CAN ARTIFICIAL INTELLIGENCE PREDICT GROWTH AND TREATMENT OUTCOMES AMONG ORTHODONTIC PATIENTS

A Dissertation

by

SAEED NASSER S ASIRI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Peter H. Buschang
Committee Members,	Larry P. Tadlock
	Reginald W. Taylor
	Emet Schneiderman
Head of Department,	Madhu K. Nair

May 2021

Major Subject: Oral Biology

ABSTRACT

The objectives of the present study were first, to synthesize the literature pertaining to artificial intelligence (AI) and machine learning (ML) applications in orthodontics, second to evaluate the possibility of predicting mandibular growth using artificial intelligence, third to assess the applicability of using artificial intelligence to predict dental treatment outcomes among Herbst patients, and finally, to predict skeletal treatment outcomes among Herbst patients using artificial intelligence. The first study was a narrative review that assessed the orthodontic literature pertaining to applications of AI and ML in orthodontics. The second study assessed the applicability of a ML method known as decision trees (DTs) for predicting maxillomandibular relationships over a five-year period using radiographs of 222 untreated subjects. The third study used DTs to predict dental treatment outcomes among 150 Herbst patients. The fourth study used a subset of 116 patients from the third study to assess possibility of using DTs to predict skeletal outcomes among Herbst patients. The first study showed that several applications of AI in orthodontics have been done, and more specifically for diagnosis and treatment planning, followed by predicting treatment outcomes, and predicting growth. The second study showed that DTs were able to successfully classify the growth of untreated subjects 85.4% of the time with the Y-axis as the most important variable for prediction. The third study demonstrated that DTs can accurately predict dental treatment outcomes among Herbst patients 81.4% of the time, and identified SN-MP, followed by overbite, and L1-MP, respectively, as the most important variables. The fourth study showed that skeletal outcomes among Herbst patients can be accurately predicted approximately 87.9% of the time. It also identified

the facial convexity angle, followed by the distance from U1 to facial plane, articular angle, and Wits, respectively, as the most important variables.

DEDICATION

This work is dedicated to my family and many friends. A special feeling of gratitude to my loving parents, Nasser and Reeah Asiri for their enduring love, support, and patience throughout my education journey. I dedicate this work to my brothers and sisters who have never left my side and are very special.

I also dedicate this dissertation to my many friends who have supported me throughout the process. I will always appreciate all they have done, especially Ahmed Aljoudi for being beside me during the past couple of years.

I dedicate this work and give special thanks to my children, Deema and Danah. You have made me stronger, better, and more fulfilled than I could have ever imagined. I love you to the moon and back.

This dissertation is dedicated to my wife, Haifa, who has been a constant source of support and encouragement during the challenges of graduate school and life. I am truly thankful for having you in my life. Thank you for encouraging me to follow my dreams and chasing them with me and keeping me positive throughout the past 8 years.

ACKNOWLEDGEMENTS

Before all, I praise and thank God for providing me with the inspiration, perseverance, and strength to pursue my dreams.

I would like to thank my committee chair, Prof. Peter H. Buschang, and my committee members, Dr. Larry Tadlock, Reginald W. Taylor, and Prof. Emet Schneiderman, for their guidance and support throughout the course of this research. Thanks also go to my friends and colleagues and the Biomedical Science Department faculty and staff for making my time at Texas A&M University a great experience. I would like to thank my sponsor (Prince Sattam Bin Abdulaziz University) for their eight years of scholarship to pursue my graduate studies.

Thank you to the two private practice orthodontists who allowed me to utilize their patient records.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Peter Buschang [advisor], Dr. Larry Tadlock, and Dr. Reginald Taylor of the Department of Orthodontics, and Professor Emet Schneiderman of the Department of Biomedical Science.

All work conducted for the dissertation was completed by the student independently.

Funding Sources

There are no outside funding contributions to acknowledge related to the research and compilation of this project.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Statement of the Problem

The terms artificial intelligence (AI) and machine learning (ML) are being used widely in in orthodontics nowadays. They are used interchangeably to describe the new tools utilized by orthodontists to help in diagnosis, treatment planning, and predicting growth and treatment outcomes. However, these terms do not mean exactly the same thing. AI is the general term that includes several subfields like reasoning, natural language processing, planning, and machine learning. [1] On the other hand, ML is a specific type of AI that is concerned with making computers learn from previous examples to perform better when introduced to new data. [2] This confusion among clinicians needs to be clarified in order to better understand ML and its different subtypes. By doing so, orthodontists can explore the possible applications of ML in orthodontics and how can that help the efficiency in the orthodontic clinic.

Applications of AI and more specifically ML in orthodontics is numerous. One area that received much attention is diagnosis and treatment planning. Several studies have been done to help orthodontists practice more efficiently and help them decide which teeth to extract for example. [3-5] Another area that received some attention is growth prediction. Using pre-treatment records, ML was successful in classifying patients into favorable and unfavorable growth patterns. [6-8] Moreover, ML models have been also used to predict treatment outcomes. [9-11] Unfortunately, there has not been a thorough review of the literature that summarized the work that has been done in those areas and how accurate these models can be to traditional statistics.

Growth Predictions

Several studies have attempted to predict growth in the orthodontic literature. Earlier studies were not accurate as they assumed that pattern extension is established early in life and the same average growth increments should be added to all patients after adjustments for age and growth. [12-14] However, this believe was found later to be not true when Bjork and coworkers found younger subjects changing from bad anteriorposterior (AP) relationships to good AP relationships over time, and vice versa. [15] Subsequently, growth predictions using multilevel models were used which developed individualized growth curves and worked well with missing data. [16, 17] Unfortunately, these methods were very complex and required longitudinal data to develop the models and make the predictions.

AI and Growth Patterns

After the successful application of AI and ML in different fields, several algorithms have been implemented to help orthodontists classify their patients' growth patterns. [12, 18, 19] For instance, an artificial neural network was developed to help classifying the growth of 43 untreated children based on changes in size and shape. [6] Nevertheless, the model was not validated on an external sample and the sample size was small. Recently, a study developed a model using cephalometric variables to classify patients' craniofacial growth into normal and abnormal. [20] They found that support vector machines algorithm could accurately classify abnormal growth patterns

99.8% of the time. Unfortunately, this study did not mention the sample size and focused on the technical aspect of developing the system. Furthermore, support vector machines were also used to classify normal or abnormal skeletal patterns based on craniomaxillary measures and they were correct only 74.5% of the time. [7] However, they excluded information about the mandible because their aim was to develop a mandibular classification system for the skeletonized remains where the mandibular bone is usually lost. Including this information in future studies could enhance the accuracy of these models.

Growth patterns classification of class III subjects has also been done. Using longitudinal data of untreated class III subjects who were classified as either good or bad growers based on the change in the sagittal relationship, a decision tree (DT) was developed. Using the same 11 cephalometric variables, the DT showed a significantly lower rate of misclassification (12.0%) than discriminant analysis (40.7%). [8] The model was able to successfully identify good and bad growth patterns 64.0% of the time when tested on new data. It would be interesting to see the accuracy of a similar approach on patients with different growth patterns, such as class I and class II subjects.

Herbst for Correcting Class II Malocclusions

One of the most common problems in the US population is bilateral class II malocclusion as it has been reported to occur in approximately 15% of the population. [21] Among those, approximately 75% of them have also class II skeletal malocclusion. [22] Those individuals usually present with mandibular retrusion and facial convexity. [23] In order to improve esthetic, the position of the chin need to be brought forward to create less convex or flatter profile. [24] The class II dental malocclusion is generally not self-correcting, different treatment modalities that correct this malocclusion should pay attention to factors that can affect soft issue esthetic and final dental and skeletal treatment outcomes.

One of the different methods used to correct class II malocclusion includes the Herbst. In 1905, Emil Herbst introduced his appliance at the international congress of dentistry in Berlin, Germany. However, it did not receive much acceptance at the time. In 1979, Pancherz reintroduced the Herbst appliance and brought attention to the possibility of mandibular growth stimulation. [25] It works by continuously posturing the mandible forward to stimulate or redirect mandibular growth. [26] Over the next years, several papers were published that evaluated the biological effects of the Herbst appliance [27] and quantified the skeletal, dental, and soft tissue changes associated with the correction of class II malocclusion. [28, 29] However, patient's response to functional appliance treatment may vary as has been shown in the literature. [23, 30] This could be attributed to several factors such as treatment timing and amount of growth remaining, [31, 32] whether the patient is hyper or hypo divergent, [33, 34] the pretreatment SNB, [35] and the ANB difference. [36]

It is clear that orthodontists need to have a tool that can help them distinguish between patients who might benefit from treatment with Herbst from those who will get worse or not improve. However, there is limited information in the literature about the predictability of treatment outcomes in patients treated with Herbst. Only one study developed a model for patients treated with Herbst and they were combined with patients treated with Twin blocks to predict changes in mandibular length. [37] They determined the gonial angle (Co-Go-Me°) as the only pretreatment predictor for mandibular length changes, with patients responding favorably if the gonial angle is below 125.5° and responding poorly if the angle is greater than 125.5°. However, their sample included subjects treated with 2 different functional appliances which both have been shown to have different effects on posterior facial height and maxillomandibular differential. [38] In addition, their primary outcome was predicting the change in mandibular length which is not the best indicator of treatment success. A subject can show an increase in the total mandibular length, but the final outcome might still be considered as unsuccessful if the patient is hyperdivergent to begin with and the vertical dimensions increase. This could result in backward rotation, and therefore unsuccessful treatment outcomes. [33] This has led to the question of what measures should be used to evaluate the success of treatment?

Successful vs Unsuccessful Treatment Outcomes

It has been stated before that an orthodontic treatment is considered successful when both the objective treatment goals and subjective patient desires are achieved. [39] In treatment with functional appliances, the definition of successful treatment outcomes differed among different studies. Most of the studies used occlusal measurements, mainly the correction of overjet and molar relations to distinguish between successful and unsuccessful treatment. [34, 36, 40-42] Few studies used a skeletal measure to evaluate treatment outcomes with functional appliances. [37, 40] Some studies used a composite treatment outcome with a criterion that have to be met in order to be classified as a successful outcome. [42-44]

One of the studies that used dental measurements reported reduction of overjet as an indicator of successful treatment outcome. [45] If the overjet was reduced to 3mm or less than one-third of its original value, a treatment was considered successful. Another study considered a case to be progressing successfully if the overjet was reduced by 50 percent or 6mm within 6 months. [41] Success was defined by achieving a bilateral class I molar relationship at a specified time (either at the end of phase I or at the end of treatment) in one study. [46]

Few studies used a skeletal measure to evaluate treatment outcomes with functional appliances. [37, 40] One study used ANB to determine successful and unsuccessful treatment outcomes with functional appliance treatment. [40] They divided individuals who demonstrated a reduction in ANB angle of 3.0° or more as favorable change, and individuals having a reduction less than 0.5° as less favorable. Their primary purpose was to identify any differences in the pretreatment variables important for predictions which will be easier to identify by having the two groups at the either ends of the spectrum. In another study, the change in total mandibular length (Co-Gn) was used to evaluate successful treatment outcomes. [37] Good responders were identified as subjects who showed a biannual increase > 5.3mm, while bad responders were determined if they had a biannual increase in Co-Gn ≤ 5.3mm. However, total mandibular length is not the best indicator of successful skeletal outcomes as explained before. The total mandibular length can increase but the treatment outcome might still be considered as unsuccessful especially in hyperdivergent patients. [33]

Some studies used a composite outcome variable to evaluate treatment. A case was considered “satisfactory” if at the end of the treatment a patient had a neutral occlusion (± 1mm), overjet and overbite less than 4mm, no observable rotation of upper incisors, occlusal contact on all teeth, crowding in lower arch not exceeding 1mm, and rotation of cuspids/premolars not exceeding 15 degrees and limited to one tooth. [43] Correction of overbite, presence of acceptable profile with upper and lower lips within 1 SD on Ricketts’ esthetic line, and the absence or little relapse for cases followed up 2-5 years after the functional appliance therapy in addition to correction of overjet and molar relationship were used in a study to evaluate treatment outcomes. [42] In the large randomized clinical trial that compared Herbst with twin-block functional appliance, they used three different criteria to evaluate treatment outcomes. [44] They used the anterorposterior skeletal discrepancy as described by the Pancherz analysis, [25] the overjet, and the Peer Assesment Rating (PAR) score to describe treatment outcomes.

AI and Treatment Outcomes

Recently, AI was implemented in orthodontics to predict treatment outcomes among class II and class III patients. Using artificial neural networks, post-treatment Peer Assessment Ratings (PAR) were predicted among class II patients based on their pre-treatment PAR index. [10] The neural network model showed better accuracy (94.0%) than linear regression (82.0%) at predicting the final PAR score. Another model was also developed to predict treatment outcomes among untreated class III patients.

[11] A machine learning model was also used to cluster patients as hypermandibular, hyperdivergent or balanced based on their cephalometric variables. When the model was applied to a treated sample, all of the unsuccessful cases belonged to either the hypermandibular or the hyperdivergent cluster. Furthermore, a ML model was found to accurately predict the prognosis of class III treatment slightly better than discriminant analysis (DA) (97.2% for ML and 92.1% for DA). [47] All these promising results using AI encourage us to us it as a prediction tool that can improve the accuracy of treatment outcomes using traditional statistical methods.

Decision Trees

Decision trees are among the most popular ML models used for classification. A major reason for their popularity is their ease of interpretation, even for non-experts. [48, 49] They can handle a mixture of categorical and continuous variables, even when there are missing values. They can also classify multiple groups. [49] A tree is built by asking a series of a yes-or-no questions. Each question has a node that leads to two other nodes, depending on the answer. The process continues until the terminal node is reached, which is the final outcome. [49] The first uppermost node pertains to the most important variable, followed by the next most important variable, and so on, until a hierarchy of variables is created. [50] For researchers, building decision trees is less time consuming than other classification algorithms. [51] Due to these reasons, DTs have often been used in medicine and bioinformatics. DTs applied in orthodontics to classify untreated Cass III subjects as good or bad growers have shown a significantly lower rate of misclassification (12.0%) than discriminant analysis (40.7%). [8]

Specific Aims

The present study will address the following aims:

Aim 1: To synthesize the literature pertaining to AI and ML in orthodontics.

Aim 2: To evaluate the possibility of using decision trees to predict favorable and unfavorable growth types from a single cephalogram.

Aim 3: To assess the possibility to predict dental treatment outcomes after orthodontic treatment with Herbst in class II patients using decision trees from pre-treatment records.

Aim 4: To assess the possibility to predict skeletal treatment outcomes after orthodontic treatment with Herbst in class II patients using decision trees from pre-treatment records.

References

1. Mueller JP and Massaron L. *Machine learning for dummies*. 2016: John Wiley & Sons.

2. Marsland S. *Machine learning: an algorithmic perspective*. Second ed. 2014: Cambridge, UK: Chapman and Hall/CRC.

3. Jung, S.-K. and T.-W. Kim, *New approach for the diagnosis of extractions with neural network machine learning.* Am J Orthod Dentofacial Orthop, 2016. 149(1): p. 127-133.

4. Suhail Y, Upadhyay M, Chhibber A, Kshitiz. *Machine Learning for the Diagnosis of Orthodontic Extractions: A Computational Analysis Using Ensemble Learning.* Bioengineering, 2020. 7(2): p. 55.

5. Xie X, Wang L, Wang A. *Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment.* Angle Orthod, 2010. 80(2): p. 262-266.

6. Lux CJ, Stellzig A, Volz D, Jäger W, Richardson A, Komposch G., *A neural network approach to the analysis and classification of human craniofacial growth.* Growth Dev Aging, 1998. 62(3): p. 95-106.

7. Niño-Sandoval TC, Guevara Perez SV, González FA, Jaque RA, Infante-Contreras C. *An automatic method for skeletal patterns classification using craniomaxillary variables on a Colombian population.* Forensic Sci Int, 2016. 261: p. 159.e1-6.

8. Auconi P, Scazzocchio M, Caldarelli G, Nieri M, McNamara JA, Franchi L. *Understanding interactions among cephalometrics variables during growth in untreated Class III subjects.* Eur J Orthod, 2017. 39(4): p. 395-401.

9. Patcas R, Bernini DAJ, Volokitin A, Agustsson E, Rothe R, Timofte R. *Applying artificial intelligence to assess the impact of orthognathic treatment on facial attractiveness and estimated age.* Int J Oral Maxillofac Surg, 2019. 48(1):77-83.

10. Zarei A, El-Sharkawi M, Hairfield M, King G. *An intelligent system for prediction of orthodontic treatment outcome*. The 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 2006, p. 2702-2706.

11. Auconi P, Scazzocchio M, Cozza P, McNamara JA Jr, Franchi L. *Prediction of Class III treatment outcomes through orthodontic data mining.* Eur J Orthod, 2015. 37(3): p. 257-67.

12. Broadbent** BH. *The face of the normal child.* Angle Orthod, 1937. 7(4): p. 183-208.

13. Brodie, A.G., *On the growth pattern of the human head. From the third month to the eighth year of life.* Am. J. Anat., 1941. 68(2): p. 209-262.

14. Johnston LE. *A simplified approach to prediction.* Am J Orthod, 1975. 67(3): p. 253-257.

15. Björk A, Pallilng M. *Adolescent age changes in sagittal jaw relation, alveolar prognathy, and incisal inclination.* Acta Odontol Scand, 1955. 12(3-4): p. 201-232.

16. Chvatal BA, Behrents RG, Ceen RF, Buschang PH. *Development and testing of multilevel models for longitudinal craniofacial growth prediction.* Am J Orthod Dentofacial Orthop, 2005. 128(1): p. 45-56.

17. Roberts RO. *Adolescent maxillomandibular relationships: Growth patterns, inter-individual variability, and predictions*. 2007: The Texas A&M University System Health Science Center.

18. Brodie AG. *Behavior of normal and abnormal facial growth patterns.* Am J Orthod Oral Surg, 1941. 27(11): p. 633-647.

19. Ricketts RM. *A principle of arcial growth of the mandible.* Angle Orthod, 1972. 42(4): p. 368-386.

20. Lakkshmanan A, Shri AA, Aruna E. *Pattern classification for finding facial growth abnormalities*. 2013 IEEE International Conference on Computational Intelligence and Computing Research, Enathi, India, 2013, pp. 1-5

21. Proffit WR, Fields HW Jr, Moray LJ. *Prevalence of malocclusion and orthodontic treatment need in the United States: estimates from the NHANES III survey.* Int J Adult Orthod Orthognath Surg, 1998. 13: p. 97-106.

22. Milacic M, Markovic M. *A comparative occlusal and cephalometric study of dental and skeletal anteroposterior relationships.* Br J Orthod, 1983. 10(1): p. 53-54.

23. Cozza P, Baccetti T, Franchi L, De Toffol L, McNamara JA Jr. *Mandibular changes produced by functional appliances in Class II malocclusion: a systematic review.* Am J Orthod Dentofacial Orthop, 2006. 129(5): p. 599.e1-12.

24. LaHaye MB, Buschang PH, Alexander RG, Boley JC. *Orthodontic treatment changes of chin position in Class II Division 1 patients.* Am J Orthod Dentofacial Orthop, 2006. 130(6): p. 732-741.

25. Pancherz H. *Treatment of Class II malocclusions by jumping the bite with the Herbst appliance: a cephalometric investigation.* Am J Orthod, 1979. 76(4): p. 423-442.

26. Barnett GA, Higgins DW, Major PW, Flores-Mir C. *Immediate Skeletal and Dentoalveolar Effects of the Crown-or Banded Type Herbst Appliance on Class II division 1 Malocclusion: A Systematic Review.* Angle Orthod, 2008. 78(2): p. 361-369.

27. Pancherz H. *The Herbst appliance—its biologic effects and clinical use.* Am J Orthod, 1985. 87(1): p. 1-20.

28. Pancherz H. *The mechanism of Class II correction in Herbst appliance treatment: a cephalometric investigation.* Am J Orthod, 1982. 82(2): p. 104-113.

29. Alves PF, Oliveira AG. *A comparison of the skeletal, dental, and soft tissue effects caused by herbst and mandibular protraction appliances in the treatment of mandibular Class II malocclusions.* World J Orthod, 2008. 9(1): e1-19.

30. Chen JY, Will LA, Niederman R. Niederman, *Analysis of efficacy of functional appliances on mandibular growth.* Am J Orthod Dentofacial Orthop, 2002. 122(5): p. 470-476.

31. Hägg U, Pancherz H. *Dentofacial orthopaedics in relation to chronological age, growth period and skeletal development. An analysis of 72 male patients with Class II division 1 malocclusion treated with the Herbst appliance.* Eur J Orthod, 1988. 10(3): p. 169-176.

32. Baccetti T, Franchi L, Toth LR, McNamara JA Jr. *Treatment timing for Twin-block therapy.* Am J Orthod Dentofacial Orthop, 2000. 118(2): p. 159-170.

33. Rogers K, Campbell PM, Tadlock L, Schneiderman E, Buschang PH. *Treatment changes of hypo-and hyperdivergent Class II Herbst patients.* Angle Orthod, 2018. 88(1): p. 3-9.

34. Pancherz H. *The mandibular plane angle in activator treatment.* Angle Orthod, 1979. 49(1): p. 11-20.

35. Parkhouse RC. *A cephalometric appraisal of cases of Angle's Class II, Division I malocclusion treated by the Andresen appliance.* Dent Pract Dent Rec, 1969. 19(12): p. 425-433.

36. Ahlgren J, Laurin C. *Late results of activator-treatment: a cephalometric study.* Br J Orthod, 1976. 3(3): p. 181-187.

37. Franchi L, Baccetti T. *Prediction of individual mandibular changes induced by functional jaw orthopedics followed by fixed appliances in Class II patients.* Angle Orthod, 2006. 76(6): p. 950-954.

38. Schaefer AT, McNamara JA Jr, Franchi L, Baccetti T. *A cephalometric comparison of treatment with the Twin-block and stainless steel crown Herbst appliances followed by fixed appliance therapy.* Am J Orthod Dentofacial Orthop, 2004. 126(1): p. 7-15.

39. Burstone CJ. *Lip posture and its significance in treatment planning.* Am J Orthod, 1967. 53(4): p. 262-284.

40. Patel HP, Moseley HC, Noar JH. *Cephalometric determinants of successful functional appliance therapy.* Angle Orthod, 2002. 72(5): p. 410-417.

41. Caldwell S, Cook P. *Predicting the outcome of twin block functional appliance treatment: a prospective study.* Eur J Orthod, 1999. 21(5): p. 533-539.

42. Ahn SJ, Kim JT, Nahm DS. *Cephalometric markers to consider in the treatment of Class II Division 1 malocclusion with the bionator.* Am J Orthod Dentofacial Orthop, 2001. 119(6): p. 578-586.

43. Bondevik O. *How effective is the combined activator-headgear treatment?* Eur J Orthod, 1991. 13(6): p. 482-485.

44. O'Brien K, Wright J, Conboy F, Sanjie Y, Mandall N, Chadwick S, Connolly I, Cook P, Birnie D, Hammond M, Harradine N, Lewis D, McDade C, Mitchell L, Murray A, O'Neill J, Read M, Robinson S, Roberts-Harry D, Sandler J, Shaw I. *Effectiveness of treatment for Class II malocclusion with the Herbst or twin-*

block appliances: a randomized, controlled trial. Am J Orthod Dentofacial Orthop, 2003. 124(2): p. 128-137.

45. Cohen AM. *A study of class II division 1 malocclusions treated by the Andresen appliance.* Br J Orthod, 1981. 8(3): p. 159-163.

46. Wheeler TT, McGorray SP, Dolce C, Taylor MG, King GJ. *Effectiveness of early treatment of Class II malocclusion.* Am J Orthod Dentofacial Orthop, 2002. 121(1): p. 9-17.

47. Kim BM, Kang BY, Kim HG, Baek SH. *Prognosis prediction for class III malocclusion treatment by feature wrapping method.* Angle Orthod, 2009. 79(4): p. 683-691.

48. Kotsiantis SB. *Decision trees: a recent overview.* Artif Intell Rev, 2013. 39(4): p. 261-283.

49. Kingsford C, Salzberg SL. *What are decision trees?* Nat Biotechnol, 2008. 26(9): p. 1011.

50. Song YY, Lu Y. *Decision tree methods: applications for classification and prediction.* Shanghai Arch Psychiatry, 2015. 27(2): p. 130.

51. Shafer J, Agrawal R, Mehta M. *SPRINT: A scalable parallel classifier for data mining*. In Proc. of the 22nd Int'l Conference on Very Large Databases, Bombay, India, September 1996.

CHAPTER II

APPLICATIONS OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN ORTHODONTICS*

Introduction

Artificial intelligence (AI) is a subfield of computer science concerned with developing computers and programs that have the ability to perceive information, reason and ultimately convert that information into intelligent actions. [1-3] AI as a science is very broad and encompasses various fields, including reasoning, natural language processing, planning and machine learning (ML). [4] Currently, machine learning is the most commonly used AI application in the medical and dental fields.[1]

Work in AI started back in 1943, [5] but it was not until 1956 that the term "artificial intelligence" was first used during a conference held at Dartmouth College. [6] A few years later, the term "machine learning" was officially applied to a checkers-playing program, considered one of the first successful self-learning tools. [7] Drawing from other fields such as statistics, mathematics, physics, biology, neuroscience and psychology, [8-11] artificial intelligence and machine learning progressed quickly.

One of the most important aspects of any intelligent system is learning. Learning is the process of improving performance or behavior by practice and experience. [12] Similarly, ML is concerned with making machines and computers capable of learning

---

* Reprinted with permission from, "Applications of artificial intelligence and machine learning in orthodontics", by Asiri SN, Tadlock LP, Schneiderman E, Buschang PH. APOS Trends in Orthodontics, 2020;10(1):17-24. Copyright [2019] by Scientific Scholar on behalf of APOS Trends in Orthodontics.

from previous experiences, data or examples. By utilizing a mixture of statistical and probabilistic tools, machines can learn from previous examples and improve their actions when new data is introduced. This could be in the form of predictions, identifying new patterns or classifying new data. [8] It is important to note that ML is not intended to mimic human behavior. Instead, it supplements human intelligence by doing tasks that are beyond human capabilities. [13] This is what makes ML superior to the rule-based expert systems that were used in the past.

Expert systems (ES) are considered among the earliest applications of artificial intelligence. As the name implies, the knowledge about a specific field is transferred from humans to computers, allowing people to consult the computer. [14] In other words, ES act as consultants that can process the input information and provide solutions based on if-then rules. ES have been used widely for diagnosis and treatment planning in medicine, [15] dentistry, [16] and orthodontics. [17] ES also facilitate the transfer of knowledge to different people in different places. However, rule-based ES are limited to information available at the time that the system was developed. Continuous updates are required to ensure that the information is correct and current. Due to the availability of more advanced technologies, such as machine learning, it is now possible to overcome the limitations associated with rule-based expert systems.

