11	0.85	-1.03	0.94
Group=0,Time=3 - Group=1,Time=3	-0.30	0.81	-1.24	1.00
Group=0,Time=3 - Group=1,Time=2	-1.03	0.40	-2.46	0.89
Group=0,Time=3 - Group=1,Time=1	-2.23	0.07	-4.17	0.17
Group=0,Time=2 - Group=0,Time=1	0.84	0.15	0.41	1.00
Group=0,Time=2 - Group=1,Time=3	0.65	0.60	0.20	1.00
Group=0,Time=2 - Group=1,Time=2	-0.08	0.95	-1.02	1.00
Group=0,Time=2 - Group=1,Time=1	-1.28	0.30	-2.73	0.79
Group=0,Time=1 - Group=1,Time=3	-0.19	0.88	-0.21	1.00
Group=0,Time=1 - Group=1,Time=2	-0.92	0.45	-1.43	1.00
Group=0,Time=1 - Group=1,Time=1	-2.12	0.09	-3.14	0.60
Group=1,Time=3 - Group=1,Time=2	-0.73	0.18	-1.22	0.68
Group=1,Time=3 - Group=1,Time=1	-1.93	**0.0002**	-2.93	**0.001**

Group=1,Time=2 - Group=1,Time=1	-1.20	0.02	-1.71	0.18

Linear mixed effects model coefficients, standard error, and significance values for estimating RMS and MaxDev in the analysis within controls when receiving delayed Schroth exercises in that last 6 months of the 1-year follow-up. Model outputs for intention to treat and per protocol analysis are reported. Covariables of the mixed effects model are the following: Age, height, weight, if participants are braced, and their curve classification.

Mixed Effect Model Outputs for Estimating RMS

	Intention to Treat			Per Protocol		
Term	**Estimate**	**Standard Error**	**p-value**	**Estimate**	**Standard Error**	**p-value**
Intercept	4.44	8.78	0.61	-4.92	14.26	0.73
Time=3	0.58	0.49	0.25	-0.77	0.57	0.19
Time=2	0.50	0.55	0.38	-0.52	0.52	0.32
Age	-0.06	0.08	0.41	1.35	0.54	**0.01**
Height	0.08	0.07	0.30	-0.06	0.13	0.64
Weight	0.64	0.94	0.49	0.12	0.10	0.26
Braced	2.20	1.74	0.21	1.23	1.25	0.33
3cp Curve	-0.89	2.09	0.67	1.28	2.56	0.62
4c Curve	1.29	1.70	0.45	0.49	3.28	0.88
4cp Curve	4.44	8.78	0.61	1.18	2.49	0.64

Mixed Effect Model Outputs for Estimating MaxDev

	Intention to Treat			Per Protocol		
Term	**Estimate**	**Standard Error**	**p-value**	**Estimate**	**Standard Error**	**p-value**
Intercept	0.95	15.67	0.95	-10.97	26.16	0.68
Time=3	1.35	0.70	0.06	-0.76	0.93	0.42
Time=2	0.47	0.71	0.51	-1.17	0.83	0.16
Age	-0.07	0.14	0.62	2.50	0.99	**0.01**
Height	0.07	0.13	0.56	-0.07	0.23	0.76
Weight	-0.13	1.70	0.94	0.11	0.19	0.56
Braced	2.04	3.23	0.53	0.85	2.30	0.71
3cp Curve	-3.18	3.57	0.37	-2.72	4.70	0.57
4c Curve	0.60	3.07	0.84	-3.13	6.02	0.61
4cp Curve	0.95	15.67	0.95	-1.77	4.57	0.70