Most algorithms used in machine learning are also being used in data mining. The difference lies in the algorithm's goal. If the goal is to optimize decisions, then the algorithms are applied to large historical data sets to look for new patterns or relationships. [18, 19] This process is called data mining. For example, data mining can help clinical practitioners find valuable information within existing patient records. By using this new information, practitioners can optimize future decisions, improve their daily practice and increase the quality of care. On the other hand, if the goal is to make predictions, then machine learning should be applied. The clinical practitioner uses available data about a certain disease to train the machine to make predictions about the diagnosis or prognosis of patients that have never been seen before. Importantly, machine learning predictive models have proven to be more accurate than statistical models. [20] The aim of the present narrative review was twofold: 1) to introduce the various types of machine learning 2) and show orthodontists how ML has been and is currently being applied. The literature was systematically searched using MEDLINE (through PubMed) and ProQuest databases, covering both the published and unpublished literature reported in English. The studies covered are comprehensive with respect to orthodontic applications.

Type of Machine Learning

Machine learning algorithms are divided into three main categories [10] (Figure 1) based on the nature of learning and the desired outcome of the algorithm:

A. *Supervised learning:* Supervised learning is mainly used for classification when the data is discrete (categorical) and for prediction (regression) if the data is continuous. It is supervised because it is based on a known outcome. With this type of learning, a model is built using a labeled set of training data (independent variables) and a known outcome (dependent) variable. [21] Since the final outcome is known, the system learns by receiving feedback signals that either confirm or reject its performance. If the algorithm encounters new input data, it will use the training data sets to link the new input data to the desired outcome. A very common example of supervised learning for classification is e-mail spam detection, where the algorithm is trained to classify newly received emails as spam or not spam. For prediction, supervised learning can be used to predict the Graduate Record Examinations (GRE) scores, for example, based on several independent variables that are related to the outcome variable, such as study time.

B. *Unsupervised learning:* This type of learning is mainly used to discover the structure of the data in order to find meaningful information. Clustering (sometimes called unsupervised classification) is the method used with this type of learning to explore the data and then organize them into groups based on similarities or relationships between variables. [21] Unlike supervised learning, the data are not labeled, and the final outcome is not known. This type of learning allows marketers to develop programs that are specific to each group of customers after clustering them based on similar interests and features. The clusters could be based on sex, age group or demographics.

C. *Reinforcement learning:* This type of learning is similar to supervised learning in that the system is provided with a feedback signal. However, the feedback signal does not provide the true value. Instead, it rewards the system based on its interaction with a dynamic environment (n.b. reinforcement learning is also known as the reward system). The system does not know anything about the behavior of the environment. By doing multiple exploratory trials-and-errors, the system learns and improves its future performance. An example of this type of learning is the chess engine. Depending on the situation (i.e., the environment), the machine decides on certain moves and will be rewarded by either winning or losing. [21]

Major Machine Learning Algorithms and Dentistry

There are several machine learning algorithms that have been used in the dental fields. Depending on the goal, the type and amount of data, different algorithms can be used. For example, if a practitioner wants to distinguish between patients who need treatment and those who do not, he/she probably would need to use a classification algorithm (e.g., support vector machine, naïve Bayes etc.) (Table 1). However, if there are many variables and a large amount of data, an algorithm like neural networks is better suited because it can handle noisy data and perform predictions even if the relationships between variables are non-linear.

Interestingly, almost all machine learning algorithms applied in orthodontics have used the supervised learning method. Most applications have sought to automate clinical procedures that perform or facilitate diagnosis and treatment planning. These applications require training with data that has a known and desired outcome, so when introduced to new data, the ML system will use the training data to predict the new input.

Artificial Intelligence and Orthodontics

Dentistry in general and orthodontic specifically has applied artificial intelligence to solve many different problems. Earlier attempts to use AI in dentistry and orthodontics were in the form of knowledge based expert systems. These systems were mainly aimed at helping non-specialist dentists develop diagnoses and treatment plans.

[22-25] These expert systems were useful in countries like England, where hospital-based orthodontists had long waiting lists and were seeing more patients than their counterparts in Europe and the US. Due to the decline in the incidence of caries that occurred at that time, dentists treated the more straightforward cases identified by the expert system and referred the more complex cases to orthodontists. However, these systems were limited because they only had been introduced to simple cases (i.e. they could not function well with new cases not already stored in the system). Currently, general dentists have more advanced ML systems available to them that can diagnose a broader range of orthodontic cases and determine treatment needs. [26] Several advanced systems have been developed to help orthodontists diagnose and treatment plan and evaluate treatment outcomes and growth.

Machine Learning for Diagnosis and Orthodontic Treatment Planning

One of the dilemmas during treatment planning is deciding whether or not to extract, with substantial variability between the orthodontists' decisions. [27] This has led to the development of several decision-support systems that reduce the subjectivity of making these decisions. Artificial neural networks (ANN) [28-30] have been used to develop such systems, and they were found to be successful at predicting the extraction decision 80% [28] of the time in one study, and 93% [29, 30] of the time in another two studies. Prediction of the detailed extraction patterns (i.e., which teeth needed to be extracted) was also shown to be possible 84% [29] of the time in one study and 83% [30] of the time in another study. Recently, a paper used ANN to also help with identifying the anchorage requirements in the cases that were determined to need extractions by the system and it was accurate 83% of the time. [30]

X-ray analysis, an integral part of diagnosis and treatment planning, has also benefited from machine learning. One of the most important applications of ML in orthodontics was the automation of landmark detections. A recent systematic review reported 5-15% better accuracy of landmark detection with machine learning than traditional methods. [31] Machine learning was also used to automate the diagnosis directly from cephalograms, including the sagittal relationships between maxilla and mandible, as well as normal and abnormal posterior-anterior facial heights ratios, overbite and overjet. [32]

Automation of x-rays analysis has also been extended to hand and wrist radiographs for estimating skeletal age. Determining the growth status of patients is essential for deciding whether or not to utilize growth during treatment. [33] A ML system applied to a sample of 360 images showed an average difference of 0.39 years between its estimate and skeletal age estimated by two expert radiologists. [34] Another study using a larger sample of 1100 images reported an average difference of 0.60 years, when compared to the readings of two experienced radiologists. [35] One study comparing the performance of different algorithms to estimate skeletal age reported a root mean square error (RMSE) of 0.24 years with ANN, and 0.25 years with genetic algorithm when compared to traditional estimates of skeletal age. [36]

Taking panoramic radiographs makes orthodontists legally liable if they overlook diagnosing a lesion or a tumor. This has led to the development of an automated neural network system that can correctly diagnose ameloblastomas and keratocystic odontogenic tumors from panoramic radiographs 83.0% of the time. [37] Five oral and maxillofacial surgeons who examined the same radiographs correctly diagnosed the problems 82.9% of the time. The difference lies in the time needed for diagnosis. The ML system required an average of 38 seconds, while the surgeons needed 23.1 minutes for each diagnosis. Another system was developed that successfully predicted odontogenic cysts, dentigurous cysts, osteomyletits, periapical cysts, and ameloblastoms 90.6%, 90.9%, 99.4 %, 89.6%, and 100% of the time respectively. [38] Currently, more and more orthodontists are using cone beam computed tomography (CBCT), which has led to the development of an automated system using the support vector machine to correctly diagnose periapical cysts and keratocystic odontogenic tumors 100% of the time. [39] Neural networks was used to estimate patients' dental ages from panoramic radiographs. [40] Its RMSE was 0.9 for girls and 1.1 for boys, while traditional regression had a RMSE of 1.3 and 1.4 for girls and boys, respectively. [40]

Panoramic and lateral cephalometric x-rays have also been used to predict maxillary canine impactions based on angular and linear measures. [41] The highest prediction accuracy was obtained with a random forest algorithm, which correctly predicted the actual eruption status of canines 88.3% of the time.

One of the challenges for less experienced orthodontists is the selection of the appropriate treatment modality and appliance, including headgears. To address this, a system was developed to help orthodontists select the headgears that should be used. [42] Compared to the selections made by 8 expert orthodontists, the system correctly identified the appropriate headgears 95.6% of the time. Recently, decision support systems were developed to determine the geometry of orthodontic springs used to close extraction spaces [43] and to determine the forces needed to align teeth, [44] but neither system has been applied clinically.

Another orthodontic challenge during treatment planning is predicting the size of unerupted teeth. To address this, a hybrid system using artificial neural networks and genetic algorithms was used to predict canines' and premolars' sizes. [45] Its maximum error was 2.4 mm in the mandible and 1.6 mm in the maxilla. The errors were often half as large as the error produced with linear regression prediction models.

Machine Learning and Treatment Outcomes

One of the more useful applications of AI in orthodontics is the prediction of soft tissues treatment outcomes. Recently, ANN was used to predict the change in lip curvature after orthodontic treatment with or without extractions. [46] Its prediction of change and the actual change that occurred differed by 29.6% and 7% for the upper and lower lips, respectively. Both predictions were much better than those based on linear regression.

The topic of beauty is controversial because it is subjective and affected by factors such as age, sex, and ethnic backgrounds. Using artificial neural networks, facial attractiveness was quantified on a scale from 0-100 (0 extremely unattractive; 100 extremely attractive) before and after orthognathic surgery. [47] The difference between the pre-and-post surgery scores was shown to be statistically significant, with facial attractiveness improving 74.7%.

Prediction of treatment outcomes in class II and class III patients have also been reported. Using artificial neural networks, predictive models were developed to predict the post-treatment Peer Assessment Rating (PAR) index in class II patients based on their pre-treatment PAR index. [48] The neural network model used in this system was able to correctly predict the final PAR score 94.0% of the time; linear regression was correct only 82.0% of the time. A system has also been developed to predict outcomes in untreated class III patients. [49] Unsupervised learning was used to cluster patients as hypermandibular, hyperdivergent or balanced based on cephalometric variables. The system was then applied to a treated sample and found that all the unsuccessful cases belonged to either the hypermandibular or the hyperdivergent cluster. Another system was able to correctly predict the prognosis of class III treatment 97.2% of the time, which was slightly better than the 92.1% reported for discriminant analysis. [50]

Machine Learning and Growth Patterns

Several methods have been introduced to help orthodontists classify their patients' growth patterns. [51-53] In 1998, an artificial neural network was used to classify the growth of 43 untreated children based on size and shape changes. [54] However, the system was not validated on an external sample. A recent study used cephalometric variables to classify patients' craniofacial growth into normal and abnormal. [55] It showed that support vector machines could correctly classify abnormal growth patterns 99.8% of the time. Another study using support vector machines to classify normal or abnormal skeletal patterns based on craniofacial measures was correct only 74.5% of the time. [56]

Classification of class III growth patterns has also been performed. Based on longitudinal data of untreated class III subjects, who were classified as either good or bad growers based on the change in the sagittal relationship, a classification tree (CT) had a significantly lower rate of misclassification (12.0%) than discriminant analysis (40.7%) which was based on 11 cephalometric variables. [57] When the system was tested on new data, it was able to successfully identify good and bad growth patterns 64.0% of the time.

Conclusions

Artificial intelligence and machine learning systems applied in orthodontics provide promising tools that can improve clinical practice. These clinical decision support systems can help orthodontists practice more efficiently, reduce variability and subjectivity. [58] The accuracy of most systems presently available is considered good to excellent with an accuracy ranging from approximately 64% to 97% in some studies. The accuracy of the lower end of this range should be expected to improve in the future as sample sizes increase and more information becomes available. Most of the systems were developed using restricted samples that reduce their generalizability. For example, patients were often excluded because they needed surgery, or had missing teeth, unusual extraction patterns or asymmetries. Future studies are needed to build predictive models that include different types of patients. Algorithms should also be expected to improve, making it possible to handle more complex data such as images. Systems based on images require more time, experience and training data than systems based on discrete or continuous data values. This is especially important in the era of digital dentistry, where all patients' records such as dental models, x-rays and facial photos are stored in computers in the form of digital images.

It is important to note that AI models can possess some limitations and drawbacks and their recommendations should be taken with careful considerations. These ML algorithms have some assumptions and limitations like any statistical model. If used incorrectly, they can give incorrect information. In addition, the quality of data is important. [59] Data with a lot of noise, missing information, and more variables than observations are considered of poor quality and can result in poor models. Moreover, a phenomenon called overfitting can happen when a model is trained so many times on a data that has few observations. [60] This can result in a model that perform poorly when introduced to new data. Keeping that in mind, orthodontists should not trust completely the output given by any AI model and should understand that these models are meant to help with the clinical judgment and not substitute the knowledge and expertise of humans.

References

1. Haugeland J. *Artificial Intelligence: The Very Idea*. 1985, Cambridge: MIT Press.

2. Morris CG. *Academic Press Dictionary of Science and Technology*. 1996: Academic Press.

3. Luger GF. *Artificial intelligence: structures and strategies for complex problem solving*. 2005, Harlow, England: Pearson Education.

4. Mueller JP, Massaron L. *Machine learning for dummies*. 2016: John Wiley & Sons.

5. McCulloch WS, Pitts W. *A logical calculus of the ideas immanent in nervous activity.* Bull Math Biol, 1943. **5**(4): p. 115-133.

6. Negnevitsky M. *Artificial intelligence: a guide to intelligent systems*. 2005: Pearson Education.

7. Samuel AL. *Some Studies in Machine Learning Using the Game of Checkers. II—Recent Progress.* IBM Journal of research and development, 2000. 44(1.2): p. 206-226.

8. Marsland S. *Machine learning: an algorithmic perspective*. Second ed. 2014: Cambridge, UK: Chapman and Hall/CRC.

9. Dhage SN, Raina CK. *A review on machine learning techniques.* International Journal on Recent and Innovation Trends in Computing and Communication, 2016. **4**(3): p. 395-399.

10. Russell SJ, Norvig P. *Artificial intelligence: a modern approach*. 2013: Malaysia; Pearson Education Limited.

11. Coppin B. *Artificial intelligence illuminated*. 2004: Jones and Bartlett Publishers.

12. Lachman SJ. *Learning is a process: Toward an improved definition of learning.* The Journal of psychology, 1997. 131(5): p. 477-480.

13. Shalev-Shwartz S, Ben-David S. *Understanding machine learning: From theory to algorithms*. 2014: Cambridge university press.

14. Aronson JE, Liang TP, MacCarthy RV. *Decision Support Systems and Intelligent Systems (7th Edition)*. 2004: Prentice-Hall, Inc.

15. Rajdeep B, Sugata S. *Rule based expert system for diagnosis of neuromuscular disorders.* Int J Advanced Networking Appl, 2012. 4(1).

16. Stheeman SE, van der Stelt PF, Mileman PA. *Expert systems in dentistry. Past performance—future prospects.* J Dent, 1992. 20(2): p. 68-73.

17. Grant PD, Freer TJ. *Expert systems in orthodontic clinical diagnosis.* Aust Orthod J 1993. 13(1): p. 32-8.

18. Sumathi S, Sivanandam SN. *Introduction to data mining principles.* Studies in Computational Intelligence, 2013. 29(3): p. 1-20.

19. Mitchell TM. *Machine learning and data mining.* Communications of the ACM, 1999. 42(11): p. 30–36.

20. Breiman L. *Statistical modeling: The two cultures (with comments and a rejoinder by the author).* Statist. Sci, 2001. 16(3): p. 199-231.

21. Raschka S. *Python machine learning*. 2015: Packt Publishing Ltd.

22. Brown ID, Adams SR, Stephens CD, Erritt SJ, Sims-Williams JH. *The initial use of a computer-controlled expert system in the treatment planning of class II division 1 malocclusion.* Br J Orthod, 1991. 18(1): p. 1-7.

23. Williams JS, Matthewman A, Brown D. *An orthodontic expert system.* Fuzzy Set Syst, 1989. 30(2): p. 121-133.

24. Sims-Williams JH, Brown ID, Matthewman A, Stephens CD. *A computer-controlled expert system for orthodontic advice.* Br Dent J, 1987. 163(5): p. 161-166.

25. Hammond RM, Freer TJ. *Application of a case-based expert system to orthodontic diagnosis and treatment planning: a review of the literature.* Aust Orthod J, 1996. 14(3): p. 150-3.

26. Thanathornwong B. *Bayesian-Based Decision Support System for Assessing the Needs for Orthodontic Treatment.* Healthc Inform Res, 2018. 24(1): p. 22-28.

27. Ribarevski R, Vig P, Vig KD, Weyant R, O'Brien K. *Consistency of orthodontic extraction decisions.* Eur J Orthod, 1996. 18(1): p. 77-80.

28. Xie X, Wang L, Wang A. *Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment.* Angle Orthod, 2010. 80(2): p. 262-266.

29. Jung SK, Kim TW. *New approach for the diagnosis of extractions with neural network machine learning.* Am J Orthod Dentofacial Orthop, 2016. 149(1): p. 127-133.

30. Li P, Kong D, Tang T, Su D, Yang P, Wang H, Zhao Z, Liu Y. *Orthodontic treatment planning based on artificial neural networks.* Sci Rep, 2019. 9(1): p. 2037.

31. Leonardi R, Giordano D, Maiorana F, Spampinato C. *Automatic cephalometric analysis: a systematic review.* Angle Orthod, 2008. 78(1): p. 145-151.

32. Arık SÖ, Ibragimov B, Xing L. *Fully automated quantitative cephalometry using convolutional neural networks.* J Med Imaging, 2017. 4(1): p. 014501.

33. Buschang PH, Roldan SI, Tadlock LP. *Guidelines for assessing the growth and development of orthodontic patients*. Semin Orthod, 2017. 23(4): p. 321-335.

34. Giordano D, Kavasidis I, Spampinato C. *Modeling skeletal bone development with hidden Markov models.* Comput Methods Programs Biomed, 2016. 124: p. 138-147.

35. Kashif M, Deserno TM, Haak D, Jonas S. *Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment.* Comput Biol Med, 2016. 68: p. 67-75.

36. Mansourvar M, Shamshirband S, Raj RG, Gunalan R, Mazinani I. *An automated system for skeletal maturity assessment by extreme learning machines.* PLoS One, 2015. 10(9): p. e0138493.

37. Poedjiastoeti W, Suebnukarn S. *Application of Convolutional Neural Network in the Diagnosis of Jaw Tumors.* Healthc Inform Res, 2018. 24(3): p. 236-241.

38. Yong TH, Lee SJ, Yi WJ. *Odontogenic cysts and tumors detection in panoramic radiographs using Deep Convolutional Neural Network (DCNN)*. In: Proceedings of the International Conference on Medical Imaging with Deep Learning. 2019. London, United Kingdom.

39. Yilmaz E, Kayikcioglu T, Kayipmaz S. *Computer-aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography.* Comput Methods Programs Biomed, 2017. 146: p. 91-100.

40. Velemínská J, Pilný A, Cepek M, Kot'ová M, Kubelková R. *Dental age estimation and different predictive ability of various tooth types in the Czech population: data mining methods.* Anthropol Anz, 2013. 70(3): p. 331-345.

41. Laurenziello M, Montaruli G, Gallo C, Tepedino M, Guida L, Perillo L, Troiano G, Lo Muzio L, Ciavarella D. *Determinants of maxillary canine impaction: Retrospective clinical and radiographic study.* J Clin Exp Dent, 2017. 9(11): p. e1304-e1309.

42. Akçam MO, Takada K. *Fuzzy modelling for selecting headgear types.* Eur J Orthod, 2002. 24(1): p. 99-106.

43. Lim D, Ong Y, Setiawan R, Idris M. *Classifier-assisted constrained evolutionary optimization for automated geometry selection of orthodontic retraction spring*. IEEE Congress on Evolutionary Computation, Barcelona, Spain, 2010, p. 1-8.

44. Omran LN, Ezzat KA, Hassanien AE. *Decision Support System for Determination of Forces Applied in Orthodontic Based on Fuzzy Logic*. The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2018). 2018, p. 158-168.

45. Moghimi S, Talebi M, Parisay I. *Design and implementation of a hybrid genetic algorithm and artificial neural network system for predicting the sizes of unerupted canines and premolars.* Eur J Orthod, 2011. 34(4): p. 480-486.

46. Nanda SB, Kalha AS, Jena AK, Bhatia V, Mishra S. *Artificial neural network (ANN) modeling and analysis for the prediction of change in the lip curvature following extraction and non-extraction orthodontic treatment.* Journal of Dental Specialities, 2015. 3(2): p. 217-219.

47. Patcas R, Bernini DAJ, Volokitin A, Agustsson E, Rothe R, Timofte R. *Applying artificial intelligence to assess the impact of orthognathic treatment on facial attractiveness and estimated age.* Int J Oral Maxillofac Surg, 2018. 8(1): p. 77-83

48. Zarei A, El-Sharkawi M, Hairfield M, King G. *An intelligent system for prediction of orthodontic treatment outcome*. The 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 2006, p. 2702-2706

49. Auconi P, Scazzocchio M, Cozza P, McNamara JA Jr, Franchi L. *Prediction of Class III treatment outcomes through orthodontic data mining.* Eur J Orthod, 2015. 37(3): p. 257-67.

50. Kim BM, Kang BY, Kim HG, Baek SH. *Prognosis prediction for class III malocclusion treatment by feature wrapping method.* Angle Orthod, 2009. 79(4): p. 683-691.

51. Brodie AG. *Behavior of normal and abnormal facial growth patterns.* Am J Orthod Oral Surg, 1941. 27(11): p. 633-647.

52. Broadbent** BH. *The face of the normal child.* Angle Orthod, 1937. 7(4): p. 183-208.

53. Ricketts RM. *A principle of arcial growth of the mandible.* Angle Orthod, 1972. 42(4): p. 368-386.

54. Lux CJ, Stellzig A, Volz D, Jäger W, Richardson A, Komposch G. *A neural network approach to the analysis and classification of human craniofacial growth.* Growth Dev Aging, 1998. 62(3): p. 95-106.

55. Lakkshmanan A, Shri AA, Aruna E. *Pattern classification for finding facial growth abnormalities*. 2013 IEEE International Conference on Computational Intelligence and Computing Research, Enathi, India, 2013, p. 1-5

56. Niño-Sandoval TC, Guevara Perez SV, González FA, Jaque RA, Infante-Contreras C. *An automatic method for skeletal patterns classification using*

craniomaxillary variables on a Colombian population. Forensic Sci Int, 2016. 261: p. 159.e1-6.

57. Auconi P, Scazzocchio M, Caldarelli G, Nieri M, McNamara JA, Franchi L. *Understanding interactions among cephalometrics variables during growth in untreated Class III subjects.* Eur J Orthod, 2017. 39(4): p. 395-401.

58. Khanna S. *Artificial intelligence: contemporary applications and future compass.* Int Dent J, 2010. 60(4): p. 269-272.

59. Botsis T, Hartvigsen G, Chen F, Weng C. *Secondary use of EHR: data quality issues and informatics opportunities.* Summit Transl Bioinform, 2010. Mar 1;2010:1-5.

60. Rodvold DM, McLeod DG, Brandt JM, Snow PB, Murphy GP. *Introduction to artificial neural networks for physicians: taking the lid off the black box.* Prostate, 2001. 46(1): p. 39-44.

61. Decaestecker C, Remmelink M, Salmon I, Camby I, Goldschmidt D, Petein M, Van Ham P, Pasteels JL, Kiss R. *Methodological aspects of using decision trees to characterise leiomyomatous tumors*. Cytometry: The Journal of the International Society for Analytical Cytology, 1996. 24(1): p. 83-92.

62. Elbestawi M, Wu S, Du R. *Automated Monitoring of Manufacturing Processes, Part 1: Monitoring Methods*. J Eng Ind. 1995;117(2): p. 121-32.

63. Rokach L, Maimon O. *Data mining and knowledge discovery handbook*. Springer, Boston, MA, 2005. p. 165-92.

64. Song Y-Y, Ying L. *Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry*. Shanghai Arch Psychiatry, 2015. 27(2): p. 130-135.

65. Quinlan JR. *Induction of decision trees*. Machine learning, 1986. 1(1): p. 81-106.

66. Quinlan JR. *C4. 5: programs for machine learning*. Elsevier; 1993.

67. Pattekari SA, Parveen A. *Prediction system for heart disease using Naïve Bayes*. Int J Adv Comput Math Sci, 2012. 3(3): p. 290-4.

68. Subbalakshmi G, Ramesh K, Rao MC. *Decision support in heart disease prediction system using naive bayes.* Ind J Compu Sci Eng, 2011. 2(2): p. 170-176.

69. Chattopadhyay S, Davis RM, Menezes DD, Singh G, Acharya RU, Tamura T. *Application of Bayesian classifier for the diagnosis of dental pain*. J Med Syst, 2012. 36(3): p. 1425-1439.

70. Suebnukarn S, Rungcharoenporn N, Sangsuratham S. *A Bayesian decision support model for assessment of endodontic treatment outcome*. Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2008. 106(3): p. e48-58.

71. Langley P, Iba W, Thompson K. *An analysis of Bayesian classifiers*. In: Proceedings of the 10th National Conference on Artificial Intelligence, 1992, pp. 223-228

72. Bhargavi P, Jyothi S. *Applying naive bayes data mining technique for classification of agricultural land soils.* Int J Comput Sci Netw Sec, 2009. 9(8): p. 117-122.

73. Boyles S, Fajardo D, Waller ST, editors. *A naive Bayesian classifier for incident duration prediction.* In: the 86th Annual Meeting of the Transportation Research Board. 2007; pp. 1-10

74. Tu JV. *Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes*. J Clin Epidemiol, 1996. 49(11): p. 1225-1231.

75. Rumelhart DE, Hinton GE, Williams RJ. *Learning representations by back-propagating errors.* Nature*,* 1986. 323(6088): p. 533-536.

76. Cruz JA, Wishart DS. *Applications of Machine Learning in Cancer Prediction and Prognosis.* Cancer Infor, 2007. 2: p. 59-77.

77. Zadeh LA. *Fuzzy logic*. Computer, 1988. 21(4): p. 83-93.

78. Torres A, Nieto JJ. *Fuzzy logic in medicine and bioinformatics*. J Biomed Biotechnol, 2006. 2006(2): p. 91908.

79. Warren J, Beliakov G, Van Der Zwaag B. *Fuzzy logic in clinical practice decision support systems.* In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, 2000: p. 1-10 .

80. Zadeh LA. *Outline of a new approach to the analysis of complex systems and decision processes.* IEEE Trans Syst Man Cybern Syst, 1973. 1: p. 28-44.

81. Godil SS, Shamim MS, Enam SA, Qidwai U. *Fuzzy logic: A "simple" solution for complexities in neurosciences?* Surg Neurol Int, 2011. 2:24

CHAPTER III

CAN DECISION TREES ACCURATELY PREDICT ANTERIOPOSTERIOR SKELETAL RELATIONS?

Introduction

To correct skeletal and dental problems, orthodontists need to be able to predict the growth potential of their patients. Most orthodontic patients are growing adolescents, comprising approximately 79% of case starts. [1] It is especially important for the orthodontists to be able to predict the patients' maxillomandibular anteroposterior (AP) relationships, which can change in either favorable or unfavorable ways. Approximately 50% of Class I and Class II patients have favorable growth patterns that can help reduce and thereby improve AP skeletal and dental relations. The other 50% have unfavorable growth patterns that often worsen these relationships. [2] Favorable growth explains why end-on or slight mesial step dental relationships in the primary dentition become Class I relations in the permanent dentition. [3] When treating Class II patients with favorable patterns, the majority (70%) of the molar corrections are due to growth rather than treatment. [4] Growth does not much, if any, of the molar corrections in patients with unfavorable patterns. On that basis, the ability to predict a patient's growth potential makes it possible to modify treatment plans and thereby improve treatment outcomes.

The orthodontic literature includes various studies that have attempted to predict facial growth, but most are not accurate [5-9] or complex. [10] Among the first methods used was the pattern extension, which assumed that each patient's growth pattern is established early in life and does not change. [8, 9] Based on that assumption, pattern extension simply added the same average growth increments to all patients, often times with adjustments for age and sex. [5-7] This notion was challenged by Björk and coworkers, who found that younger subjects with good AP relationships could develop sagittal problems over time, just as the relations of those with horizontal discrepancies between the maxilla and the mandible could improve. [11] This has led to the development of growth predictions methods using multi-level models, which developed individualized growth curves and were able to work well with missing data. [10, 12] However, these methods are complex and require longitudinal cephalograms to make accurate predictions.

With the emergence of AI and ML, it has been possible to overcome many of the limitations associated with traditional statistical methods and produce simpler, more accurate predictions. These new methods can be used when traditional methods cannot. For example, they can use predictor variables that are correlated, which traditional statistical models such as logistic regression [13] and multiple regression [14] cannot. They also do not have to make the assumptions that traditional statistics must make. In addition, noise in the data caused by errors and missing values, which can affect the accuracy of traditional predictive models, have less effects on ML models. When applied to the same data, ML models are often more accurate than traditional methods. For example, the medical literature has shown that ML outperforms logistic regression by approximately 7% [15] and discriminant analysis by 10%. [16] In orthodontics, ML correctly predicted the prognosis of class III treatment 97.2% of the time, whereas discriminant analysis was correct 92.1% of the time. [17] One of the best methods used in ML for classification are decision trees (DTs).

Decision trees are among the most popular ML models used for classification. A major reason for their popularity is their ease of interpretation, even for non-experts. [18, 19] They can handle a mixture of categorical and continuous variables, even when there are missing values. They can also classify multiple groups. [19] A tree is built by asking a series of a yes-or-no questions. Each question has a node that leads to two other nodes, depending on the answer. The process continues until the terminal node is reached, which is the final outcome. [19] The first uppermost node pertains to the most important variable, followed by the next most important variable, and so on, until a hierarchy of variables is created. [20] For researchers, building decision trees is less time consuming than other classification algorithms. [21] Due to these reasons, DTs have often been used in medicine and bioinformatics. DTs applied in orthodontics to classify untreated Cass III subjects as good or bad growers have shown a significantly lower rate of misclassification (12.0%) than discriminant analysis (40.7%). [22]

The purpose of this study was to determine the applicability of DTs for the classification of growth patterns in a sample of 222 untreated Cass I and Class II subjects who were followed longitudinally. The objective was to develop accurate predictions and identify the most important variables contributing to the predictions.

Materials and Methods

The sample includes 222 untreated subjects (116 males, 106 females) with Class I or Class II dental occlusion/malocclusion. They were all French-Canadians, drawn from three school districts representing the socioeconomic backgrounds of the Montreal area at large. [23]

Subjects were selected based on the availability of two longitudinal cephalograms. The mean ages at T1 and T2 were 10.4 ± 1 years and 15.3 ± 0.6 years, respectively. One technician traced and digitized all of the cephalograms. On each tracing, twelve landmarks were identified (Table 2). To describe the horizontal and vertical positions of landmarks, rectangular coordinates (X, Y) were computed registering on sella and orienting 7 degrees below S-N (Figure 2). Size measurements were corrected for radiographic enlargement. Reliability of landmark locations ranged between 95% and 98%. [2] The T2 tracing was superimposed onto the T1 tracing using stable natural structures in the anterior cranial base and cranium, [24] with a reliability greater than 98%. [25] After superimpositioning, the natural reference line (RL) was on the T1 was transferred to the T2 tracing and used for orientation.

The subjects' skeletal relationships were based on the horizontal distance between ANS (the maxillary skeletal base) and Pg (the mandibular base). These landmarks were selected because they can be reliably located and are less affected by tooth movements. Importantly, these landmarks have been previously validated for evaluating maxillomandibular relationships. [2] The horizontal distance between ANS and Pg (ANSPgh) was the primary outcome variable of the present study. Each subject's AP relationship at 15 years of age was classified as either favorable and unfavorable based on whether $ANSPgh_{15}$ was below (the favorable group) or above (the unfavorable group) his/her sex specific mean value, respectively.

Eleven predictor variables, previously used to distinguish between different adolescent facial growth patterns, [26-28] were computed at age 10 and used to predict favorable and unfavorable relations at age 15 (Table 2). Some of the predictors described horizontal and vertical facial patterns, while others pertained to mandibular features, such as the shape of the condylar inclination and the symphyseal shape. Due to an expected relationship, $ANSPgh_{10}$ was also included as a potential predictor variable (Table 3).

Decision trees (DTs) are common machine learning algorithms used in medicine [29] and dentistry, [30] including orthodontics. [22] As previously indicated, their branching structure makes them easier to interpret and understand, even to non-specialists. In addition, they provide probabilities for each decision at each node. The probability of favorable outcome is reported on the left side of each node, while the probability of unfavorable outcome is indicated on the right side of each node. The percentage of subjects meeting these conditions is presented below each of the nodes. Pruning was undertaken due to the possibility of overfitting the data of DTs. [31] Pruning identifies the less relevant variables and reduces the number of branches. The accuracy of the larger trees can be improved by pruning. [32]

To ensure that the models apply to other subjects that were not included in the training sample, 20% (44 subjects) of the total sample was randomly chosen and used as the validation sample. The remaining 80% (178 subjects) was used for training the model. The trees were produced using the R package ‘rpart’ version 4.1-15. [33]

Results

The DTs for the 178 subjects in the training sample were built using 4 of the 12 variables, with the most important variable on top (Figure 3). All of them were angular measurements (Y-axis, ANS-N-PG, NSB, and MPA), producing 7 terminal nodes. The unpruned tree showed that 51.1% of the subjects had a Y-axes smaller than 68.4 degrees, while 48.9% had Y-axes greater than or equal to 68.4 degrees.

If a subject had a Y-axis less than 68.4°, then a favorable growth outcome could be expected 73.6% of the time. If the Y-axis is less than 68.4° and ANS-N-PG was less than 11.4°, then the likelihood of a favorable outcome increased to 80.5%. It further increased to 90% if the subject on had a MPA less than 36.2°. If a subject has a Y-axis greater than or equal to 68.4°, then an unfavorable growth outcome was expected 48.9% of the time. The probability increased to 81.8% if a subject has also an ANS-N-PG greater than or equal to 7.64°. If the NSB angle was greater than or equal to 128°, the probability of an unfavorable relationship increased to 89.4%. The accuracy of the model was 85.39% on the training data set and 75% on the validation data set (Table 4; Figure 4).

The pruned tree included 3 variables and 5 terminal nodes (Figure 5). If a subject had a Y-axis greater than or equal to 68.4°, the branching, the contributing variables and the probabilities were the same as in the unpruned tree. If a subject had a Y-axis less than 68.4° , the probability of a favorable relationship at 15 years of age was also 73.6%, as with the unpruned tree. The probability of a favorable relationship increased to 80.5% if the ANS-N-PG angle was less than 11.4°. The accuracy of the pruned model was

83.15% on the training data set and 81.82% on the validation data set (Table 4; Figure 6).

Discussion

The horizontal distance between the maxilla and the mandible provides an excellent measure of AP basal bone relationships. This measure was used to eliminate the problems associated with other measures of maxilomandibular relationship, such as the ANB angle and wits appraisal. The ANB angle is based on three landmarks that can affect its accuracy; it is also affected by incisor position. [34] The Wits appraisal is also affected by tooth position, both horizontally and vertically. [35] Pg and ANS represent the anteriormost postions of the jaws. Moreover, the distance ANSPgh has been shown to provide a valid measure for determining the jaw responsible for longitudinal changes of AP skeletal relationships. [2] As previously indicated, these skeletal landmarks are easily located on lateral cephalograms, they are more stable than dental landmarks and they are no influenced by in incisor position. [36]

Favorable and unfavorable growth were easily distinguished using the 15-year mean values as cut-offs. Values at 15 years of age were used because they pertain to AP relationships after most maxillomandibular growth changes have taken place. The use of means made it possible to dichotomize subjects, as previously done to the classify patients needing and not needing treatment. [37, 38] This simplifies the comparison and makes it possible to use decision trees, which perform better with binary outcomes. [39] At 15 years of age, ANSPgh was approximately 7-8 mm smaller among subjects with favorably than unfavorably AP relationships, a difference that is both statistically significant and clinically important. Most orthodontists would agree that a 17-18 mm distance between ANS and Pg constitutes an unfavorable growth pattern.

Artificial intelligence (AI) and machine learning (ML) have been recently used to classify growth and has shown excellent results. [40-42] The increasing use of these methods can be attributed to the previously described advantages they have over traditional statistics. The present study used DTs to classify subjects because 1) their flow chart structures are easily understood and 2) they hold clear advantages over other models. [43] The DTs used in the present study had lower rates of misclassification (12.1-24.2%) than previously demonstrated using discriminant analysis (40.7%). [22]

Pruning helps to enhance the validity of the DT. There is a risk of overfitting (not being able to generalize beyond the training sample) associated with larger unpruned trees that include more variables and have more complex branches. [31] In other words, overfitting makes it less likely that the model can be applied to other subjects. This risk is reduced, and the classification accuracy of the model can be improved by pruning. [32] This was clearly seen in the present study, where the misclassification rate in the unpruned tree increased from 14.61% in the training sample to 25% in the validation sample (Figure 4). In contrast, the pruned tree showed only a slight 1.3% increase in misclassification between the training and validation samples (Figure 6). The pruned tree is also easier to understand and apply clinically.

The Y-axis is the most important variable for predicting AP skeletal relations. In the present study, the Y-axis alone was able to distinguish between favorable and unfavorable relationships almost 75% of the time. It was most important because it incorporates both horizontal and vertical aspects of growth. [2] Ricketts originally identified the Y-axis as an important variable to describe the mandible's future growth direction. [9] He found that the Y-axis of prognathic and retrognathic patients decreased and increased, respectively, during growth. [44] A recent regression model also reported the Y-axis as the most important variable explaining 41% of the variation of the horizontal relationship of ANSPgh at T2. [45]

Including both ANS-N-Pg and the Y-axis increased the probability of identifying favorable and unfavorable outcomes from 73.6% to 80.5%. Multiple regression has shown that the Y-axis explained 41% of the variation of the horizontal relationship of ANSPgh at T2, and ANS-N-Pg explained an additional 16%. [45] If a subject has a small (< 68.4) Y-axis and a large (≥ 11.4) ANS-N-Pg angle, the probability of an unfavorable AP relationship at 15 years of age was approximately 89% (Figure 3). If a subject had a large (≥ 68.40) Y-axis but a small (< 7.64) ANS-N-Pg angle, then the probability of favorable outcome was 80%. This shows that there are subjects with favorable AP skeletal relations whose mandibular growth is oriented in a more vertical direction.

The cranial base angle further improves the accuracy of identifying subjects with unfavorable maxillomandibular relationships. The probability of unfavorable relationships increases from approximately 82% to slightly more than 89% among subjects with large cranial base angles (i.e., greater than 128 degree). Larger cranial base angles have been previously related to retrognathisim. [46-48] Cranial base angles are larger among Class II division 1 than Class I and Class III subjects, respectively. [49]

Conclusions

- The unpruned and pruned decision trees were able to successfully classify the growth of untreated subjects 85.4% and 83.2% of the time, respectively.
- The validation sample showed that the pruned decision tree was more accurate than the unpruned (81.8% vs 75.0%) tree.
- The decision trees identified the Y-axis as the most important variable to classify growth, followed, in order, by ANS-N-PG, NSB, and MPA.

References

1. Keim RG, Gottlieb EL, Vogels DS 3rd, Vogels PB.*2017 JCO orthodontic practice study: Part 1 Trends.* J Clin Orthod, 2017. 51(10): p. 639-656.

2. Buschang PH, Martins J. *Childhood and adolescent changes of skeletal relationships.* Angle Orthod, 1998. 68(3): p. 199-208.

3. Bishara SE, Hoppens BJ, Jakobsen JR, Kohout FJ. *Changes in the molar relationship between the deciduous and permanent dentitions: a longitudinal study.* Am J Orthod Dentofacial Orthop. 1988. 93(1): p. 19-28.

4. Harris EF, Dyer GS, Vaden JL. *Age effects on orthodontic treatment: skeletodental assessments from the Johnston analysis.* Am J Orthod Dentofacial Orthop., 1991. 100(6): p. 531-536.

5. Broadbent BH, Broadbent BH, Golden WH. *Bolton standards of dentofacial developmental growth*. 1975: CV Mosby.

6. Broadbent** BH. *The face of the normal child.* Angle Orthod, 1937. 7(4): p. 183-208.

7. Brodie AG. *On the growth pattern of the human head. From the third month to the eighth year of life.* American Journal of Anatomy, 1941. 68(2): p. 209-262.

8. Johnston LE. *A simplified approach to prediction.* Am J Orthod, 1975. 67(3): p. 253-257.

9. Ricketts RM. *Planning treatment on the basis of the facial pattern and an estimate of its growth.* Angle Orthod, 1957. 27(1): p. 14-37.

10. Chvatal BA, Behrents RG, Ceen RF, Buschang PH. *Development and testing of multilevel models for longitudinal craniofacial growth prediction.* Am J Orthod Dentofacial Orthop, 2005. 128(1): p. 45-56.

11. Björk A, Pallilng M. *Adolescent age changes in sagittal jaw relation, alveolar prognathy, and incisal inclination.* Acta Odontol Scand, 1955. 12(3-4): p. 201-232.

12. Roberts RO. *Adolescent Maxillomandibular Relationships: Growth Patterns, Inter-Individual Variability, and Predictions*. The Texas A&M University System Health Science Center; 2007.

13. Stoltzfus JC. *Logistic regression: a brief primer.* Acad Emerg Med, 2011. 18(10): p. 1099-1104.

14. Osborne JW, Waters E. *Four assumptions of multiple regression that researchers should always test.* Pract. Assess. Res. Evaluation, 2002. 8(2): p. 1-9.

15. Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. *Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards.* Crit Care Med, 2016. 44(2): p. 368-74.

16. Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I, de Mendonça A. *Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests.* BMC Res Notes, 2011. 4(1): p. 1-14.

17. Kim BM, Kang BY, Kim HG, Baek SH. *Prognosis prediction for class III malocclusion treatment by feature wrapping method.* Angle Orthod, 2009. 79(4): p. 683-691.

18. Kotsiantis SB. *Decision trees: a recent overview.* Artif Intell Rev, 2013. 39(4): p. 261-283.

19. Kingsford C, Salzberg SL. *What are decision trees?* Nat Biotechnol, 2008. 26(9): p. 1011.

20. Song YY, Lu Y. *Decision tree methods: applications for classification and prediction.* Shanghai Arch Psychiatry, 2015. 27(2): p. 130-5.

21. Shafer J, Agrawal R, Mehta M. *SPRINT: A scalable parallel classifier for data mining*. In Proc. of the 22nd Int'l Conference on Very Large Databases, Bombay, India, September 1996.

22. Auconi P, Scazzocchio M, Caldarelli G, Nieri M, McNamara JA, Franchi L. *Understanding interactions among cephalometrics variables during growth in untreated Class III subjects.* Eur J Orthod, 2017. 39(4): p. 395-401.

23. Demirjian A, Dubuc MB, Jenicek M. *Etude comparative de la croissance de l'enfant canadien d'origine française à Montréal.* Can J Public Health, 1971. 62(2): p. 111-119.

24. Björk A, Skieller V. *Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years.* Eur J Orthod, 1983. 5(1): p. 1-46.

25. Buschang PH, LaPalme L, Tanguay R, Demirjian A. *The technical reliability of superimposition on cranial base and mandibular structures. .* Eur J Orthod, 1986. 8(3): p. 152-156.

26. Björk A. *Prediction of mandibular growth rotation.* Am J Orthod, 1969. 55(6): p. 585-599.

27. Skieller V, Björk A, Linde-Hansen T. *Prediction of mandibular growth rotation evaluated from a longitudinal implant sample.* Am J Orthod, 1984. 86(5): p. 359-370.

28. Ricketts RM. *The influence of orthodontic treatment on facial growth and development.* Angle Orthod, 1960. 30(3): p. 103-133.

29. Lee VJ, Lye DC, Sun Y, Leo YS. *Decision tree algorithm in deciding hospitalization for adult patients with dengue haemorrhagic fever in Singapore.* Trop Med Int Health, 2009. 14(9): p. 1154-1159.

30. Chiang HJ, Tseng CC, Torng CC. *A retrospective analysis of prognostic indicators in dental implant therapy using the C5. 0 decision tree algorithm.* J. Dent. Sci, 2013. 8(3): p. 248-255.

31. Bellazzi R, Zupan B. *Predictive data mining in clinical medicine: current issues and guidelines.* Int J Med Inform, 2008. 77(2): p. 81-97.

32. Strobl C, Malley J, Tutz G. *An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests.* Psychol Methods, 2009. 14(4): p. 323-48.

33. Therneau T, Atkinson B, Ripley B, Ripley MB. *Package 'rpart'.* Available online: cran.r-project.org/web/packages/rpart/rpart.pdf (accessed on 18 March 2021). 2015 Jun 29.

34. Jacobson A. *The "Wits" appraisal of jaw disharmony.* Am J Orthod, 1975. 67(2): p. 125-138.

35. Nanda RS, Merrill RM. *Cephalometric assessment of sagittal relationship between maxilla and mandible.* Am J Orthod Dentofacial Orthop, 1994. 105(4): p. 328-344.

36. Ricketts RM. *Cephalometric synthesis: An exercise in stating objectives and planning treatment with tracings of the head roentgenogram.* Am J Orthod, 1960. 46(9): p. 647-673.

37. Costa RN, Abreu MH, Magalhães CS, Moreira AN. *Validity of two occlusal indices for determining orthodontic treatment needs of patients treated in a public university in Belo Horizonte, Minas Gerais State, Brazil.* Cad Saude Publica, 2011. 27(3): p. 581-590.

38. Beglin FM, Firestone AR, Vig KW, Beck FM, Kuthy RA, Wade D. *A comparison of the reliability and validity of 3 occlusal indexes of orthodontic treatment need.* Am J Orthod Dentofacial Orthop, 2001. 120(3): p. 240-246.

39. Nourzad SH, Pradhan A. *Binary and multi-class classification of fused LIDAR-imagery data using an ensemble method*. In Construction Research Congress 2012: Construction Challenges in a Flat World 2012 (pp. 909-918).

40. Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li SX, Negahban SN, Krumholz HM. *Analysis of machine learning techniques for heart failure readmissions.* Circ Cardiovasc Qual Outcomes, 2016. 9(6): p. 629-640.

41. Shameer K, Johnson KW, Yahi A, Miotto R, Li LI, Ricks D, Jebakaran J, Kovatch P, Sengupta PP, Gelijns S, Moskovitz A, Darrow B, David DL, Kasarskis A, Tatonetti NP, Pinney S, Dudley JT. *Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: a case-study using Mount Sinai heart failure cohort*. Pac Symp Biocomput, 2017. 22: p. 276-287.

42. Asiri SN, Tadlock LP, Schneiderman E, Buschang PH. *Applications of artificial intelligence and machine learning in orthodontics.* APOS Trends in Orthodontics, 2020. 10(1): p. 17-24.

43. Agrawal GL, Gupta H. *Optimization of C4. 5 decision tree algorithm for data mining application.* Int. J. Emerging Technol. Adv. Eng, 2013. 3(3): p. 341-345.

44. Ricketts RM. *Cephalometric analysis and synthesis.* Angle Orthod, 1961. 31(3): p. 141-156.

45. Bleyer JS. *A Predictive Model of" Favorable" Versus" Unfavorable" Growth For Orthodontic Treatment Planning*. Master Thesis, 2019.

46. Björk, A. *The nature of facial prognathism and its relation to normal occlusion of the teeth.* Am J Orthod., 1951. 37(2): p. 106-124.

47. Järvinen S. *Saddle angle and maxillary prognathism: a radiological analysis of the association between the NSAr and SNA angles.* Br J Orthod, 1984. 11(4): p. 209-213.

48. Sassouni V. *A classification of skeletal facial types.* Am J Orthod, 1969. 55(2): p. 109-123.

49. Hopkin GB, Houston WJ, James GA. *The cranial base as an aetiological factor in malocclusion.* Angle Orthod, 1968. 38(3): p. 250-255.

CHAPTER IV

PREDICTION OF DENTAL TREATMENT OUTCOMES AMONG HERBST PATIENTS USING DECISION TREES

Introduction

One of the most common problems in the US population is class II dental malocclusion, which occurs in approximately 15% of the population. [1] Among those, approximately 75% have also class II skeletal malocclusion. [2] Class II subjects usually present with mandibular retrusion and facial convexity. [3] In order to improve esthetics, the position of the chin need to be brought forward to create less convex or flatter profile. [4] The class II skeletal malocclusion is generally not self-correcting. [5, 6] This is also true for full step class II dental relationships. Subjects with full step relationships in the primary dentition always have Class II dental relationship in the permanent dentition. [7] Longitudinal studies also indicates that approximately one third of children with end-to-end or slight mesial step deciduous molar relationships develop Class II molar relationships in the permanent dentition. [7, 8] The various treatment modalities performed to correct this malocclusion should be based on evidence-based information of what will and will not correct problems.

Excluding surgery, there are two common treatment modalities to correct skeletal class II problems, headgears and functional appliances, either fixed or removable. Fixed functional appliances eliminate the compliance problem associated with the removable functional appliances. [9] Herbst is one of the more popular fixed functional appliances used to correct class II malocclusion. In 1905, Emil Herbst introduced his appliance at the international congress of dentistry in Berlin, Germany, but it was not accepted until 1979, when Pancherz reintroduced it. [10] Since then, numerous papers have evaluated the biological effects of the Herbst appliance [11-13] and quantified the skeletal and dental changes associated with the correction of class II malocclusion. [14-17] A recent meta-analysis of the Herbst treatment reported 0.56 ◦ and 1.10 ◦ decreases of the SNA and ANB angles, respectively, and 1.1 ◦ and 0.2 ◦ increases of the SNB and mandibular plane angles, respectively. [18] Overjet was found to decrease by 4.8 mm, while overbite and molar relationships decreased by 1.7 mm and 5.7 mm, respectively.

Orthodontic treatment is considered successful when both the treatment objectives and patient desires have been achieved. [19] In treatment with functional appliances, the definition of successful treatment outcomes differed among studies. Most studies have used occlusal measurements, mainly the correction of overjet and molar relations, to distinguish between successful and unsuccessful treatment. [20-24] Skeletal measures have also been used to determine successful treatment outcomes with functional appliances. [22, 25] Other studies used a composite treatment outcome that had to be met in order to determine successful outcomes. [24, 26, 27] Importantly, a treatment considered to be successful based on occlusal measurements might not be considered as a successful case if the skeletal outcome was evaluated only (Table 5). For example, patients successfully treated dentally often become more hyperdivergent during Herbst treatment, with little or no improvement in the anteroposterior skeletal dimension. On that basis, dental and skeletal outcomes should be evaluated separately.

Previous studies have demonstrated that slightly more than 50% of the patients treated with functional appliances have unfavorable treatment outcomes. [22, 24] This clearly indicates that models are needed to predict treatment outcomes of patients based on pre-treatment features. To date, traditional statistics have been used to predict treatment outcomes in patients treated with functional appliances without making a clear distinguish whether the outcome was entirely skeletal or dental. For example, multiple regression model showed that overbite and SNB explained 52% of the variation in overjet changes among patients treated with twin block appliance. [23] When discriminant analyses were also used to classify successful and unsuccessful outcomes among various functional appliances (twin block, stainless steel crown Herbst, and an acrylic splint Herbst), the gonial angle (Co-Go-Me◦) was found to be able to correctly predict outcomes 80% of the time. [25] In addition, the patients were classified into bad and good responders based on changes in the mandibular length which is not the best indicator of treatment success. A subject can show an increase in the total mandibular length, but the final outcome might still be considered as unsuccessful if the patient is hyperdivergent to begin with and the vertical dimensions increases. This could result in backward rotation, and therefore unsuccessful treatment outcomes. [28]

Recently, artificial intelligence (AI) was used to predict treatment outcomes among class II and class III patients, with better outcomes than could be attained with traditional statistics. Using artificial neural networks (ANN), post-treatment Peer Assessment Ratings (PAR) were predicted among class II patients based on their pre-treatment PAR index. [29] The ANN model used was able to accurately predict the final

PAR score 94.0% of the time; which was better than the 82.0% reported for linear regression. AI was also used to correctly predict the prognosis of class III treatment 97.2% of the time, which was slightly better than the 92.1% obtained with discriminant analysis. [30]

Decision trees (DTs), a commonly used AI method, have been used and have also shown excellent results at predicting treatment outcomes for dental implants [31] and classifying growth in untreated subjects. [32, 33] DTs offer several over other methods. They are non-parametric, which means they do not make assumptions about the variables' distributions and the relationship between the independents and dependent variables. [34] Moreover, decision trees are non-linear, which allows the same variable to be used multiple times. This makes it possible to detect relationships that cannot be detected with linear methods. DTs also can handle categorical and numerical variables, unlike other machine learning algorithms. [34] Importantly, DTs are relatively easy to interpret and understand, even by non-specialists. Their if-then approach makes predicting treatment outcomes easy and efficient when new observations are available to clinicians. [35] They need to get the new patient's pretreatment measurements, and then follow the path of the tree leading to the final outcome which takes less than a minute. In term of accuracy, DTs were recently shown to outperform discriminant analysis for predicting growth in untreated subjects. [32]

The purpose of the present study was to determine the applicability of DTs for predicting dental treatment outcomes among class II patients treated with Herbst previously. The objective was to develop accurate predictions and identify the most important variables contributing to the predictions.

Materials and Methods

The sample included 150 patients (98 males, 52 females) who had been previously treated (within the past 20 years) with the Herbst and fixed appliances. Their records were collected from two private practices, and from the Graduate Orthodontic Department of Texas A&M University. The study was approved by Texas A&M University IRB (approval # 2019-1238D). To be included in the study, each subject had to meet the following criteria:

1) at least half step class II molar and canine relationships (unilateral or bilateral) before treatment;
2) between the ages of 9 and 16 years;
3) treated previously with the Herbst (for at least 6 months) and fixed appliances;
4) good quality pretreatment (T1) and posttreatment (T2) radiographs taken with lips lightly touching
5) no syndromes, craniofacial anomalies, or congenitally missing teeth; and
6) no orthognathic or cosmetic facial surgery.

The mean ages of the subjects at T1 and T2 were 12.5 ±1.5 years and 15.3 ±1.5 years, respectively. All patients were treated using a standard cantilever Herbst appliances with full-coverage stainless steel crowns on the maxillary and mandibular first molars, telescopic cantilever arms from the mandibular first molars, and 0.040 mm stainless steel lower lingual arches with occlusal rests on the mandibular first premolars. During the

Herbst phase, fixed appliances were bonded to the maxillary second premolars 5-5 and the mandibular canines 3-3. After the Herbst removal, fixed appliances were bonded to the remaining teeth (U&L 6-6). The mean treatment time with the Herbst was 12.4 ±2.9 months and the mean total treatment duration was 32.2 ±7.2 months.

Evaluations

All pretreatment and posttreatment cephalograms were digitized by the primary investigator using Dolphin Imaging software, version 11.0 (Dolphin Imaging, Chatsworth, CA, USA). Nine angular measurements and 7 linear measurements were computed from the pretreatment cephalometric as predictor variables. The stage of cervical vertebral maturation (CVS) was determined from the pretreatment cephalograms, [36] along with 3 additional variables (Class II dental relationships [unilateral or bilateral], severity of Class II [full-step, half-step], and the presence/absence of posterior crossbite) evaluated on the pretreatment intraoral photographs (Table 6 & Figure 7). All linear measurements were adjusted to eliminate magnification. For reliability, the lateral cephalograms of 8 randomly selected patients were remeasured. The method error ranged between 0.14 and 0.32 for the linear measurements and between 0.33 and 0.60 for the angular measurements.

Subjects were classified as having favorable or unfavorable dental outcomes based on their post-treatment (T2) dental relationships. In order to be considered as a favorable dental outcome, three dental criteria had to be met: 1) bilateral Class I molar relationships (with the maxillary first molar mesiobuccal cusp tip aligning within 1 mm of the buccal grooves of the mandibular molars), 2) bilateral Class I canine relationships (with the maxillary canine cusp tip aligning within 1 mm of the embrasure of the mandibular canine), and 3) an overjet less than 4 mm (or less than 50% of the initial overjet if the pre-treatment overjet was greater than 8 mm). If one or more of these three criteria were not satisfied, the subject was classified as having an unfavorable dental outcome. Overjet was measured on the final cephalogram, and molar and canine relationships were determined from the final intraoral photos.

Statistical Analyses

Decision trees (DTs) are machine learning algorithms that have been shown to be successful in making predictions in dentistry [31] and orthodontics. [32] The tree-like structure makes it easier to visualize the model and better understand the interactions between the different variables. The probability of having a favorable dental outcome is depicted at each node of the tree, along with the percentage of subjects satisfying the previous condition. When using DTs, there is a process called pruning that eliminates the less important variables and reduces the size of the tree to make it less complex. Pruning can help improve the accuracy of the model, especially when the DT is complex and has many branches that make it prone to overfitting. [37]

The DT model was developed using 113 randomly selected subjects (75%). These subjects were used as the training data. The remaining 37 subjects (25%) was used to validate the model (testing data). The DTs were produced using the R package 'rpart' version 4.1-15. [38] SPSS version 25 (SPSS Inc, Chicago, Ill) was used for the traditional statistical analyses. Based on their skewness and kurtosis, the distributions were all normal. Independent sample T-tests were used to compare the pretreatment differences and the change between the favorable and unfavorable groups. Pearson chi-square test was used to compare the frequencies of favorable and unfavorable outcomes.

Results

The unfavorable dental outcomes were primarily due to the inability to achieve Class I canine relationships (94.6%), followed by the inability to achieve Class I molar relationships (70.3%), and lastly by excessive overjet (21.6%) (Table 7).

The frequencies of posterior crossbites and full step Class II molar relationships were significantly higher frequencies among patients with unfavorable than favorable outcomes (Table 8). Favorable treatment outcomes were more likely among patients treated at the university clinic (70.0%) than among those treated in the private orthodontic practices (Table 9), but the between-group differences were not statistically significant (prob=0.06).

Only the pretreatment between-group difference of SNB was statistically significant (Table 10). Differences of SNA and Wits were not statistically significant after Bonferroni adjustments for multiple comparisons. None of the treatment changes showed statistically significant between-group differences (Table 11).

Decision Trees

Based on the 137 subjects in the training sample, the DT identified 7 predictor variables as the most important (Figure 8). Four of them were angular measurements (SN-MP, L1-MP, U1-NA, U1-SN), two were linear measurements (Overbite, Pg-NB), and one was the dichotomous variable describing the severity of pretreatment dental Class II relationships (whether it was full step or not).

The SN-MP angle, which was the most important variable, called for the first decision. The 15% of the patients who had SN-MP angles greater than or equal to 30° could expect to have favorable and unfavorable outcomes 12% and 88% of the time, respectively. The remaining 85%, who had SN-MP angles less than 30°, could expect favorable outcome 56% of the time. None of the 6% of patients who had SN-MP angles < 30° and an overbite ≥ 7.3 mm were expected to have favorable outcomes. The likelihood of a favorable outcome increases to 61% if the SN-MP angle was < 30° and the overbite was < 7.3 mm. Of the 21% who also had L1/MP angle > 99, those that had Pg-NB distances ≥ 2.6 were all expected to have unfavorable dental outcomes. Those with Pg-NB distances < 2.6 mm had a 59% chance of having favorable outcomes.

If the subject had a SN-MP angle < 30°, an overbite < 7.3mm, a L1-MP < 99°, and a full step Class II molar relationship, a favorable outcome was expected 54% of the time. The probability of a favorable outcome increased to 83% if the U1-NA angle was < 20°. It decreased to 40% if the U1-NA angle was ≥ 20° and further decreased to 24% if the U1-SN angle was < 114°. However, if the U1-SN angle was ≥ 114°, the probability of a favorable outcome was 75%. The 25% of patients who had SN-MP angles < 30°, overbite < 7.3mm, L1-MP < 99°,and less than a full-step Class II had favorable outcomes 86% of the time. The accuracy of the model was 81.4% when applied to the training data set and 78.4% when applied to the testing data set (Table 12 & Figure 9).

The pruned tree included 3 variables and 4 terminal nodes (Figure 10). The branching of the tree, the contribution of the variables, and the probabilities were the same as previously described for the unpruned tree. The pruning process simply eliminated the least important variables. The accuracy of the pruned model was 70.8% on the training data set and 67.6% on the testing data set (Table 12 & Figure 11).

The cephalometric measurements of the subjects in each of the 8 terminal nodes of the unpruned tree are provided in Table 13. These measurements suggest that subjects with Class II division 1 malocclusion and end on molar relationships had the highest likelihood of favorable dental outcome (86%) followed by subjects with Class II division 2 malocclusion and full step molar relationships (83%). Features of Class II division I malocclusion were indicated in the 2 groups that had U1-SN $\geq 114^{\circ}$ and U1-SN $< 114^{\circ}$. However, their likelihoods of favorable outcomes differed.

Discussion

Approximately 50% of the Herbst cases should be expected to have unfavorable dental treatment outcomes. This is consistent with frequencies previously reported for other functional appliances; treatment outcomes of the Herbst have not previously been reported. For example, slightly more than half of the patients (15 out of 28 cases) treated with 3 different functional appliances had unsuccessful treatment outcomes, based on changes in the ANB angle. [22] Based on a composite of dental and skeletal outcome variables, slightly more than half of the patients treated with bionator appliances had unfavorable outcomes. [24] Given such high proportions of patients with unfavorable outcomes, it is imperative to identify pre-treatment features that could help identify the patients better suited for other treatments (e.g. surgery).

Unfavorable dental outcomes were most commonly due to the inability to establish or maintain Class I dental relationships. Unfavorable outcomes could also have been due to the severity of the skeletal problems. The patients in the present study had smaller SNBs, and higher frequencies of full step Class II molar relationships and posterior crossbites, all of which made it more difficult to correct the occlusal relationships. Assuming Class I had been attained during treatment, there could have been relapse, which is expected after the fixed appliance phase among patients treated with Herbst,[39] especially among those with persisting lip-tongue habits and unstable cuspal interdigitation at the end of treatment.[40] Finally, the unfavorable outcomes could also have been due to the relatively strict criteria used in the present study (i.e. maxillary canine and first molar mesiobuccal cusp tips needed to align within 1 mm of the embrasure and the buccal grooves of the mandibular canines and molars, respectively) for classification.

Overjet was the least likely component contributing to unfavorable dental outcomes. Only 21.6% of the patients with unfavorable dental outcomes had insufficient overjet correction. This could have been due to the fact that overjet correction with Herbst can be achieved several ways, including the restraint of the anterior maxillary movement, anterior movement of the mandible, posterior movement of the maxillary incisors, and anterior movement of the mandibular incisors. [39] Any or all of these movements make it easier to correct overjet than molar and canine relationships.

Patients with favorable treatment outcomes had larger pretreatment SNB angles than those with unfavorable outcomes. While no other study has focused on posttreatment outcomes, successful reduction in overjet [23] and ANB [22] have been associated with smaller SNB angles. This indicates that subjects with less retognathic mandibles have better chance of having favorable dental outcomes and that suggests with more retrognthic mandibles have greater potential for treatment changes.

Decision trees provide an excellent and accurate way to predict treatment outcomes. [31, 32] In the present study, the accuracy for the unpruned tree was 81.4% for the training sample and 78.4 % when validated on the testing sample. Recently, DTs have been shown to more accurately predict anterorposterior skeletal relations than discriminant analysis. [33] DTs have also been shown to be substantially better at predicting favorable and unfavorable growth among untreated Class III subjects than discriminant analysis (87.9% vs 59.3%). [32]

SN-MP angle is the most important variable in determining the success of dental outcomes among patients treated with Herbst appliances. In the present study, patients who had SN-MP angles ≥ 30° had a 12% chance of a favorable outcome, while those with angles < 30° had a 56% chance of having favorable dental outcomes. Divergence is the most important factor because it is directly related to the AP position of the mandible and chin. [4, 41] If hyperdivergence results in less anterior chin displacement, it will also produce less AP displacements of the molar and canines. In comparison, hypodivergent patients should be expected to exhibit greater anterior displacement of the teeth. [42] The surgical literature has also shown that the autorotation of the mandible associated with Le Fort 1 maxillary impaction to treat open bite can change a class II to class I relationships, and even possibly cause a class III relationship in subjects starting with class I relationship. [43]

In addition to the mandibular plane angle, overbite is an important determinant of treatment outcomes. A combination of increased mandibular plane angle and open bite has previously been associated with unsuccessful treatment outcomes with activator appliances. [21] This association was not identified in the present study because there were too few open bite cases. In the present study, patients with overbite < 7.3 mm had favorable treatment outcomes 61% of the time. This is consistent with previous studies showing favorable treatment outcome after twin block therapy among patients with overbite of 4.6 ± 0.6 mm. [23] However, it is important to emphasize that pretreatment overbites ≥ 7.3mm lead to unfavorable outcomes all of the time. In other words, there is a “sweet spot” between open bite and excessive deep bite that allows for the production of successful treatment outcomes.

The third variable that has to be considered is L1-MP. Patients with proclined lower incisors and less prominent chins were less likely to have favorable dental treatment outcomes. Patients with L1-MP ≥ 99◦ were less likely to have favorable treatment outcomes, especially if the distance Pg-NB was ≥ 2.8 mm. Since overjet reduction among Herbst patients is primarily dental, [44] treatment, which proclines the mandibular incisors an average 10.8 degrees, [45] may exceed the biological limit among those who start treatment with proclined lower incisors. Correction is even more difficult if the lower teeth are proclined and the mandible is retrognathic. However, if the mandibular incisors are not proclined, the probability of favorable outcome is highly likely (86% of the time), assuming the patients have average or less than average SN-MP angles, no more than moderate overbite and less than full-step molar relations.

Additional considerations are necessary for patients with full-step Class II molar relationships. When upper incisor proclination is average or less than average, favorable outcomes are expected 83% of the time, suggestive of a Class II division 2 malocclusions. In the present study, the 11% of patients who had full-step molar relation and retroclined incisors had overbites of 5 mm, interincisal angles of 145°s, and SN-MP angle of 23°.

Full step Class II patients with average or proclined upper incisors are very likely (75%) to have favorable treatment outcomes, whereas those with retroclined incisors were not. The cephalometric features of these two groups indicated large overjet, slight to moderate overbite and small interincisal angles, all of which are suggestive of Class II division 1 malocclusions (Table 13). During treatment with functional appliances, the maxillary incisors typically retrocline and the mandibular incisors procline. [45, 46]

The initial inclinations of the incisors affect treatment outcomes because they are largely related to how functional appliances work. However, previous studies have not examined the relationship between initial incisor positions and treatment outcomes with Herbst. The current study had more favorable outcomes among patients who had U1-SN angle ≥ 114° than among patients who had U1-SN angle < 114°, despite the fact that both groups represented the Class II division 1 patients. These suggest that the first group (U1-SN angle ≥ 114°) had more favorable growth patterns compared to the other group that had more unfavorable growth patterns. It is well established that favorable mandibular growth can improve the sagittal discrepancies even without treatment. [28, 47] However, it is difficult to predict which subjects will fall into that category. Using a Herbst for patients with unfavorable growth patterns can cause undesirable treatment outcomes.[28]

Conclusions

- Approximately 50% of the Herbst patients have unfavorable dental treatment outcome.
- Lack of Class I canine relationships (94.6%) is the primary reason for having unfavorable dental outcomes.
- Decision trees can predict successful or unsuccessful dental treatment outcomes approximately 80% of the time.
- The primary variables that determine successful or unsuccessful dental treatment outcomes are the SN-MP, followed by overbite, and L1-MP, respectively.

References

1. Proffit WR, F.H., Moray LJ., *Proffit WR, Fields HW Jr, Moray LJ.* Int J Adult Orthod Orthognath Surg, 1998. 13 (2): p. 97-106.

2. Milacic M, Markovic M. *A comparative occlusal and cephalometric study of dental and skeletal anteroposterior relationships.* Br J Orthod, 1983. 10(1): p. 53-54.

3. Cozza P, Baccetti T, Franchi L, De Toffol L, McNamara JA Jr. *Mandibular changes produced by functional appliances in Class II malocclusion: a systematic review.* Am J Orthod Dentofacial Orthop, 2006. 129(5): p. 599. e1-599. e12.

4. LaHaye MB, Buschang PH, Alexander RG, Boley JC. *Orthodontic treatment changes of chin position in Class II Division 1 patients.* Am J Orthod Dentofacial Orthop, 2006. 130(6): p. 732-741.

5. Baccetti T, Franchi L, McNamara JA Jr, Tollaro I. *Early dentofacial features of Class II malocclusion: a longitudinal study from the deciduous through the mixed dentition.* Am J Orthod Dentofacial Orthop, 1997. 111(5): p. 502-509.

6. Bishara SE, Jakobsen JR, Vorhies B, Bayati P. *Changes in dentofacial structures in untreated Class II division 1 and normal subjects: a longitudinal study.* Angle Orthod, 1997. 67(1): p. 55-66.

7. Bishara SE, Hoppens BJ, Jakobsen JR, Kohout FJ. *Changes in the molar relationship between the deciduous and permanent dentitions: a longitudinal study.* Am J Orthod Dentofacial Orthop, 1988. 93(1): p. 19-28.

8. Arya BS, Savara BS, Thomas DR. *Prediction of first molar occlusion.* Am J Orthod, 1973. 63(6): p. 610-621.

9. Bishara SE, Ziaja RR. *Functional appliances: a review.* Am J Orthod Dentofacial Orthop, 1989. 95(3): p. 250-258.

10. Pancherz H. *Treatment of Class II malocclusions by jumping the bite with the Herbst appliance: a cephalometric investigation.* Am J Orthod. 1979. 76(4): p. 423-442.

11. Pancherz H. *The Herbst appliance—its biologic effects and clinical use.* Am J Orthod. 1985. 87(1): p. 1-20.

12. Pancherz H. *The effects, limitations, and long-termdentofacial adaptations to treatment with the herbst appliance*. Semin Orthod, 1997. 3(4):232-43.

13. Franchi L, Baccetti T, McNamara Jr JA. *Treatment and posttreatment effects of acrylic splint Herbst appliance therapy.* Am J Orthod Dentofacial Orthop, 1999. 115(4): p. 429-438.

14. Pancherz H. *The mechanism of Class II correction in Herbst appliance treatment: a cephalometric investigation.* Am J Orthod, 1982. 82(2): p. 104-113.

15. Alves PF, Oliveira AG. *A comparison of the skeletal, dental, and soft tissue effects caused by herbst and mandibular protraction appliances in the treatment of mandibular Class II malocclusions.* World J. Orthod, 2008. 1;9(1).

16. Valant JR, Sinclair PM. *Treatment effects of the Herbst appliance.* Am J Orthod Dentofacial Orthop, 1989. 95(2): p. 138-147.

17. Schiavoni R, Grenga V, Macri V. *Treatment of Class II high angle malocclusions with the Herbst appliance: a cephalometric investigation.* Am J Orthod Dentofacial Orthop, 1992. 102(5): p. 393-409.

18. Yang X, Zhu Y, Long H, Zhou Y, Jian F, Ye N, Gao M, Lai W. *The effectiveness of the Herbst appliance for patients with Class II malocclusion: a meta-analysis.* Eur J Orthod, 2015. 38(3): p. 324-333.

19. Burstone CJ. *Lip posture and its significance in treatment planning.* Am J Orthod, 1967. 53(4): p. 262-284.

20. Ahlgren J, Laurin C. *Late results of activator-treatment: a cephalometric study.* Br. J. Orthod, 1976. 3(3): p. 181-187.

21. Pancherz HA. *The mandibular plane angle in activator treatment.* Angle Orthod, 1979. 49(1): p. 11-20.

22. Patel HP, Moseley HC, Noar JH. *Cephalometric determinants of successful functional appliance therapy.* Angle Orthod, 2002. 72(5): p. 410-417.

23. Caldwell S, Cook P. *Predicting the outcome of twin block functional appliance treatment: a prospective study.* Eur J Orthod, 1999. 21(5): p. 533-539.

24. Ahn SJ, Kim JT, Nahm DS. *Cephalometric markers to consider in the treatment of Class II Division 1 malocclusion with the bionator.* Am J Orthod Dentofacial Orthop, 2001. 119(6): p. 578-586.

25. Franchi L, Baccetti T. *Prediction of individual mandibular changes induced by functional jaw orthopedics followed by fixed appliances in Class II patients.* Angle Orthod, 2006. 76(6): p. 950-954.

26. Bondevik O. *How effective is the combined activator-headgear treatment?* Eur. J. Orthod, 1991. 13(6): p. 482-485.

27. O'brien K, Wright J, Conboy F, Sanjie Y, Mandall N, Chadwick S, Connolly I, Cook P, Birnie D, Hammond M, Harradine N. *Effectiveness of treatment for Class II malocclusion with the Herbst or twin-block appliances: a randomized, controlled trial.* Am J Orthod Dentofacial Orthop, 2003. 124(2): p. 128-137.

28. Rogers K, Campbell PM, Tadlock L, Schneiderman E, Buschang PH. *Treatment changes of hypo-and hyperdivergent Class II Herbst patients.* Angle Orthod, 2018. 88(1): p. 3-9.

29. Zarei A, El-Sharkawi M, Hairfield M, King G. *An intelligent system for prediction of orthodontic treatment outcome*. The 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 2006, p. 2702-2706.

30. Kim BM, Kang BY, Kim HG, Baek SH. *Prognosis prediction for class III malocclusion treatment by feature wrapping method.* Angle Orthod, 2009. 79(4): p. 683-691.

31. Chiang HJ, Tseng CC, Torng CC. Torng, *A retrospective analysis of prognostic indicators in dental implant therapy using the C5. 0 decision tree algorithm.* J. Dent. Sci, 2013. 8(3): p. 248-255.

32. Auconi P, Scazzocchio M, Caldarelli G, Nieri M, McNamara JA, Franchi L. *Understanding interactions among cephalometrics variables during growth in untreated Class III subjects.* Eur J Orthod, 2017. 39(4): p. 395-401.

33. Asiri S. *Can Decision Trees Accurately Predict Anteroposterior Skeletal Relations?* Dissertation, 2021.

34. Kim K. *A hybrid classification algorithm by subspace partitioning through semi-supervised decision tree.* Pattern Recognit, 2016. 60: p. 157-163.

35. Agrawal GL, Gupta H. *Optimization of C4. 5 decision tree algorithm for data mining application.* Int. J. Emerging Technol. Adv. Eng, 2013. 3(3): p. 341-345.

36. Baccetti T, Franchi L, McNamara Jr JA. *The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics*. Semin Orthod, 2005. 11(3):p. 119-129

37. Strobl C, Malley J, Tutz G. *An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests.* Psychol. Methods, 2009. 14(4): p. 323.

38. Therneau T, Atkinson B, Ripley B, Ripley MB. *Package 'rpart'.* Available online: cran.r-project.org/web/packages/rpart/rpart.pdf (accessed on 18 March 2021). 2015 Jun 29.

39. Wigal TG, Dischinger T, Martin C, Razmus T, Gunel E, Ngan P. *Stability of Class II treatment with an edgewise crowned Herbst appliance in the early mixed dentition: Skeletal and dental changes.* Am J Orthod Dentofacial Orthop, 2011. 140(2): p. 210-223.

40. Pancherz H. *The nature of Class II relapse after Herbst appliance treatment: a cephalometric long-term investigation.* Am J Orthod Dentofacial Orthop, 1991. 100(3): p. 220-233.

41. Buschang PH, Jacob HB. *Mandibular rotation revisited: what makes it so important?* Semin Orthod, 2014. 20(4):p. 299-315.

42. Bjok A, Skieller V. *Facial development and tooth eruption: an implant study at the age of puberty.* Am J Orthod, 1972. 62(4): p. 339-383.

43. Reyneke JP, Ferretti C. *Anterior open bite correction by Le Fort I or bilateral sagittal split osteotomy.* Oral Maxillofac Surg Clin North Am, 2007. 19(3): p. 321-338.

44. Baysal A, Uysal T. *Dentoskeletal effects of Twin Block and Herbst appliances in patients with Class II division 1 mandibular retrognathy.* Eur J Orthod, 2014. 36(2): p. 164-172.

45. Hansen K, Koutsonas TG, Pancherz H. *Long-term effects of Herbst treatment on the mandibular incisor segment: a cephalometric and biometric investigation.* Am J Orthod Dentofacial Orthop, 1997. 112(1): p. 92-103.

46. McNamara JA Jr, Bookstein FL, Shaughnessy TG. *Skeletal and dental changes following functional regulator therapy on Class II patients.* Am J Orthod, 1985. 88(2): p. 91-110.

47. Björk A, Pallilng M. *Adolescent age changes in sagittal jaw relation, alveolar prognathy, and incisal inclination.* Acta Odontol Scand. 1955. 12(3-4): p. 201-232.

48. Burkhardt DR, McNamara JA Jr, Baccetti T. *Maxillary molar distalization or mandibular enhancement: a cephalometric comparison of comprehensive orthodontic treatment including the pendulum and the Herbst appliances.* Am J Orthod Dentofacial Orthop, 2003. 123(2): p. 108-116.

49. Pancherz H, Hensen K. *Occlusal changes during and after Herbst treatment: a cephalometric investigation.* Eur J Orthod, 1986. 8(4): p. 215-28.

50. de Almeida MR, Henriques JF, Rodrigues de Almeida R, Ursi W, McNamara Jr JA. *Short-term treatment effects produced by the Herbst appliance in the mixed dentition*. Angle Orthod, 2005. 75(4): p. 540-7.

51. McNamara Jr JA, Howe RP, Dischinger TG. *A comparison of the Herbst and Fränkel appliances in the treatment of Class II malocclusion*. Am J Orthod Dentofacial Orthop, 1990. 98(2): p. 134-44.

52. Sidhu MS, Kharbanda OP, Sidhu SS. *Cephalometric analysis of changes produced by a modified Herbst appliance in the treatment of Class II division 1 malocclusion*. Br J Orthod, 1995. 22(1): p. 1-2.

CHAPTER V

PREDICTION OF SKELETAL TREATMENT OUTCOMES AMONG HERBST PATIENTS USING DECISION TREES

Introduction

Functional appliances are often used to treat class II dental and skeletal malocclusions. [1-3] They work by altering the sagittal and vertical positions of the mandible, producing both orthopedic and orthodontic changes. [4-6] The Herbst appliance, which is among the most widely used functional appliance, works by inhibiting maxillary growth, altering mandibular growth, and producing dentoalveolar changes. [2, 7, 8] The fact that it is fixed eliminates the compliance concerns associated with removable appliances. [4, 9] The skeletal effect of the Herbst appears to be mainly on the maxilla (due to the headgear effect), which accounts for a greater amount of ANB reduction than the mandible. [1, 10, 11] Importantly, the treatment outcomes with Herbst are not always favorable. [11] Some patients finish treatment with mandibles in a worse position than if it had been left untreated. This happens to hyperdivergent patients treated with Herbst appliances, who typically show true backward rotation, which brings the chin and mandible back. [11] Criteria are needed to determine whether Herbst patients are likely to have favorable or unfavorable treatment outcomes remain unknown.

Several criteria have been used in the past to classify treatment outcomes with functional appliances as favorable or unfavorable. These criteria, whether skeletal or dental, have focused primarily on the changes in the anteroposterior (AP) dimension.

The most commonly dental criteria for success for class II patients are normal molar and canine relationships. [12, 13] A reduction of overjet to less than 3 mm at the end of treatment, [14] and a percentage reduction relative to pretreatment overjet, [15] have also been used. Skeletally, a 4◦ ANB angle has often been used to distinguish between class I and class II patients is well established. [16-19] However, a 4◦ cut-off cannot be used for cases with large initial ANB angles, for whom a 50% reduction is considered possible and favorable. [20] An ANB change of 3◦ or more [21] and an increase in mandibular length of at least 5.3 mm [22] have also been used to describe successful treatment, due to the different criteria that have been used, there is substantial variability in the proportion of patients reported to have been successfully treated with functional appliances (35%-67%). [12, 14, 21, 22] Most often, a cases are successfully treated based on occlusal criteria, but remain class II skeletally (Table 14). This indicates that skeletal and dental outcomes should be evaluated separately based on a clearly defined criterion.

In addition to the AP dimension, the vertical dimension also has to be considered when evaluating skeletal treatment outcomes. Based on a recent meta-analysis, the mean changes in SN-MP angle with Herbst appliances range between -0.09° to 0.04°. [23] In other words, approximately 50% of the patients should be expected to have an unfavorable increase in the SN-MP angle. That being the case, any increase of the SN-MP angle greater than 1° should be considered unfavorable.

Studies have identified pre-treatment features that predict treatment outcomes with other functional appliances. Overbite and the SNB angle have been shown to explain 52% of the reduction of overjet with twin block appliance. [15] Pretreatment lip thicknesses were found to be the most important variables for discriminating between the good and poorly treated bionator cases, followed by the FMIA, the U1 to facial plane, and the articular angle and the L1 to A-Pg. [12] The one study that attempted to predict skeletal treatment outcomes with functional appliances used the change in overall mandibular length to classify patients as good or bad responders. [22] However, mandibular length does not describe the AP relationship, and it could be misleading. Overall mandibular length increases among class II patients due to a redirection of condylar growth, with little or no effect on AP skeletal relations, especially among hyperdivergent patients. The gonial angle (Co-Go-Me◦) has been identified as the best pre-treatment predictor of good and bad responders to Herbst and Twin-Blocks therapy. [22] However, Herbst and Twin-Blocks have been shown to have different skeletal effects. [24] The prediction of skeletal outcomes among Herbst patients is currently not possible because it has not been attempted.

Previous studies predicting treatment outcomes have relied on traditional statistics, which appear to be less accurate than artificial intelligence (AI) and machine learning (ML). For example, a ML model was able to accurately predict the prognosis of Class III treatment better than discriminant analysis (DA) (97.2% for ML and 92.1% for DA). [25] Based on the pre-treatment PAR index, a neural network model was successful at predicting the final PAR score 94.0% of the time, whereas linear regression was accurate only 82.0% of the time. [26] When used among untreated class III patients, decision trees (DTs) have been shown to be more accurate than discriminant analysis (87.9% vs 59.3%) in classifying class III patients as good and bad growers. [27]

Among the ML algorithms, DTs algorithms are the easiest to understand and interpret, even by non-experts. [28, 29] In addition, they can detect non-linear relationships between the variables, which is often the case in clinical studies due to variable interactions. [30] DTs can be used even when there are missing data and can work well with both categorical and numerical variables. [31, 32] Unlike traditional statistics, they do not require underlying assumptions because they are non-parametric. [33] To date, DTs have not been used to predict skeletal treatment outcomes of class II Herbst patients.

The purpose of the present study was to determine the applicability of DTs for predicting skeletal treatment outcomes among class II patients treated with Herbst previously. The objective was to develop accurate predictions and identify the most important variables contributing to the predictions.

Materials and Methods

This retrospective study utilized the cephalometric records of patients who had been previously treated with the Herbst and fixed appliances. The records were obtained from two private practices, and the Graduate Orthodontic Department of Texas A&M University. The study was approved by Texas A&M University IRB (approval # 2019-1238D). The selection criteria for the study included 1) pretreatment class II skeletal relations [ANB > 4°]; 2) at least half-step pretreatment class II molar and canine relationships (unilateral or bilateral); 3) pretreatment ages between 9 and 16 years; 4) treated with the Herbst (for at least 6 months) and fixed appliances; 5) good quality pretreatment [T1] and posttreatment [T2] cephalograms taken with lips lightly touching; 6) no syndromes, craniofacial anomalies, or congenitally missing teeth; and 7) no orthognathic or cosmetic facial surgery.

A total of 116 patients (77 males, 39 females) were selected. The mean ages of the subjects at T1 and T2 were 12.51 ± 1.7 years and 15.21 ± 1.6 years, respectively. The treatment time with the Herbst was 12.3 ± 2.8 months and the treatment duration was 31.6 ± 6.7 months. All patients were treated using standard cantilever Herbst appliances with full-coverage stainless steel crowns on the maxillary and mandibular first molars, telescopic cantilever arms from the mandibular first molars, and 0.040 mm stainless steel lower lingual arches with occlusal rests on the mandibular first premolars. During the Herbst phase, fixed appliances were bonded from the maxillary second premolar to second premolar, and from the mandibular canine to canine. After the Herbst removal, fixed appliances were bonded to the remaining teeth (maxillary and mandibular 6-6).

Evaluations

Cephalograms were digitized by the primary investigator using Dolphin Imaging software, version 11.0 (Dolphin Imaging, Chatsworth, CA, USA). In addition to age, sex, and the stage of cervical vertebral maturation (CVS), [34] 22 pre-treatment measurements were computed as predictor variables, including 9 angular, 12 linear, and one ratio measurements (Table 15 & Figure 12). All linear measurements were adjusted to eliminate magnification. For reliability, the lateral cephalograms of 8 randomly selected patients were remeasured. The method errors ranged between 0.14 and 0.32 for the linear measurements and between 0.33 and 0.60 for the angular measurements.

Subjects were classified as having favorable or unfavorable skeletal outcomes based on the post-treatment values of ANB and SN-MP. To be considered as a favorable skeletal outcome, two criteria had to have been met: 1) a posttreatment ANB angle $< 4°$, or a 50% decrease of the ANB if the pretreatment ANB was greater than 8°, and 2) $\leq 1°$ increase of the SN-MP angle. If either of these criteria was not satisfied, the subject was classified as having an unfavorable skeletal outcome.

Decision Trees and Statistical Analyses

Decision trees (DTs) are machine learning algorithms that have been shown to be successful in making predictions and understanding interactions among variables in dentistry [35] and orthodontics. [27] They classify by asking a series of yes-no questions. Each question is located at a node and each node points to another node until the terminal node is reached, depending on the answer to that question. A tree-like structure develops based on this hierarchy. This structure helps to visualize the variables' relationships and interactions. The most important variable is located at the top of the tree, followed by the next important variable and so on, until the terminal node is reached. [29] Each node of the tree denotes the outcome (either favorable or unfavorable), the probability of that outcome, and the percentage of subjects satisfying the previous condition. When using DTs, there is a process called pruning that eliminates the less important variables and reduces the size of the tree to make it less complex. Pruning can help improve the accuracy of the model, especially when the DT is complex and has many branches that make it prone to overfitting. [36] Confusion matrices are used along with the DTs to display the proportions of true positive, true negative, false positive, and false negative cases.

Of the 116 subjects that met the inclusion criteria, 91 (80%) were randomly selected and used to develop the DT model (training sample). The remaining 25 subjects (20%) were used to validate the model (testing sample). The DTs were produced using the R package ‘rpart’ version 4.1-15. SPSS version 25 (SPSS Inc, Chicago, Ill) was used for the traditional statistical analyses. Based on their skewness and kurtosis, the distributions were all normal. Independent sample T-tests were used to compare the pretreatment differences and the changes between the favorable and unfavorable groups. Pearson chi-square tests were used to compare the frequencies of favorable and unfavorable outcomes.

Results

Of the 63 subjects who had unfavorable skeletal outcomes, the AP dimension was the primary determining factor for 51 subjects, the change of SN-MP was unfavorable in 39 subjects (Table 16). There were no statistically significant sex and maturity stage (CVS) differences between subjects with favorable & unfavorable skeletal treatment outcomes (Table 17). There also were no differences in the frequencies of favorable skeletal outcomes among the three sites where the records were obtained (Table 18).

After Bonferroni adjustments for multiple comparisons, the pretreatment between-group differences showed that the unfavorable group had significantly larger

SN-MP, PP-MP, N-A-Pg, and S-Ar-Go angles, along with a significantly smaller anteroposterior facial height ratio (Table 19). Three pretreatment linear measurements, including ANS-Me, U1-NPg, and L1-NPg, were significantly larger among the unfavorable than favorable group. The changes of SNB, SN-MP, PP-MP, S-Go/N-Me, Ar-Go, and N-A-Pg showed statistically significant between-group differences, with the favorable group having more pronounced changes (Table 20).

Decision Trees

Only 4 variables were identified in the unpruned DT predicting the skeletal treatment outcomes among Herbst patients (Figure 13). There were two angular (N-A-Pg◦ and S-Ar-Go◦) and two linear (U1-NPg and Wits) measurements. The angle of facial convexity (N-A-Pg◦) was the most important variable. Among the 52% patients having facial convexity angle < 10◦, favorable skeletal treatment outcomes were expected 72% of the time. The remaining 48% with a facial convexity angle ≥ 10◦ had an 86% chance of having unfavorable outcomes.

The next two variables were U1-NPg and the S-Ar-Go angle, followed by Wits. If a patient had a convexity angle ≥ 10◦ and U1-NPg ≥ 6.8mm, the probability of an unfavorable treatment outcome was 97.0%. However, the probability of a favorable outcome was 62% if the patients' upper incisor was less than 6.8 mm from the facial plane. Patients with N-A-Pg angles < 10◦ and S-Ar-Go angles < 132◦ had a 95% probability of having a favorable outcome. This probability decreased to 56% if the S-Ar-Go ≥ 132◦. If the Wits was ≥ 5 mm, there was a 71% chance of having an unfavorable outcome. On the other hand, there was an 85% probability of a favorable outcome if the Wits was < 5 mm. The accuracy of the unpruned model was 87.9% when applied to the training data set and 84.0% when applied to the testing data set (Table 21 & Figure 14).

The pruned tree included only 3 variables and 4 nodes (Figure 15). The branching of the tree, the contribution of the variables, and the probabilities were the same as previously described for the unpruned tree. For those with N-A-Pg angles ≥ 10◦, the terminal node was reached immediately with an 86% chance of having unfavorable skeletal treatment outcome. There was a 95% change of have a favorable outcome when N-A-Pg < 10◦ and S-Ar-Go < 132◦. The accuracy of the pruned model was 85.7% on the training data set and 84.0% on the testing data set (Table 21 & Figure 16).

Discussion

Most patient (81%) in the present study had unfavorable skeletal treatment outcomes, due primarily to the orthodontists' inability to correct AP relations. Based on the criteria applied in the present study, the majority of the previous Herbst studies have reported unfavorable skeletal treatment outcomes (Table 14). [37-43] There are several possible reasons why most class II Herbst patients end treatment with unfavorable skeletal outcomes. The mandibular effects of the Herbst depend on mandibular divergence, with the mandibles of hyperdivergent patients rotating backward. [11] Many class II patients have unfavorable growth patterns, with AP skeletal relationships worsening over time. [44] In addition, the Herbst's appears to have a greater treatment effect on the maxilla than mandible (i.e. it produces a headgear effect). [1, 10, 11] Perhaps most importantly, orthodontists typically stop treatment when class I dental relationships (class I molar and canine) have been attained, whether or not the class II skeletal problems have been corrected.

Patients with unfavorable skeletal outcomes usually exhibit hyperdivergent growth patterns. Pretreatment between-group differences in the present study showed larger mandibular, palatal and gonial angles, greater convexity, longer lower face heights, and smaller posterior-to-anterior face height ratios among the unfavorable group, all of which are indicative of a hyperdivergent growth pattern. Previous studies that examined treatment outcomes of functional appliances have also shown hyperdivergent patterns among patients with unfavorable outcomes, including longer lower facial heights, [21] larger articular angles, [12] and larger angles of facial convexity. [12] Divergence is important because it is closely related to the positions of the chin and mandible. [40, 45] The Herbst appliance often reposition the chins and mandibles of hyperdivergent patients backward, resulting in unfavorable outcomes. [11] This shows that orthodontists must consider the pretreatment vertical dimension, in addition to the AP, when determining treatment objectives and evaluating skeletal treatment outcomes.

The angle of facial convexity and protrusion of the maxillary incisors were the two primary factors responsible for unfavorable treatment outcomes. Patients with facial convexity angles $\geq 10^{\circ}$ and maxillary incisors protrusion ≥ 6.8 mm almost always (97% of the time) had unfavorable outcomes. These cut-off values are smaller than cut-offs previously reported (i.e. 12° and 14 mm for N-A-Pg and U1-NPg, respectively) for unfavorable functional appliance treatment. [12] However, these cut-offs were based on the means and standard deviations, which could be problematic because the distribution might not be normally distributed due to the presence of outliers. [46] With decision trees, the cut-off values depend on the interactions between all the variables included in the model, which produces a more realistic and more accurate values that relate to the outcome variable. [32]

In contrast, hypodivergent patients treated with the Herbst tend to have favorable outcomes. In the present study, patients with favorable outcomes showed significantly greater decreases in all measures of divergence and greater chin projection. Hypodivergent patients treated with Herbst typically undergo forward true mandibular rotation, [11] which reduced divergence and brought the chin and mandible more forward. [40, 45]

The angle of facial convexity and the articular angle are the primary factors determining favorable skeletal outcomes. Patients with facial convexity angles < 10◦ and articular angles < 132◦ almost always (95%) had favorable outcomes. Significantly smaller facial convexity and articular angles have been previously reported among the favorably treated group than the unfavorably treated group with functional appliance. [12] Both measurements have been used to describe the anteroposterior jaw discrepancy. [47, 48] The convexity angle reflects the sagittal protrusion of the maxilla in relation to the facial profile (the convex or concave face), while the articular angle reflects the degree of retrusion or protrusion of the mandible in relation to the cranial base. Hence, class II subjects with less protruded maxillae and less retrognathic mandibles prior to treatment are more likely to have favorable skeletal outcomes. These variables are the most important for predicting favorable outcomes.

When the articular angle is ≥ 132◦, favorable or unfavorable outcomes depend on the Wits. There was an 85% chance of a favorable outcome if the pretreatment Wits was < 5 mm. Patients with more pronounced skeletal discrepancies (Wits ≥ 5 mm) were more likely to have unfavorable skeletal outcomes (71%). It is possible that the larger overjet among the group with the larger Wits (overjet of 8.7 mm and 6.4 mm for the unfavorable and favorable groups, respectively). Overjet has been positively correlated with the Wits, explaining 56% of its variation,[49] which is not surprising since the line of reference for the Wits is the functional occlusal plane, which is a dental parameter.

Decision trees provide an excellent way to predict treatment outcomes. [35, 50] The accuracy of the DT model in the present study was 87.9 %, which makes it the most accurate model currently available for predicting treatment outcomes with functional appliances. The predictive accuracies of previous studies that used traditional statistics to predict treatment outcomes with functional appliance ranged from 52% [15] to 80.4%. [22] DTs in the present study could have been more accurate because they can detect non-linear relationships among multiple covariates. [30] DTs can handle both numeric and categorical data unlike regression models, [31, 32] which gives it an advantage since several of the clinical variables are categorical.

The overall goal of the present study was to provide orthodontists with a way to predict patients' skeletal treatment outcomes. Both DTs developed in the present study (unpruned and pruned) were validated, making them applicable to other patients. They are also easy to apply to individual patients. For example, the pretreatment angles of facial convexity (which was the top variable in the DT) for both cases # 1 and 2 were < 10◦, and their articular angles were < the 132◦ cut-off value, which predicted a favorable treatment outcome 95% of the time (Figures 17 and 18). The angles of facial convexity of case #3 and case #4 were ≥ than the 10◦ cut off-value, and their maxillary incisors were more than 6.8mm from the facial plane, indicating unfavorable skeletal treatment outcomes 97% of the time (Figures 19 and 20). Based on just a few variables, skeletal outcomes with Herbst treatment can be predicted with high level of accuracy, which should help orthodontists practice more efficiently, reduce variability, and eliminate subjectivity when treating Class II skeletal patients.

Conclusions

- Unfavorable posttreatment's AP relationship was the primary factor associated with having unfavorable skeletal treatment outcomes.
- Patients with unfavorable outcomes tended to be more hyperdivergent than patients in the favorable group.
- Decision trees can accurately predict favorable and unfavorable skeletal treatment outcomes among Herbst patients approximately 87.9% of the time.
- The most important variables for the prediction of skeletal treatment outcomes is N-A-Pg, followed by U1-NPg, S-Ar-Go, and Wits, respectively.

References

1. McNamara JA Jr, Bookstein FL, Shaughnessy TG. *Skeletal and dental changes following functional regulator therapy on Class II patients.* Am J Orthod, 1985. 88(2): p. 91-110.

2. Pancherz H. *Treatment of Class II malocclusions by jumping the bite with the Herbst appliance: a cephalometric investigation.* Am J Orthod, 1979. 76(4): p. 423-442.

3. Pancherz H. *A cephalometric analysis of skeletal and dental changes contributing to Class II correction in activator treatment.* Am J Orthod, 1984. 85(2): p. 125-134.

4. Bishara SE, Ziaja RR. *Functional appliances: a review.* Am J Orthod Dentofacial Orthop, 1989. 95(3): p. 250-258.

5. Mills JR. *The effect of functional appliances on the skeletal pattern.* Br J Orthod, 1991. 18(4): p. 267-275.

6. Chen JY, Will LA, Niederman R. *Analysis of efficacy of functional appliances on mandibular growth.* Am J Orthod Dentofacial Orthop, 2002. 122(5): p. 470-476.

7. Pancherz H. *The Herbst appliance—its biologic effects and clinical use.* Am J Orthod 1985. 87(1): p. 1-20.

8. Pancherz H. *The mechanism of Class II correction in Herbst appliance treatment: a cephalometric investigation.* Am J Orthod, 1982. 82(2): p. 104-113.

9. Barton S, Cook PA. *Predicting functional appliance treatment outcome in Class II malocclusions–a review.* Am J Orthod Dentofacial Orthop, 1997. 112(3): p. 282-286.

10. Hansen K, Koutsonas TG, Pancherz H. *Long-term effects of Herbst treatment on the mandibular incisor segment: a cephalometric and biometric investigation.* Am J Orthod Dentofacial Orthop, 1997. 112(1): p. 92-103.

11. Rogers K, Campbell PM, Tadlock L, Schneiderman E, Buschang PH. *Treatment changes of hypo-and hyperdivergent Class II Herbst patients.* Angle Orthod, 2018. 88(1): p. 3-9.

12. Ahn SJ, Kim JT, Nahm DS. *Cephalometric markers to consider in the treatment of Class II Division 1 malocclusion with the bionator.* Am J Orthod Dentofacial Orthop, 2001. 119(6): p. 578-586.

13. Bondevik O. *How effective is the combined activator-headgear treatment?* Eur J Orthod, 1991. 13(6): p. 482-485.

14. Cohen AM. *A study of class II division 1 malocclusions treated by the Andresen appliance.* Br J Orthod, 1981. 8(3): p. 159-163.

15. Caldwell S, Cook P. *Predicting the outcome of twin block functional appliance treatment: a prospective study.* Eur J Orthod, 1999. 21(5): p. 533-539.

16. Steiner CC. *Cephalometrics for you and me.* Am J Orthod, 1953. 39(10): p. 729-755.

17. Cha BK, Kim CH, Baek SH. *Skeletal sagittal and vertical facial types and electromyographic activity of the masticatory muscle.* Angle Orthod, 2007. 77(3): p. 463-470.

18. Meyer-Marcotty P, Reuther T, Stellzig-Eisenhauer A. *Bridging of the sella turcica in skeletal Class III subjects.* Eur J Orthod, 2010. 32(2): p. 148-153.

19. Uzuner FD, Aslan BI, Dinçer M. *Dentoskeletal morphology in adults with Class I, Class II Division 1, or Class II Division 2 malocclusion with increased overbite.* Am J Orthod Dentofacial Orthop, 2019. 156(2): p. 248-256. e2.

20. Steiner CC. *Cephalometrics in clinical practice.* Angle Orthod, 1959. 29(1): p. 8-29.

21. Patel HP, Moseley HC, Noar JH. *Cephalometric determinants of successful functional appliance therapy.* Angle Orthod, 2002. 72(5): p. 410-417.

22. Franchi L, Baccetti T. *Prediction of individual mandibular changes induced by functional jaw orthopedics followed by fixed appliances in Class II patients.* Angle Orthod, 2006. 76(6): p. 950-954.

23. Yang X, Zhu Y, Long H, Zhou Y, Jian F, Ye N, Gao M, Lai W. *The effectiveness of the Herbst appliance for patients with Class II malocclusion: a meta-analysis.* Eur J Orthod, 2015. 38(3): p. 324-333.

24. Schaefer AT, McNamara JA Jr, Franchi L, Baccetti T. *A cephalometric comparison of treatment with the Twin-block and stainless steel crown Herbst appliances followed by fixed appliance therapy.* Am J Orthod Dentofacial Orthop, 2004. 126(1): p. 7-15.

25. Kim, B.-M., et al., *Prognosis prediction for class III malocclusion treatment by feature wrapping method.* Angle Orthod, 2009. 79(4): p. 683-691.

26. Zarei A, El-Sharkawi M, Hairfield M, King G. *An intelligent system for prediction of orthodontic treatment outcome*. The 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 2006, p. 2702-2706.

27. Auconi P, Scazzocchio M, Caldarelli G, Nieri M, McNamara JA, Franchi L. *Understanding interactions among cephalometrics variables during growth in untreated Class III subjects.* Eur J Orthod, 2017. 39(4): p. 395-401.

28. Kotsiantis SB. *Decision trees: a recent overview.* Artif Intell Rev, 2013. 39(4): p. 261-283.

29. Kingsford C, Salzberg SL. *What are decision trees?* Nat Biotechnol, 2008. 26(9): p. 1011.

30. Doupe P, Faghmous J, Basu S. *Machine learning for health services researchers.* Value Health, 2019. 22(7): p. 808-815.

31. Pal M, Mather PM. *An assessment of the effectiveness of decision tree methods for land cover classification.* Remote Sens Environ, 2003. 86(4): p. 554-565.

32. Ramezankhani A, Hadavandi E, Pournik O, Shahrabi J, Azizi F, Hadaegh F. *Decision tree-based modelling for identification of potential interactions between type 2 diabetes risk factors: a decade follow-up in a Middle East prospective cohort study.* BMJ Open, 2016. 6(12): p. e013336.

33. Friedl, M.A. and C.E. Brodley, *Decision tree classification of land cover from remotely sensed data.* Remote sensing of environment, 1997. 61(3): p. 399-409.

34. Baccetti T, Franchi L, McNamara Jr JA. *The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics*. Semin Orthod, 2005. 11(3):p. 119-129.

35. Chiang HJ, Tseng CC, Torng CC. Torng, *A retrospective analysis of prognostic indicators in dental implant therapy using the C5. 0 decision tree algorithm.* J. Dent. Sci, 2013. 8(3): p. 248-255.

36. Strobl, C., J. Malley, and G. Tutz, *An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests.* Psychological methods, 2009. 14(4): p. 323.

37. Burkhardt DR, McNamara JA Jr, Baccetti T. *Maxillary molar distalization or mandibular enhancement: a cephalometric comparison of comprehensive orthodontic treatment including the pendulum and the Herbst appliances.* Am J Orthod Dentofacial Orthop, 2003. 123(2): p. 108-116.

38. Valant JR, Sinclair PM. *Treatment effects of the Herbst appliance.* Am J Orthod Dentofacial Orthop, 1989. 95(2): p. 138-147.

39. Pancherz H, Hansen K. *Occlusal changes during and after Herbst treatment: a cephalometric investigation.* Eur J Orthod, 1986. 8(4): p. 215-228.

40. LaHaye MB, Buschang PH, Alexander RG, Boley JC. *Orthodontic treatment changes of chin position in Class II Division 1 patients.* Am J Orthod Dentofacial Orthop, 2006. 130(6): p. 732-741.

41. de Almeida MR, Henriques JF, de Almeida RR, Weber U, McNamara JA Jr. *Short-term treatment effects produced by the Herbst appliance in the mixed dentition.* Angle Orthod, 2005. 75(4): p. 540-547.

42. McNamara JA Jr, Howe RP, Dischinger TG. *A comparison of the Herbst and Fränkel appliances in the treatment of Class II malocclusion.* Am J Orthod Dentofacial Orthop, 1990. 98(2): p. 134-144.

43. Sidhu MS, Kharbanda OP, Sidhu SS. *Cephalometric analysis of changes produced by a modified Herbst appliance in the treatment of Class II division 1 malocclusion.* Br J Orthod, 1995. 22(1): p. 1-12.

44. Buschang PH, Martins J. *Childhood and adolescent changes of skeletal relationships.* Angle Orthod, 1998. 68(3): p. 199-208.

45. Buschang PH, Jacob HB. *Mandibular rotation revisited: what makes it so important?* Semin Orthod, 2014. 20(4):p. 299-315.

46. Ratcliff R. *Methods for dealing with reaction time outliers.* Psychol Bull, 1993. 114(3): p. 510 -32.

47. Downs WB. *Variations in facial relationships: Their significance in treatment and prognosis.* Am J Orthod, 1948. 34(10): p. 812-840.

48. Björk A. *Some biological aspects of prognathism and occlusion of the teeth.* Acta Odontol Scand, 1950. 9(1): p. 1-40.

49. Zupancic S, Pohar M, Farcnik F, Ovsenik M. *Overjet as a predictor of sagittal skeletal relationships.* Eur J Orthod, 2008. 30(3): p. 269-273.

50. Chu CS, Lee NP, Adeoye J, Thomson P, Choi SW. *Machine learning and treatment outcome prediction for oral cancer.* J Oral Pathol Med, 2020. 49(10): p. 977-985.

51. Wigal TG, Dischinger T, Martin C, Razmus T, Gunel E, Ngan P. *Stability of Class II treatment with an edgewise crowned Herbst appliance in the early mixed dentition: Skeletal and dental changes.* Am J Orthod Dentofacial Orthop, 2011. 140(2): p. 210-223.

CHAPTER VI

CONCLUSIONS

Since the present study evaluated several aspects of artificial intelligence and machine learning application in orthodontics, the conclusions will be summarized by chapters starting with chapter II:

Based on the literature review regarding the application of artificial intelligence and machine learning in orthodontics, the following conclusions can be drawn:

1. Artificial intelligence and machine learning systems applied in orthodontics provide promising tools that can improve clinical practice.
2. These clinical decision support systems can help orthodontists practice more efficiently, reduce variability and subjectivity.
3. Algorithms should be expected to improve, making it possible to handle more complex data such as images.
4. AI models are meant to help with the clinical judgment and not substitute the knowledge and expertise of humans.

Chapter III evaluated the applicability of decision trees for predicting maxillomandibular relationship among untreated Class I and Class II subjects, the following conclusions can be drawn:

1. The unpruned and pruned decision trees were able to successfully classify the growth of untreated subjects 85.4% and 83.2% of the time, respectively.
2. The validation sample showed that the pruned decision tree was more accurate than the unpruned (81.8% vs 75.0%) tree.

3. The decision trees identified the Y-axis as the most important variable to classify growth, followed, in order, by ANS-N-PG, NSB, and MPA.

Chapter IV investigated the ability of decision trees for predicting dental treatment outcomes among patients with Class II dental relationship treated with Herbst, and based on the findings of the present study the following conclusions can be drawn:

1. Approximately 50% of the Herbst patients have unfavorable dental treatment outcome.
2. Lack of Class I canine relationships (94.6%) is the primary reason for having unfavorable dental outcomes.
3. Decision trees can predict successful or unsuccessful treatment outcomes approximately 80% of the time.
4. The primary variables that determine successful or unsuccessful dental treatment outcomes are the SN-MP, followed by overbite, and L1-MP, respectively.

Chapter V assessed the possibility of using decision trees for predicting skeletal treatment outcomes among patients with Class II skeletal relationship treated with Herbst, the following conclusions can be drawn:

1. Unfavorable posttreatment's AP relationship was the primary factor associated with having unfavorable skeletal treatment outcomes.
2. Patients with unfavorable outcomes tended to be more hyperdivergent than patients in the favorable group.

3. Decision trees can accurately predict favorable and unfavorable skeletal treatment outcomes among Herbst patients approximately 87.9% of the time.
4. The most important variables for the prediction of skeletal treatment outcomes are N-A-Pg, followed by U1-NPg, S-Ar-Go, and Wits, respectively.

APPENDIX A

FIGURES

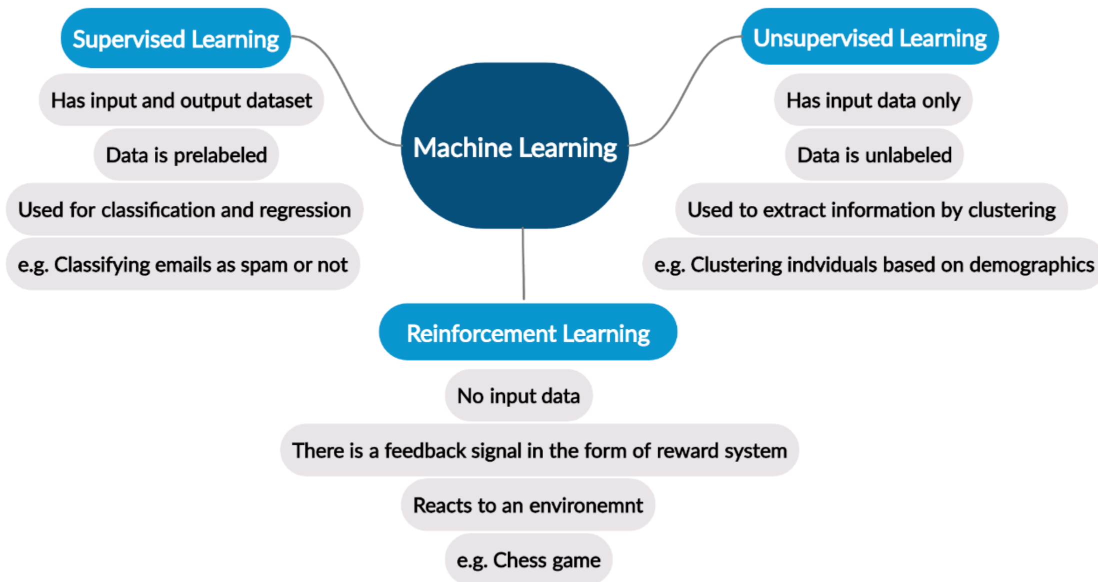

Figure 1. Types of machine learning based on the training method. Reprinted with permission from, “Applications of artificial intelligence and machine learning in orthodontics”, by Asiri SN, Tadlock LP, Schneiderman E, Buschang PH. APOS Trends in Orthodontics, 2020;10(1):17-24. Copyright [2019] by Scientific Scholar on behalf of APOS Trends in Orthodontics.

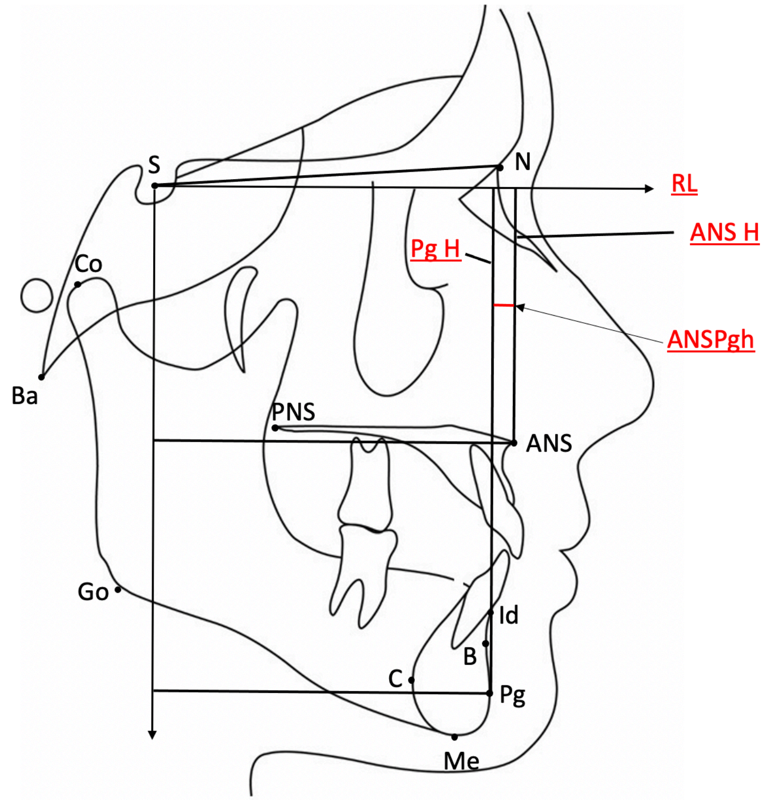

Figure 2. Landmarks evaluated on subject cephalograms and the horizontal relationship between ANS and Pg transferred to the natural structure reference line.

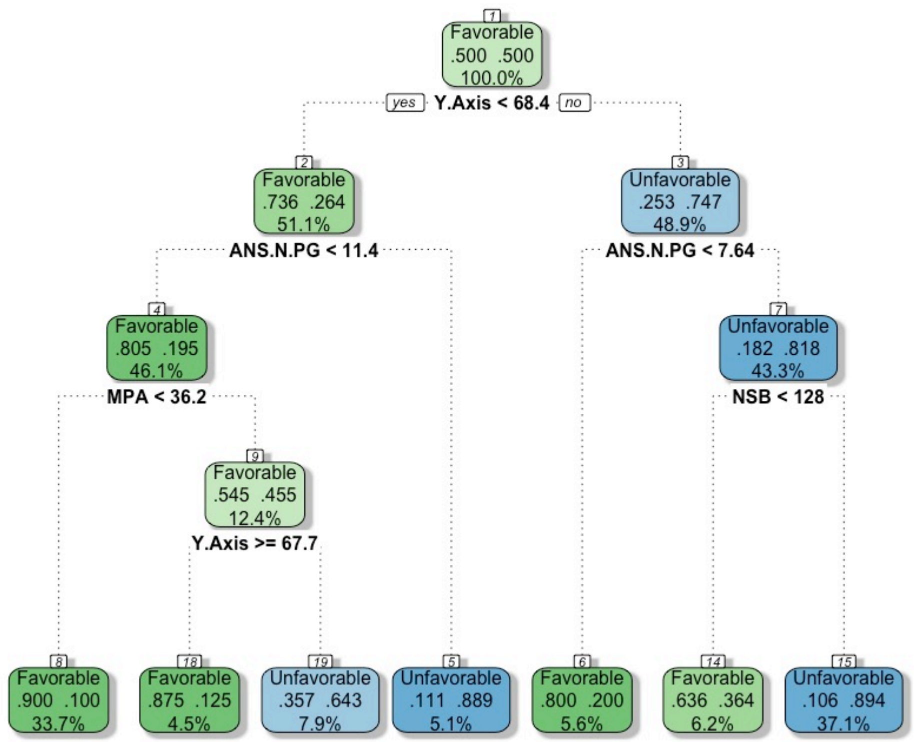

Figure 3. Unpruned tree for the 178 longitudinal untreated class I and class II subjects at age 10 to determine the skeletal relationship at age 15.

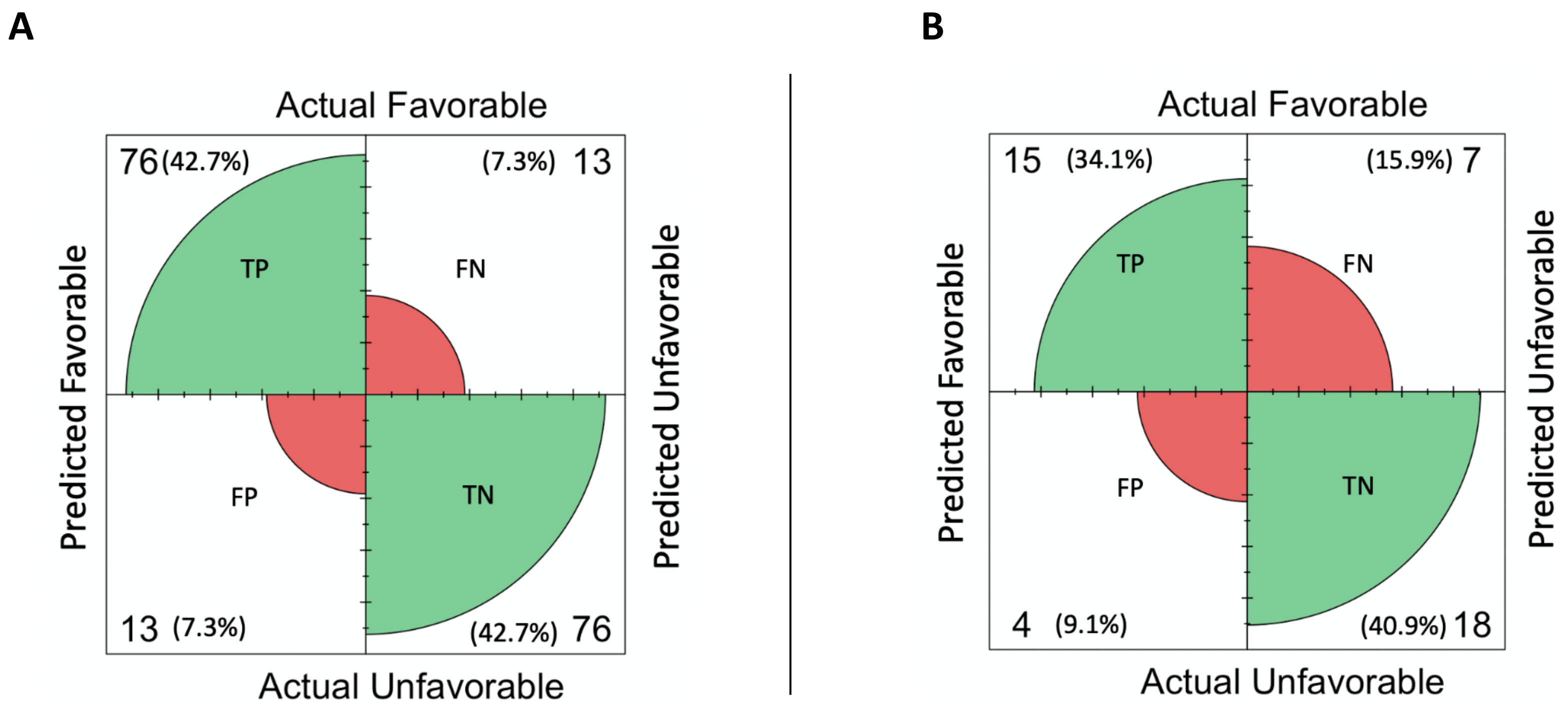

Figure 4. Confusion matrices of the A) training sample, and B) testing sample showing the absolute and the relative (%) number of subjects having true positive (TP), false positive (FP), true negative (TN), and false negative (FN) results of the unpruned model.

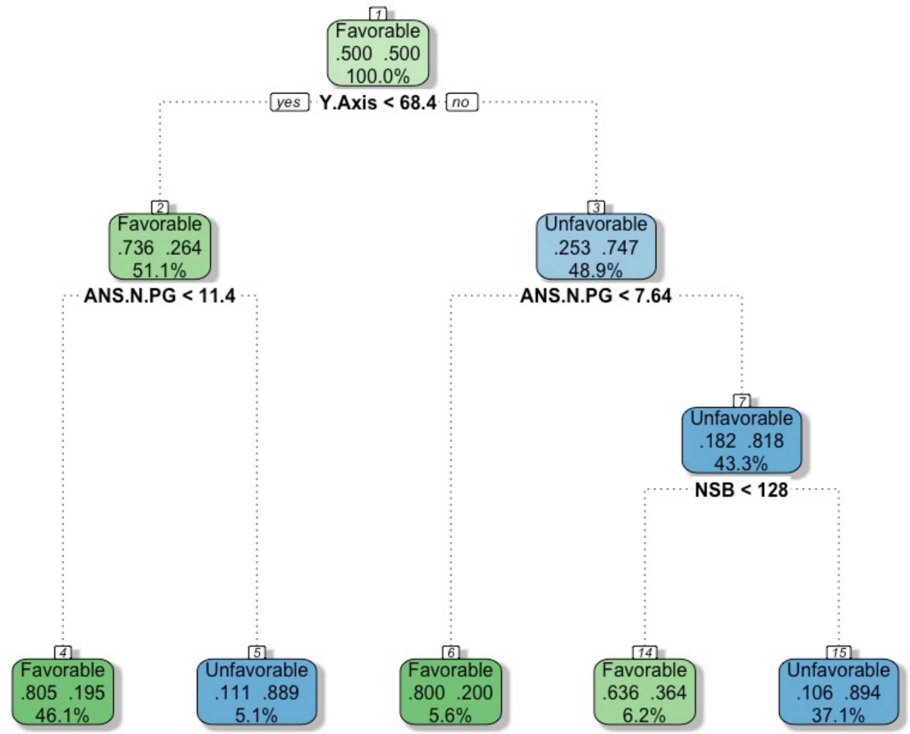

Figure 5. Pruned tree for the 178 longitudinal untreated class I and class II subjects at age 10 to determine the skeletal relationship at age 15.

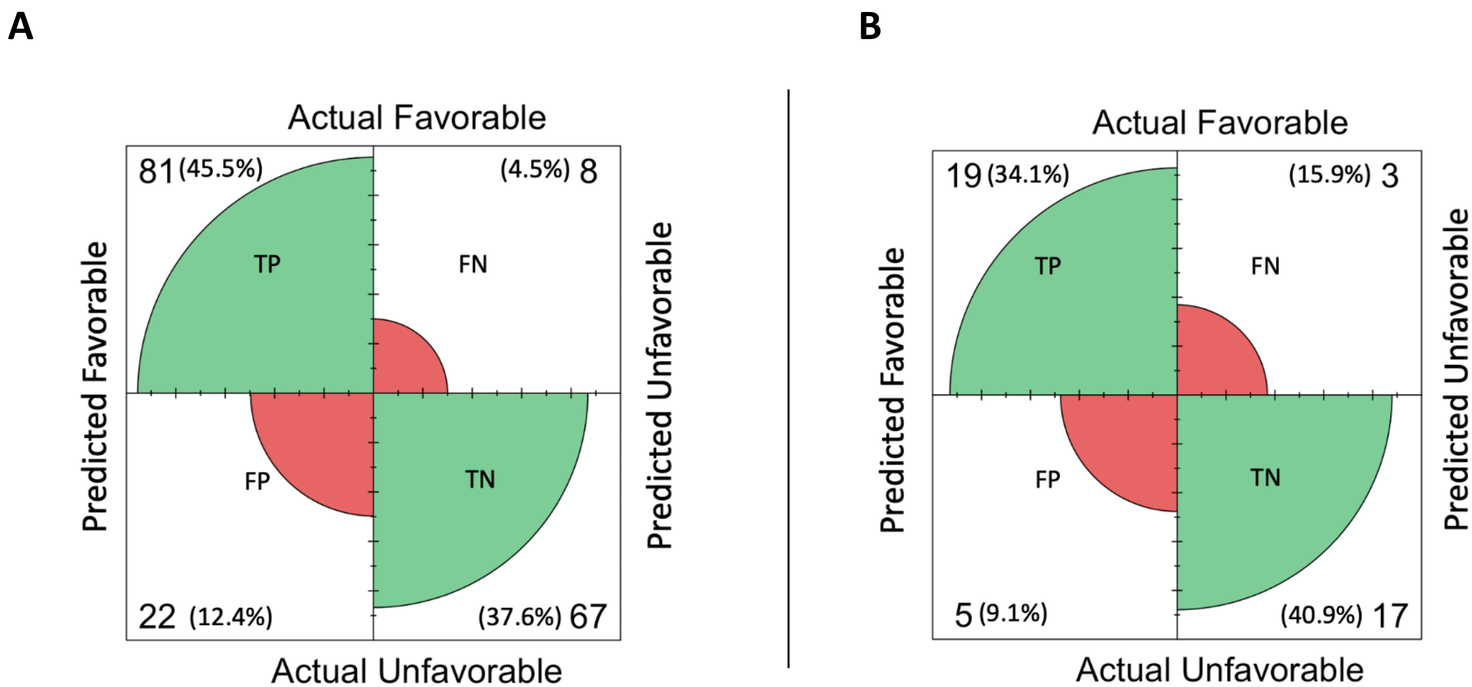

Figure 6. Confusion matrices of the A) training sample, and B) testing sample showing the absolute and the relative (%) number of subjects having true positive (TP), false positive (FP), true negative (TN), and false negative (FN) results of the pruned model.

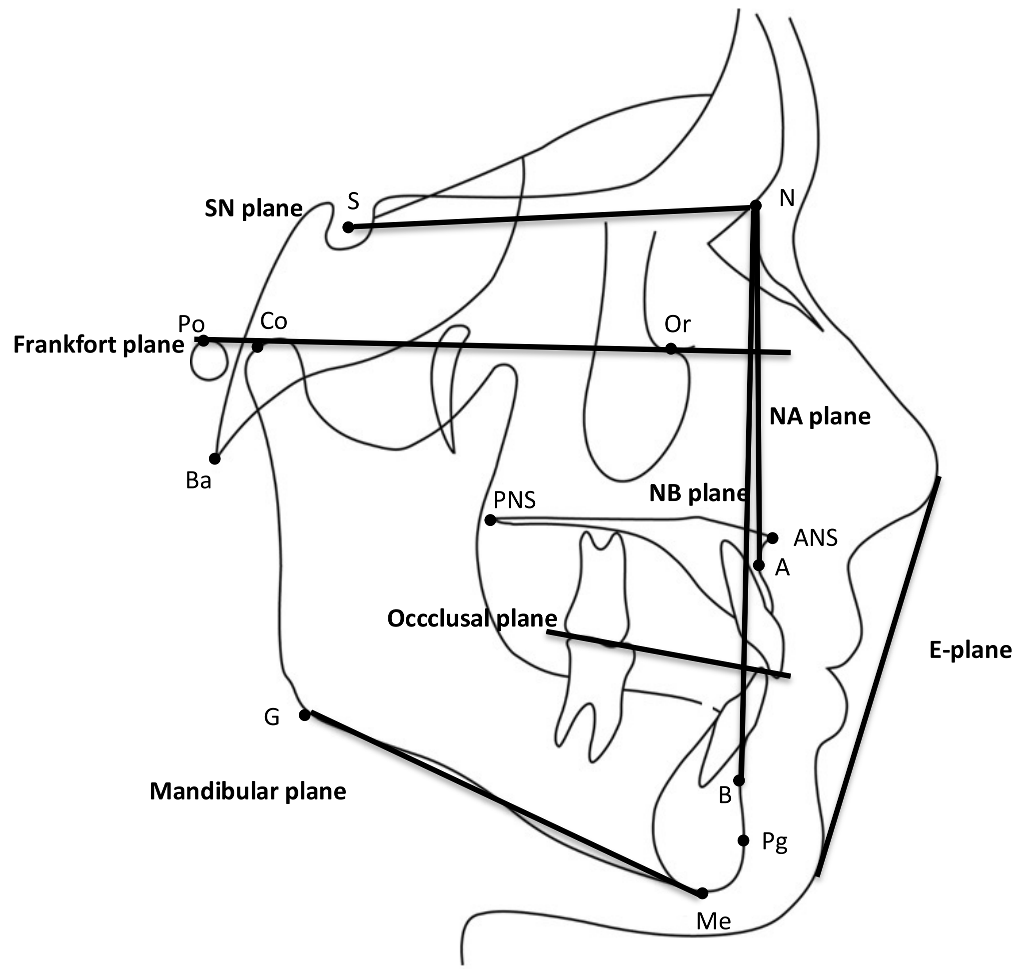

Figure 7. Landmarks and planes used for the growth study

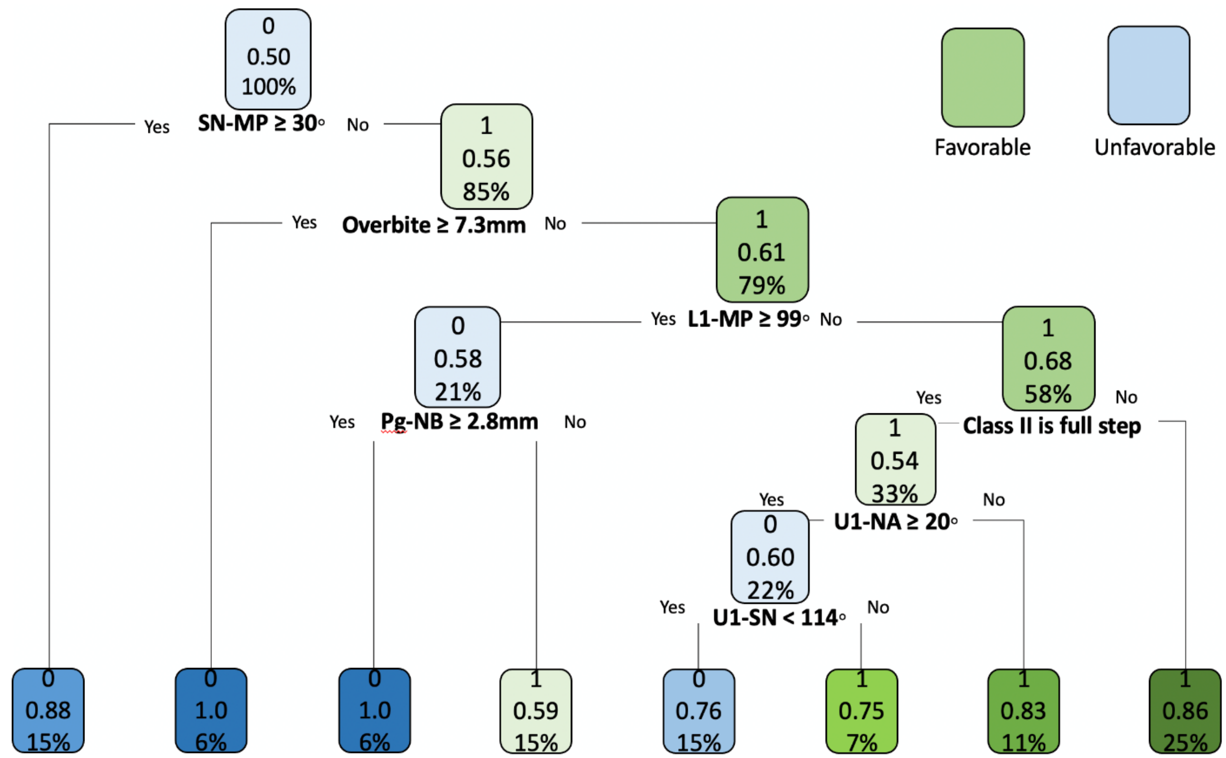

Figure 8. Decision tree for the unpruned model with each node showing the type of outcome on top (0 for unfavorable and 1 for favorable), probability of a favorable outcome in the middle, and the percentage of people satisfying that condition at the bottom.

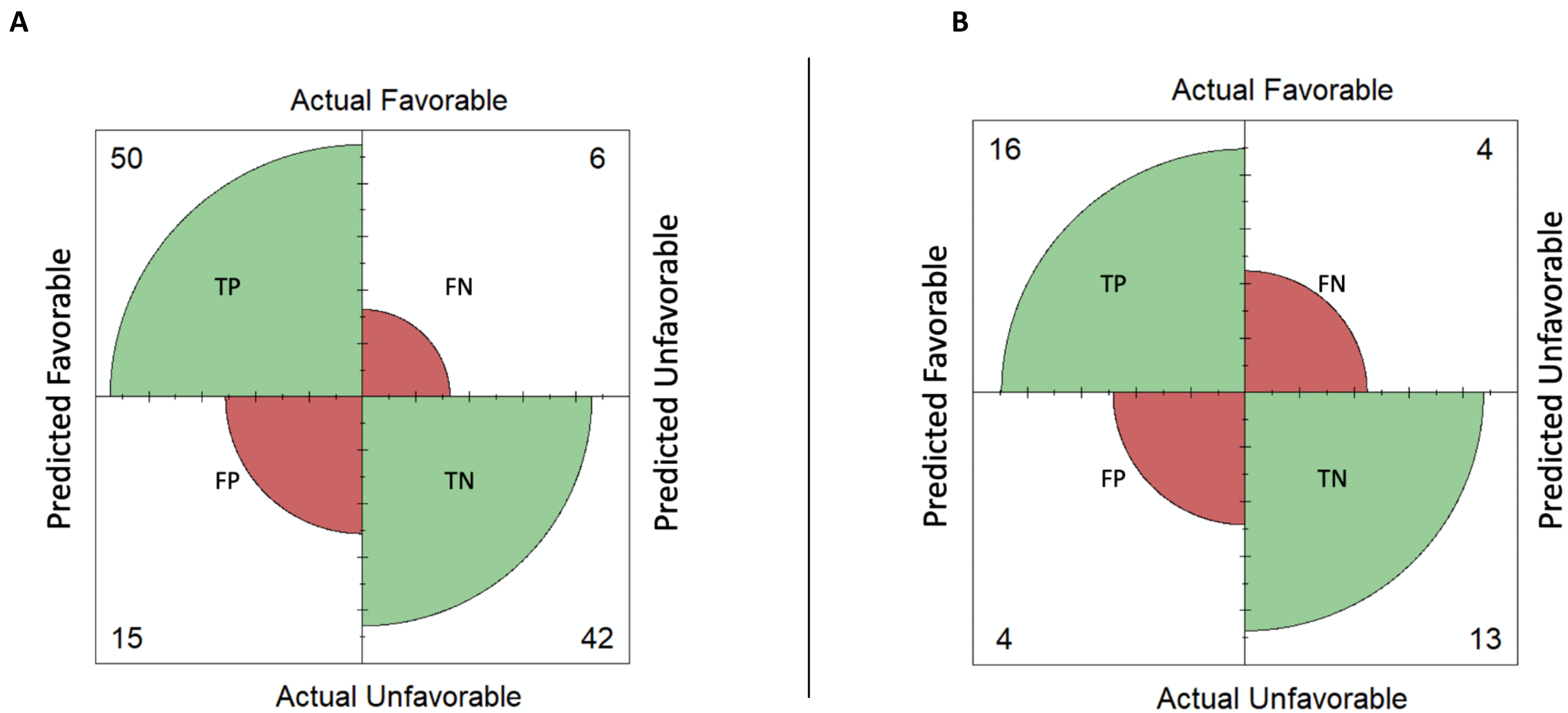

Figure 9. Confusion matrices of the A) training sample, and B) testing sample showing the number of subjects having true positive (TP), false positive (FP), true negative (TN), and false negative (FN) results of the unpruned model.

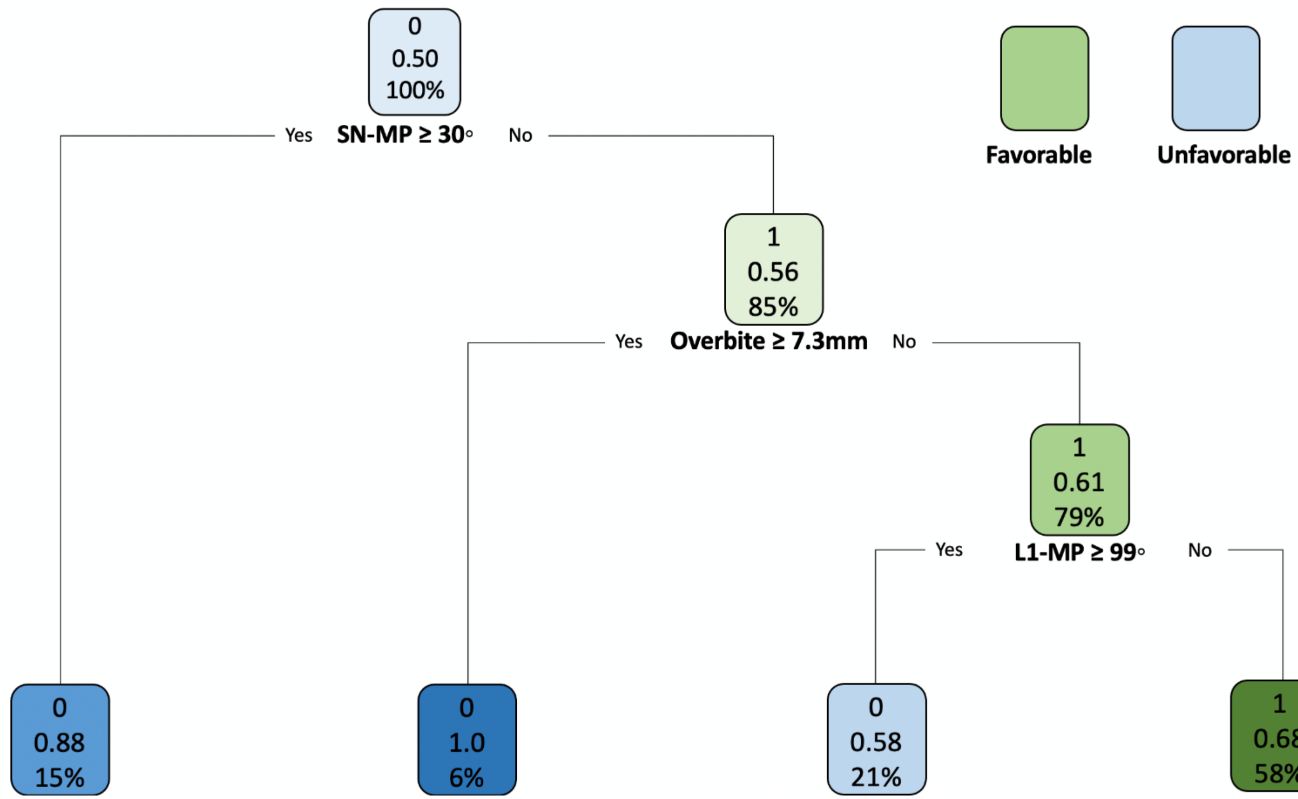

Figure 10. Decision tree for the pruned model with each node showing the type of outcome on top (0 for unfavorable and 1 for favorable), probability of a favorable outcome in the middle, and the percentage of people satisfying that condition at the bottom.

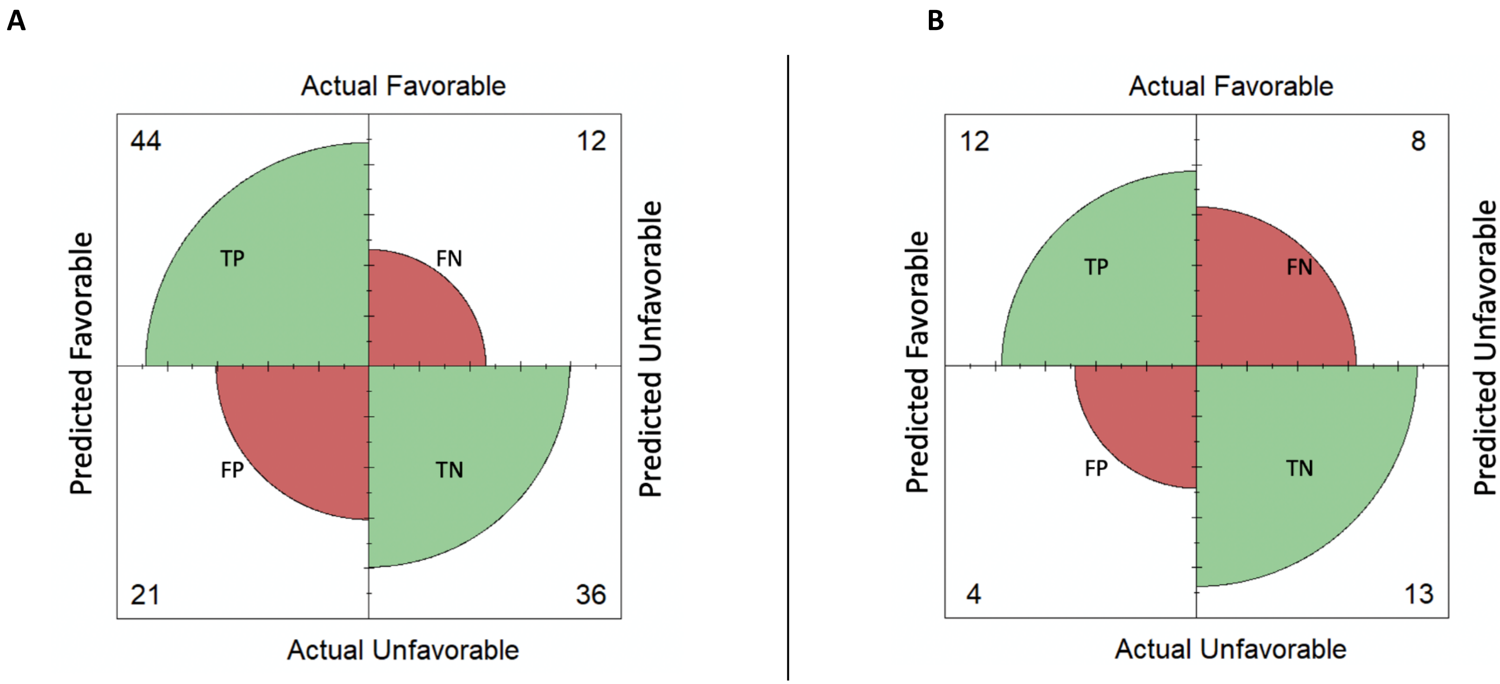

Figure 11. Confusion matrices of the A) training sample, and B) testing sample showing the number of subjects having true positive (TP), false positive (FP), true negative (TN), and false negative (FN) results of the pruned model.

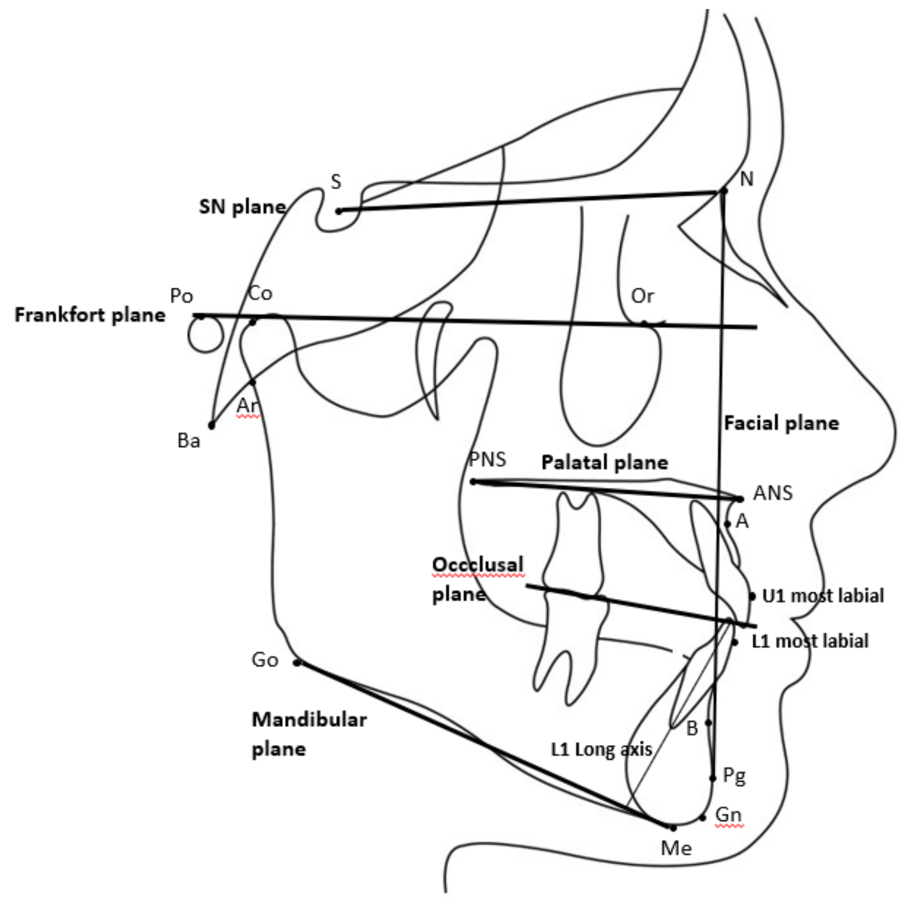

Figure 12. Landmarks & planes used to locate the variables used in the skeletal outcome study

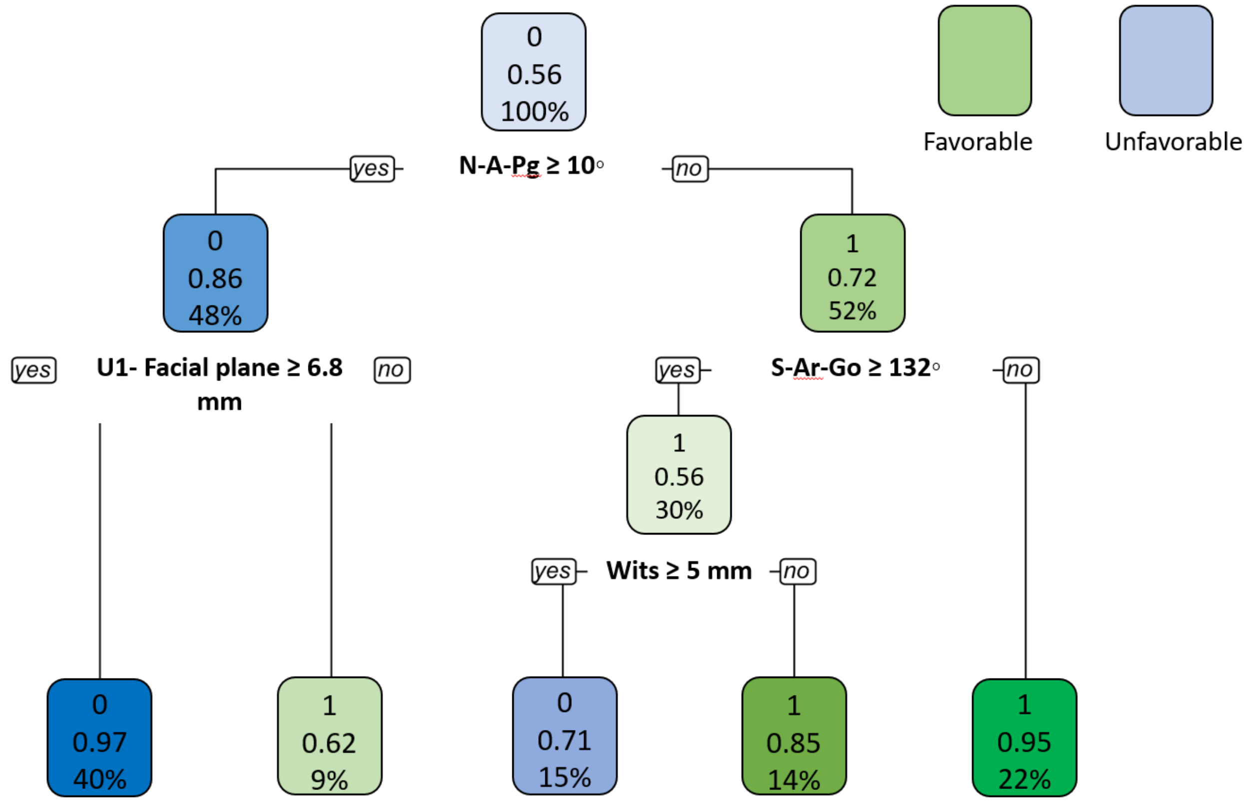

Figure 13. Decision tree for the unpruned model, with each node showing the type of outcome on top (0 for unfavorable and 1 for favorable), the probability of a favorable or unfavorable outcome in the middle, and the percentage of patients satisfying that condition at the bottom.

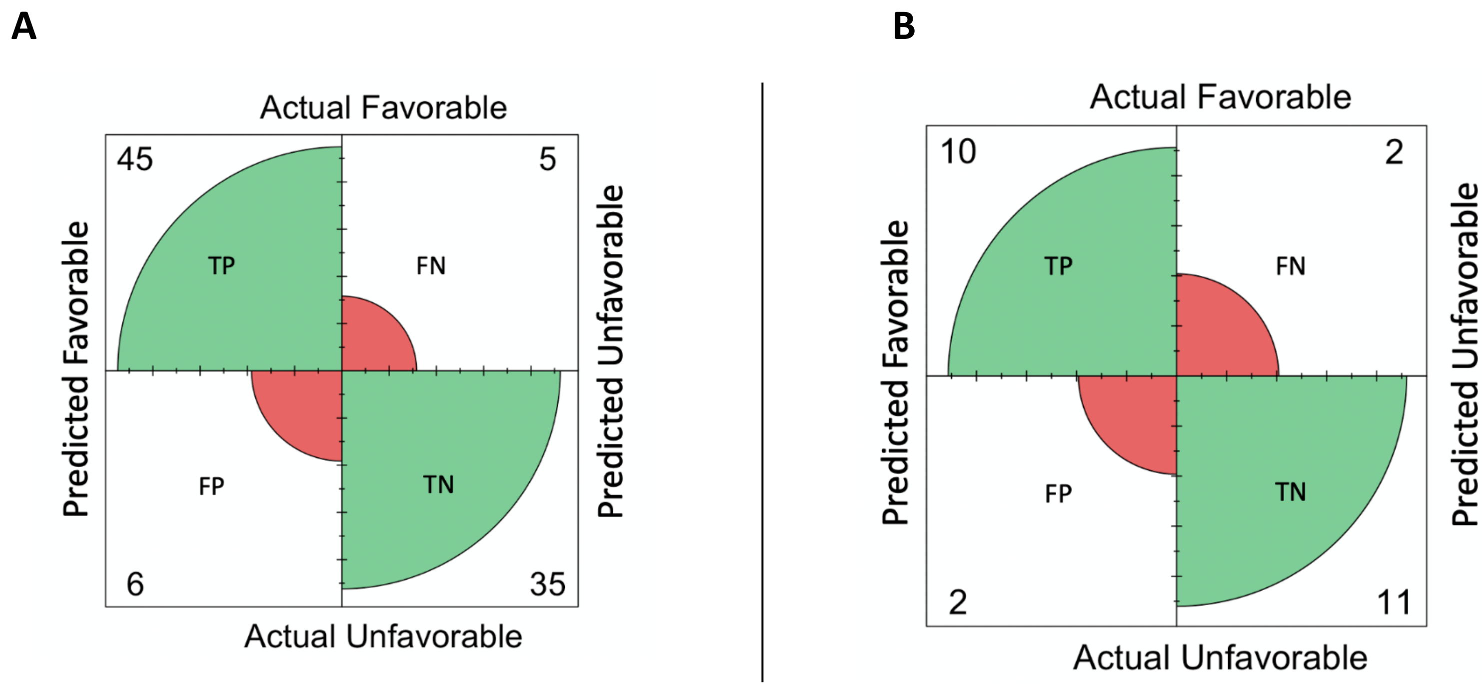

Figure 14. Confusion matrices of the A) training sample, and B) testing sample showing the number of subjects having true positive (TP), false positive (FP), true negative (TN), and false negative (FN) results of the unpruned model.

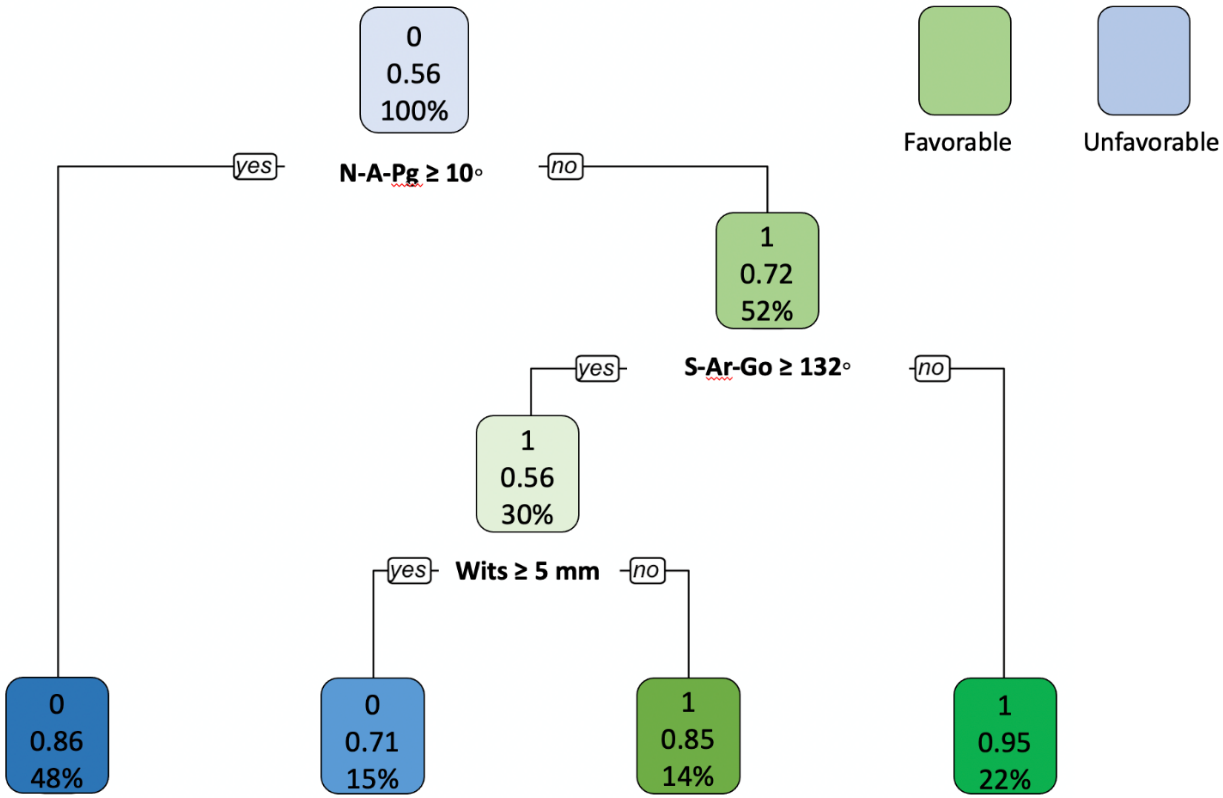

Figure 15. Decision tree for the unpruned model, with each node showing the type of outcome on top (0 for unfavorable and 1 for favorable), the probability of a favorable or unfavorable outcome in the middle, and the percentage of patients satisfying that condition at the bottom.

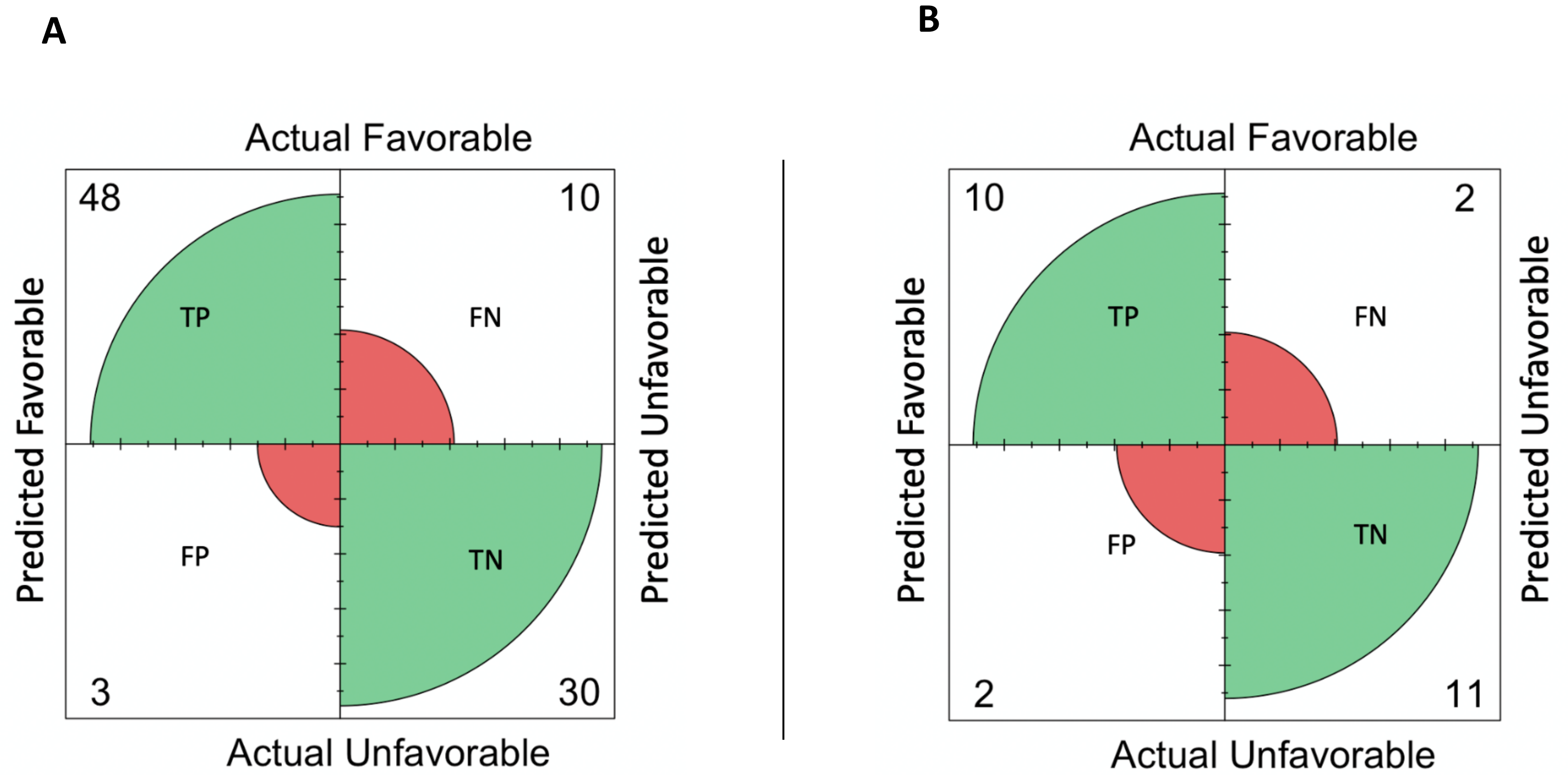

Figure 16. Confusion matrices of the A) training sample, and B) testing sample showing the number of subjects having true positive (TP), false positive (FP), true negative (TN), and false negative (FN) results of the pruned model.

A

Pretreatment measurement	
ANB (◦)	6.3
SN-MP (◦)	23.9
N-A-Pg (◦)	8.2
S-Ar-Go (mm)	131.8

B

Posttreatment measurement	
ANB (◦)	3.9
SN-MP (◦)	24.1

Figure 17. Case #1: Pretreatment (A) and posttreatment (B) lateral cephalometric radiographs and measurements of a patient predicted to have favorable skeletal outcomes based on the facial convexity angle (N-A-Pg) and the articular angle (S-Ar-Go).

A

Pretreatment measurement	
ANB (◦)	5.1
SN-MP (◦)	27.8
N-A-Pg (◦)	7
S-Ar-Go (mm)	126.8

B

Posttreatment measurement	
ANB (◦)	2.8
SN-MP (◦)	27.6

Figure 18. Case #2: Pretreatment (A) and posttreatment (B) lateral cephalometric radiographs and measurements of a patient predicted to have favorable skeletal outcomes based on the facial convexity angle (N-A-Pg) and the articular angle (S-Ar-Go).

A

Pretreatment measurement	
ANB (◦)	6.3
SN-MP (◦)	23.1
N-A-Pg (◦)	10.8
U1-NPg (mm)	13.3

B

Posttreatment measurement	
ANB (◦)	4.3
SN-MP (◦)	24.2

Figure 19. Case #3: Pretreatment (A) and posttreatment (B) lateral cephalometric radiographs and measurements of a patient predicted to have unfavorable skeletal outcomes based on the facial convexity angle (N-A-Pg) and the U1 to facial plane (U1-NPg).

A

Pretreatment measurement	
ANB (◦)	6.1
SN-MP (◦)	28.9
N-A-Pg (◦)	11.8
U1-NPg (mm)	12.2

B

Posttreatment measurement	
ANB (◦)	5.5
SN-MP (◦)	31.7

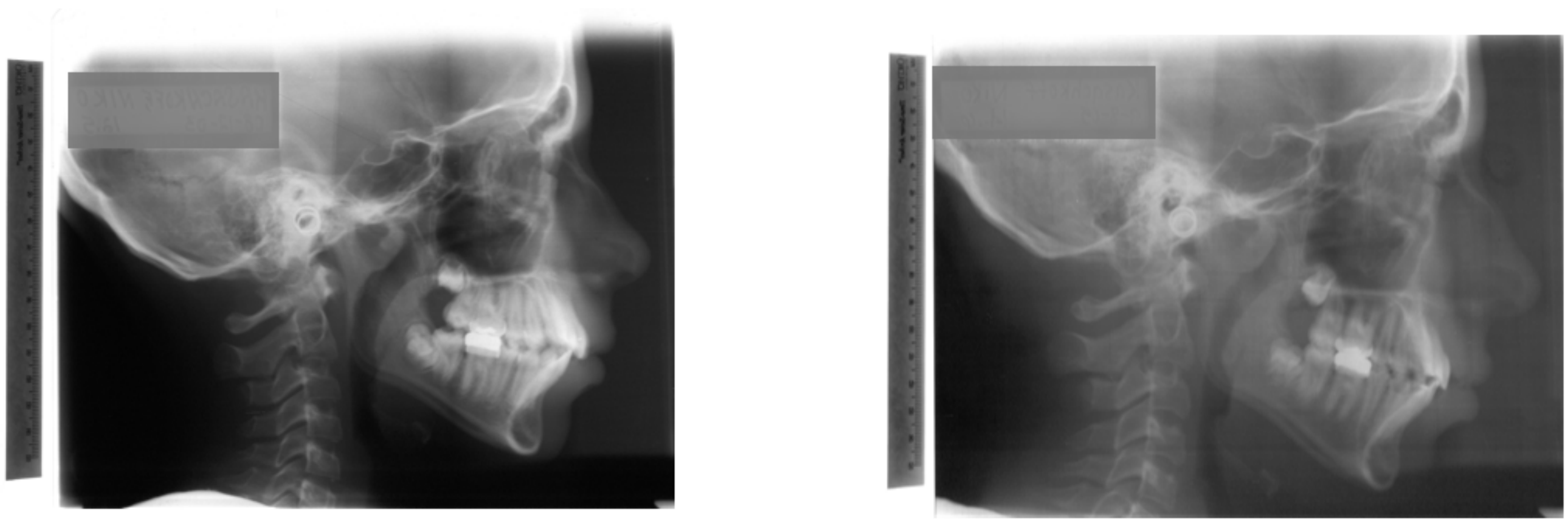

Figure 20. Case #4: Pretreatment (A) and posttreatment (B) lateral cephalometric radiographs and measurements of a patient predicted to have unfavorable skeletal outcomes based on the facial convexity angle (N-A-Pg) and the U1 to facial plane (U1-NPg).

APPENDIX B

TABLES

Table 1. Summary of major machine learning algorithms applied in orthodontics. Reprinted with permission from, “Applications of artificial intelligence and machine learning in orthodontics”, by Asiri SN, Tadlock LP, Schneiderman E, Buschang PH. APOS Trends in Orthodontics, 2020;10(1):17-24. Copyright [2019] by Scientific Scholar on behalf of APOS Trends in Orthodontics.

Machine learning algorithm	**Uses/applications**	**Pros**	**Cons**
Decision trees	Used mainly for classification Applied in medical diagnosis [61] and manufacturing monitoring [62]	Simple and easy to understand even by non-experts [63] They are non-parametric and can handle both nominal and numeric input attributes [63] Can be used when data are missing, skewed, or have errors [64] Order of training instances is not important [65] Pruning reduces overfitting and improves prediction accuracy [65] Order of training has no effect on training [65]	Most algorithms require the target attribute to have only discrete values [63] They perform poorly when many complex interactions exist [63] Oversensitivity to the training set, irrelevant attributes and to noise [66]
Naïve Bayes	Used mainly for classification Applied in medicine [67,68] and dentistry [69,70] for decision support and risk assessment	Simple and easy to understand [71] Order of training has no effect on training [71] It is based on statistical modeling [71] Requires small amount of data for training [[72] Fast and can deal with discrete and continuous attributes [72] Robust to outliers [73]	Accuracy is affected by redundant attributes and class frequency [71] Normal distribution is assumed for numeric attributes [71] Attributes are assumed to be conditionally independent [71]
Neural network	Used for classification and regression. Applied in dentistry and medicine for diagnosis [37]	Boolean functions (AND, OR, and NOT) can be used with neural networks Can handle noisy inputs and allows changing input features during data collection [74] Successful with complex non-linear relationships between predicted variable and input data [74]	Overfitting is common especially with too many variables [75] Have limited ability to identify causal relationship [74] Require more computational resources [74]

Table 1. Continued.

Support vector machine	Used for classification and regression Applied in dentistry for classification of skeletal patterns [56]	Resistant to overfitting [10] Can model nonlinear functions [10] Can be used with non-linear relationships between predicted variable and input data	Training is slow Structure of algorithm is difficult to understand
Genetic algorithm	Used for search and optimization problems Applied in dentistry and medicine mainly for prediction	Simple algorithm and easy to apply [76] Always try to find the best solution	Not efficient for finding the best solution There are complications in representing training and output data
Fuzzy logic	Concerned with finding the truth by approximate modes of reasoning rather than exact reasoning [77] Used to deal with imprecision and uncertainty present in many fields including medicine [78]	Mimics human thinking and can be written in a form similar to natural language [79] Allows for the degree of belonging to either 0 or 1, with 1 representing complete membership and 0 for non-membership Can use both numerical variables and linguistic variables [80]	Requires a lot of data and expertise to develop [81] Analysis is difficult because fuzzy outputs can be interpreted in different ways [81]

Table 2. Landmark and measurement definitions and abbreviations.

Name	Definition	Abbreviation
	Landmarks	
Anterior Nasal Spine	Most anterior point of the maxilla	ANS
B Point	Point of deepest curvature between infradentale and pogonion	B
Basion	Midpoint of the anterior margin of the foramen magnum	Ba
C Point	Point of deepest curvature of the lingual portion of the mandibular symphysis	C
Condylion	Most superior point of the mandibular condyle	Co
Gonion	Midpoint of the angle of the mandible, defined by bisection of the angle formed by the tangents to the posterior border of the ramus and the inferior border of the mandible	Go
Infradentale	The intersection point of the anterior lower incisor and the crestal bone	Id
Menton	The most inferior point of the mandibular symphisis	Me
Nasion	Junction of the frontonasal suture at the most posterior point on the curve at the bridge of the nose	N
Pogonion	Most anterior point of the bony chin	Pg
Sella	Center of the sella turcica of the sphenoid bone by inspection	S

Table 2. Continued.

Measurements		
Mandibular plane angle	Angle formed by the intersection of line Go-Me with line S-N	MPA
Y-axis	Angle formed by the intersection of line S-Gn and S-N	Y-Axis
Posterior to anterior face height	Ratio of the distance from S to Go divided by the distance from N to Me	PAFH
ANS-N-Pg	Angle formed between the points ANS, N, and Pg	ANS-N-Pg
Condylar Inclination	Angle formed between the line Go-S and S-N	CondInc
Gonial Angle	Angle formed between Ar, Go, and Me	GonAng
Symphysial Ratio	Ratio of the distance from C to Pg divided by the distance from Id to Me	HVSym
Symphysial Angle	Angle formed between Id, B, and Pg	SymA
Palatal Plane Angle	Angle formed between the line ANS-PNS and S-N	PPA
Cranial Base Angle	Angle formed between N, S, and Ba	NSBa

Table 3. The mean and the standard deviation (SD) for the Horizontal relationship between ANS and Pg (mm) in males & females at age 10 and 15.

	Male						Female					
	Overall		Favorable		Unfavorable		Overall		Favorable		Unfavorable	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ANSPgh at 10 years old	12.90 (N= 116)	4.05	9.70 (N=55)	2.49	16.10 (N=61)	2.34	13.50 (N= 106)	3.81	10.40 (N=54)	2.25	16.10 (N=53)	2.75
ANSPgh at 15 years old	12.70 (N=116)	5.03	9.2 (N=66)	3.15	17.40 (N=52)	2.70	14.70 (N= 106)	5.05	11.00 (N=66)	3.51	18.20 (N=40)	3.33

Table 4. Model performance for the unpruned & pruned trees on both the training and testing data sets.

	Unpruned tree		Pruned tree	
Data / Statistics	Training	Test	Training	Test
Accuracy %	85.39%	75.0%	83.15%	81.82%
Kappa	0.71	0.5	0.66	0.64
Sensitivity	0.85	0.68	0.91	0.86
Specificity	0.85	0.82	0.75	0.77
Positive predictive value	0.85	0.79	0.79	0.79
Negative predictive value	0.85	0.72	0.89	0.85

Table 5. Several Herbst studies reporting the correction of dental relationships and overjet

Reference	N	Molar relationship	Canine relationship	Overjet
Burkhardt et al[48]	30	Corrected	N/A	Corrected
Valant and Sinclair[16]	32	Corrected	N/A	N/A
Pancherz and Hensen[49]	40	Corrected	Corrected	Corrected
Pancherz[14]	22	Corrected	Corrected	Corrected
LaHaye et al[4]	19	Corrected	Corrected	Corrected
Wigal et al[39]	22	Corrected	N/A	Corrected
de Almeida et al[50]	30	Corrected	Corrected	Corrected
McNamara et al[51]	45	Corrected	N/A	Corrected
Sidhu et al[52]	8	Corrected	Corrected	Corrected

Table 6. Pretreatment predictor variables used to build the model

Variables	Units	Descriptions
S-N-A	◦	Angle formed between the landmarks sella, nasion and point A, indicating antero-posterior position of maxilla with respect to anterior cranial base
S-N-B	◦	Angle formed between the landmarks sella, nasion and point B, indicating antero-posterior position of mandible with respect to anterior cranial base
AO-BO	mm	Wits: the distance between the perpendiculars from landmarks A and B on the maxilla and mandible, respectively, onto the occlusal plane.
SN-MP	◦	Angle formed by the intersection of Go-Me plane with the S-N plane
OJ	mm	Overjet: the distance between maxillary incisor most labial and mandibular incisor edge parallel to occlusal plane
OB	mm	Overbite: the distance between maxillary incisor edge and mandibular incisor edge perpendicular to occlusal plane
U1-SN	◦	The axial inclination of the most labial maxillary central incisor in relation to the cranial base.
U1-NA	◦	The angle formed by the long axis of the upper incisor to a line from nasion to point A
L1-NB	mm	The perpendicular distance from lower incisor to NB plane
Pg-NB	mm	The perpendicular distance from Pg to NB plane
U1-L1	◦	The interincisal angle which is formed by the intersection of the long axis of the maxillary & mandibular incisors.
L1-MP	◦	Angle formed by intersection of mandibular incisor to mandibular plane
FMIA	◦	Angle formed by extending mandibular incisor long axis to the Frankfort horizontal plane (Po-Or)
CVS	NA	Stages of Cervical Vertebral Maturation[1]
Unilateral/bilateral Class II	NA	Whether the Class II molar relationship is unilateral or bilateral
Severity of Class II molar	NA	Whether the Class II molar relationship is full step or less than that
Posterior crossbite	NA	Whether or not there was a posterior crossbite
UL- E plane	mm	The perpendicular distance from the most anterior portion on the margin of the upper lip to a line drawn tangent to the tip of the nose and the soft tissue chin (E-plane)
LL- E plane	mm	The perpendicular distance from the most anterior portion on the margin of the lower lip to a line drawn tangent to the tip of the nose and the soft tissue chin (E-plane)
N-A-Pg	◦	Soft tissue convexity which is an acute angle formed by the intersection of NA and A-Pg planes

Table 7. Frequencies (%) of dental outcomes of subjects classified as having unfavorable outcomes.

	Overjet	Canines	Molars
Favorable outcome	58 (78.4%)	4 (5.4%)	22 (29.7%)
Unfavorable outcome	16 (21.6%)	70 (94.6%)	52 (70.3%)
Total	74 (100%)		

Table 8. Cervical vertebra stages (CVS), type and severity of Class II malocclusion and the presence or absence of posterior crossbite among subjects with favorable & unfavorable treatment outcomes (%).

Subjects with favorable outcome (50.7%)						
CVS	1 (14.5%)	2 (10.5%)	3 (17.1%)	4 (43.4%)	5 (11.8%)	6 (2.6%)
Class II	Unilateral (11.8%)			Bilateral (88.2%)		
Severity of Class II molar*	End on (51.3%)			Full step (48.7%)		
Posterior crossbite*	No (92.1%)			Yes (7.9%)		
Subjects with unfavorable outcome (49.3%)						
CVS stage	1 (12.2%)	2 (16.2%)	3 (18.9%)	4 (36.5%)	5 (13.5%)	6 (2.7%)
Class II	Unilateral (9.4%)			Bilateral (90.6%)		
Severity of Class II molar*	End on (33.8%)			Full step (66.2%)		
Posterior crossbite*	No (79.7%)			Yes (20.3%)		

*Significant differences where the unfavorable group had significantly higher frequencies of full step class II molar relationship and posterior crossbite.

Table 9. Proportion of favorable & unfavorable dental outcomes among 3 dental sites where records were collected, with A and B representing the private practices.

		Dental site			Total
		A	B	C	
Favorable	Count	38	17	21	76
	% within Dental office	46.3%	44.7%	70.0%	50.7%
	% of Total	25.3%	11.3%	14.0%	50.7%
Unfavorable	Count	44	21	9	74
	% within Dental office	53.7%	55.3%	30.0%	49.3%
	% of Total	29.3%	14.0%	6.0%	49.3%
Total	Count	82	38	30	150
	% within Dental office	100.0%	100.0%	100.0%	100.0%
	% of Total	54.7%	25.3%	20.0%	100.0%

- P value of 0.06 showing no significance in term of the frequencies of favorable outcomes among the 3 dental sites.

Table 10. Pretreatment differences between groups have favorable and unfavorable dental outcomes

	Favorable (N=76)		Unfavorable (N=74)		
Variable	**Mean**	**SD**	**Mean**	**SD**	Prob
S-N-A (°)	81.8	3.6	80.6	3.7	0.05
S-N-B (°)	76.9	3.5	75.4	3.4	**0.009***
AO-BO (mm)	4.1	2.5	5.0	2.8	0.042
SN-MP (°)	24.5	4.3	26.1	6.1	0.077
OJ (mm)	7.1	2.7	7.7	2.7	0.227
OB (mm)	4.3	1.5	4.2	2.2	0.694
U1-SN (°)	102.9	11.5	102.1	10.9	0.661
U1-NA (°)	21.1	11.7	21.6	10.7	0.787
L1-NB (mm)	3.7	2.3	3.6	2.3	0.728
Pg- NB (mm)	2.3	1.7	2.5	1.9	0.392
Interincisal angle (°)	132.7	13.8	132.5	14.3	0.928
L1-MP (°)	92.8	6.6	92.3	7.7	0.656
FMIA (°)	66.1	8.1	65.6	7.8	0.724
UL-E plane (°)	-1.3	2.2	-1.6	2.6	0.368
LL-E plane (°)	-0.7	2.7	-1.2	2.8	0.321
N-A-Pg (°)	128.2	5.6	127.6	4.7	0.447

*Significant differences were found only in SNB angle after adjustment for multiple comparisons.

Table 11. Mean change & standard deviation (SD) of subjects with favorable & unfavorable dental treatment outcomes.

	Favorable dental outcome (N=76)		Unfavorable dental outcome (N=74)		
Variable	**Mean**	**SD**	**Mean**	**SD**	**Prob**
S-N-A (°)	-0.9	1.7	-0.9	1.6	.887
S-N-B (°)	1.1	1.8	1.0	1.5	.650
AO-BO (mm)	-4.3	2.2	-4.3	2.5	.951
SN-MP (°)	-.53	2.2	-0.1	2.5	.274
OJ (mm)	-4.3	2.7	-4.3	2.7	.968
OB (mm)	-2.9	1.4	-2.6	2.1	.435
U1-SN (°)	3.5	11.7	1.9	11.8	.409
U1-NA (°)	4.3	11.9	2.7	11.7	.407
L1-NB (mm)	1.4	1.9	1.2	2.2	.558
Pg to NB (mm)	0.4	1.0	0.3	1.0	.793
Interincisal angle (°)	-8.3	14.7	-6.5	15.1	.471
L1-MP (°)	5.5	7.1	4.7	7.1	.527
FMIA (°)	-6.1	8.0	-5.1	7.4	.427
UL- E plane (°)	-2.6	2.0	-2.5	2.1	.888
LL- E plane (°)	-0.9	2.2	-0.9	2.9	.840
N-A-Pg (°)	-0.1	3.0	-0.1	2.9	.860

- No significant differences were found after adjusting for multiple comparisons.

Table 12. Model performance of the unpruned & pruned trees for both the training and testing data sets for the dental outcome.

	Unpruned tree		**Pruned tree**	
Data Statistics	Training	Testing	Training	Testing
Accuracy %	81.4%	78.4%	70.8%	67.6%
Kappa	0.63	0.56	0.42	0.36
Sensitivity	0.74	0.76	0.63	0.76
Specificity	0.89	0.80	0.78	0.60
Positive predictive value	0.87	0.76	0.75	0.62
Negative predictive value	0.77	0.80	0.68	0.75

Table 13. Means and standards deviations (SD) for the subjects at each of the 8 terminal nodes.

	(A) SN-MP ≥ 30		(B) Overbite ≥ 7.3mm		(C) Pg-NB ≥ 2.8mm		(D) Pg-NB < 2.8mm		(E) U1-SN < 114◦		(F) U1-SN ≥ 114◦		(G) U1-NA < 20◦		(H) Class II is full step	
Variable	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
SNA (°)	78.9	3.4	83.2	3.5	82.8	2.1	81.5	3.3	81.4	3.4	81.2	2.9	82.3	4.4	81.7	4.7
SNB (°)	72.8	2.5	77.3	2.5	77.6	2.1	76.4	3.1	76.9	2.1	76.8	4.5	77.3	4.5	77.6	3.9
Wits	4.9	3.1	5.2	2.4	5.3	1.9	4.4	1.9	4.6	3.1	6.1	3.9	3.9	2.6	3.8	1.9
SN-MP (°)	33.1	3.1	20.2	5.4	21.5	2.0	23.6	3.8	23.6	3.6	23.4	4.1	23.0	3.6	23.5	4.2
Overjet (mm)	7.4	3.2	7.0	0.8	6.4	2.1	6.5	2.1	8.1	2.2	10.5	3.5	5.9	1.3	7.2	2.3
Overbite (mm)	3.1	1.9	8.5	1.2	4	1.1	3.9	1.3	3.7	1.6	4.2	1.9	5.0	1.2	4.3	1.6
U1-SN (°)	98.2	7.5	86.2	12.2	105.2	5.6	106.9	8.0	106.1	4.3	117.8	4.4	94.5	9.7	104.6	12.7
U1-NA (°)	19.3	6.8	4.1	11.9	22.4	5.3	25.4	9.1	24.7	3.5	36.6	3.1	12.1	8.2	22.8	12.4
L1-NB (mm)	5.0	2.6	1.2	2.7	5.1	1.0	5.7	1.0	2.8	1.3	2.8	1.6	2.8	2.0	2.8	1.9
Pg-NB (mm)	1.2	1.7	3.8	2.5	3.8	0.7	1.7	0.9	2.8	1.5	2.7	1.7	3.1	1.2	2.6	1.8
Interincisal angle (°)	131.4	11.5	157.3	18.9	124.2	6.2	120.1	8.8	131.9	6.5	119.3	5.2	144.9	10.6	134.1	12.8
L1-MP (°)	90.2	7.6	89.2	9.3	102.1	2.3	102.4	2.6	91.3	5.2	92.4	4.7	90.6	5.3	90.8	4.9

Table 13. Continued

FMIA (°)	61.9	6.2	73.8	11.6	57.7	2.4	57.3	3.9	68.1	6.3	67.7	5.8	70.1	5.6	69.5	5.9

Table 14. The means at T2 and the changes for ANB and SN-MP in subjects treated with Herbst and had successful dental outcomes

Reference	N	ANB T2	Change in ANB	SN-MP T2	Change in SN-MP	Molar relationship	Canine relationship	Overjet
Burkhardt et al[37]	30	4.1	-1.4	23.2	-0.3	Corrected	N/A	Corrected
Schiavoni et al[51]	19	N/A	-1.8	N/A	-1.1	Corrected	Corrected	Corrected
Valant and Sinclair[38]	32	4.0	-1.9	21.4	0.6	Corrected	N/A	N/A
Pancherz and Hensen[39]	40	4.6	-1.5	31.5	0.4	Corrected	Corrected	Corrected
Pancherz[7]	75	N/A	-2.0	N/A	0.1	Corrected	Corrected	Corrected
Pancherz[8]	22	3.9	-1.9	31.6	0.2	Corrected	Corrected	Corrected
LaHaye et al[40]	19	5.1	-0.7	34.5	0.3	Corrected	Corrected	Corrected
Wigal et al[51]	22	2.6	-2.0	34.3	0.1	Corrected	N/A	Corrected
de Almeida et al[41]	30	5.2	-1.4	33.9	0.1	Corrected	Corrected	Corrected
McNamara et al[42]	45	4	-1.9	23.7	-0.3	Corrected	N/A	Corrected
Sidhu et al[43]	8	4.4	-2.1	N/A	N/A	Corrected	Corrected	Corrected

Table 15. Pretreatment predictor variables used to build the model for skeletal outcomes

Variables	**Units**	**Descriptions**
Age	Years	Subject's age before the start of orthodontic treatment
Sex	N/A	Whether the patient is male or female
CVS	N/A	Stages of Cervical Vertebral Maturation[1]
S-N-A	°	Angle formed between the landmarks sella, nasion and point A, indicating antero-posterior position of maxilla with respect to anterior cranial base
S-N-B	°	Angle formed between the landmarks sella, nasion and point B, indicating antero-posterior position of mandible with respect to anterior cranial base
Wits	mm	The distance between the perpendiculars from landmarks A and B on the maxilla and mandible, respectively, onto the functional occlusal plane.
SN-MP	°	Angle formed by the intersection of Go-Me plane with the S-N plane
PP-MP	°	Inclination of the palatal plane in relation to the mandible plane
Mx-Md Diff	mm	Maxillomandibular differential is the difference between total mandibular length (Co-Gn) and midfacial length (Co-A)
SGo:NMe	N/A	Facial height ratio of the linear measurements from S-Go to N-Me
N-ANS	mm	Upper Anterior Facial Height (UAFH): the distance from Nasion (N) to Menton (Me)
ANS-Me	mm	Lower Anterior Facial Height (LAFH): the distance from Anterior Nasal Spine (ANS) to Menton (Me)
Ar-Go	mm	Lower posterior facial height (LPFH): the distance between Articluare (Ar) and Gonion (Go)

Table 15. Continued

Ar-Go-Me	°	The gonial angle: formed by line connecting Articulare to Gonion to Gnathion
S-Ar-Go	°	The articular angle: formed by line connecting Sella to Articulare to Gonion
Cd-Go	mm	Mandibular ramus height, distance between point condylion and point gonion
Ar-Gn	mm	Mandibular length measured from Ar to Gn
Cd-Gn	mm	Effective mandibular length: the distance from the posterior border of the Cond to Gn
Overjet	mm	The distance between maxillary incisor most labial and mandibular incisor edge parallel to occlusal plane
Overbite	mm	The distance between maxillary incisor edge and mandibular incisor edge perpendicular to occlusal plane
U1- NPg	mm	Linear distance from the most prominent anterior point on the labial surface of the upper incisor to the facial plane (N-Pg)
L1- NPg	mm	Linear distance from the most prominent anterior point on the labial surface of the lower incisor to the facial plane N-Pg plane.
LMIP	°	Angle formed by intersection of mandibular incisor to mandibular plane
FMIA	°	Angle formed by extending mandibular incisor long axis to the Frankfort horizontal plane
N-A-Pg	°	Facial convexity angle which is an acute angle formed by intersection of N-A and A-Pg lines

Table 16. Frequencies (%) of skeletal outcomes of subjects classified as having unfavorable outcomes.

		Vertical		
		Favorable	Unfavorable	Total
AP	Favorable	53	12	65 (55.2%)
	Unfavorable	24	27	51 (44.8%)
	Total	77 (66.4%)	39 (33.6%)	116 (100%)

Table 17. Frequencies of CVS and sex among subjects with favorable & unfavorable skeletal treatment outcomes (%).

Subjects with favorable outcome (45.7%)						
CVS	**1** (17.0%)	**2** (15.1%)	**3** (15.1%)	**4** (35.8%)	**5** (17.0%)	**6** (0.0%)
Sex	Male			Female		
Subjects with unfavorable outcome (54.3%)						
CVS	**1** (9.5%)	**2** (15.9%)	**3** (15.9%)	**4** (46.0%)	**5** (9.5%)	**6** (3.2%)
Sex	Male			Female		

- Both statistically insignificant with prob > .05

Table 18. Proportion of favorable and unfavorable skeletal outcomes among 3 sites where records were collected, with A and B representing the private practices.

		Dental Site			Total
		A	B	C	
Favorable	Count	28	14	11	53
	% within Dental office	50.9%	35.9%	50.0%	45.7%
	% of Total	24.1%	12.1%	9.5%	45.7%
Unfavorable	Count	27	25	11	63
	% within Dental office	49.1%	64.1%	50.0%	54.3%
	% of Total	23.3%	21.6%	9.5%	54.3%
Total	Count	55	39	22	116
	% within Dental office	100.0%	100.0%	100.0%	100.0%
	% of Total	54.7%	25.3%	20.0%	100.0%

- Chi square insignificant with p value of 0.321.

Table 19. Pretreatment differences between groups have favorable and unfavorable skeletal outcomes

		Favorable (N= 53)		Unfavorable (N= 63)		Prob
Variable	**Units**	**Mean**	**SD**	**Mean**	**SD**	
Age	Years	12.3	1.4	12.7	1.9	0.3
S-N-A	°	82.2	2.9	82.2	3.3	0.9
S-N-B	°	77.1	3.2	75.6	3.2	0.013
Wits	mm	4.6	2.2	5.7	2.7	0.016
Mx-Md Diff	mm	20.1	3.9	20.1	2.6	0.971
SN-MP	°	23.6	4.2	27.7	5.3	**<0.001***
PP-MP	°	21.9	5.1	25.5	5.9	**0.001***
S-Go/N-Me	N/A	68.7	3.8	65.8	4.8	**<0.001***
N-ANS	mm	48.4	3.2	49.0	3.8	0.359
ANS-Me	mm	61.4	5.1	65.1	6.2	**0.001***
Ar-Go	mm	50.1	4.6	49.6	5.5	0.615
Ar-Go-Me	°	120.6	5.5	121.6	7.1	0.380
N-A-Pg	°	7.9	2.6	12.5	3.7	**<0.001***
S-Ar-Go	°	133.5	5.6	137.6	6.3	**<0.001***
Co-Go	mm	59.1	4.8	58.9	5.9	0.789
Co-Gn	mm	109.9	6.2	109.0	8.7	0.545
Ar-Gn	mm	104.8	6.0	103.4	8.3	0.314
Overjet	mm	7.0	2.6	7.8	2.8	0.125
Overbite	mm	4.6	1.4	3.8	2.5	0.040
LIMP	°	93.5	6.8	93.2	7.7	0.841
FMIA	°	65.9	7.4	62.2	7.5	0.010
U1- NPg	mm	8.0	3.3	10.9	3.6	**<0.001***
L1- NPg	mm	1.7	2.5	3.7	3.4	**<0.001***

- *Significantly different after adjustment for multiple comparisons.

Table 20. Mean change & standard deviation (SD) of subjects with favorable & unfavorable skeletal treatment outcomes.

		Favorable (N= 53)		**Unfavorable (N= 63)**		Prob
Variable	**Units**	**Mean**	**SD**	**Mean**	**SD**	
S-N-A	°	-1.0	1.4	-1.7	1.6	0.008
S-N-B	°	1.6	1.4	0.1	1.6	**<0.001***
Wits	mm	-4.2	1.9	-4.9	2.6	0.139
Mx-Md Diff	mm	5.0	2.1	4.1	2.6	0.051
SN-MP	°	-1.7	1.8	1.5	2.1	**<0.001***
PP-MP	°	-2.2	2.6	0.8	2.3	**<0.001***
S-Go/N-Me	N/A	2.2	1.8	-0.8	5.4	**<0.001***
N-ANS	mm	1.8	2.1	2.2	2.3	0.378
ANS-Me	mm	0.5	10.5	3.2	2.8	0.049
Ar-Go	mm	4.8	2.9	3.1	2.8	**0.002***
Ar-Go-Me	°	-0.9	2.8	0.5	2.1	0.003
N-A-Pg	°	-6.4	3.1	-4.1	3.2	**<0.001***
S-Ar-Go	°	-0.4	4.2	0.5	3.2	0.218
Co-Go	mm	4.6	2.9	3.2	3.1	0.014
Co-Gn	mm	5.6	3.9	4.8	4.4	0.332
Ar-Gn	mm	5.8	4.1	3.1	13.2	0.148
Overjet	mm	-4.0	2.6	-4.7	2.7	0.135
Overbite	mm	-3.0	1.4	-2.4	2.4	0.091
LIMP	°	4.2	7.7	4.5	6.8	0.795
FMIA	°	-3.2	8.3	-6.2	7.4	0.037
U1- NPg	mm	-3.2	3.3	-2.9	3.3	0.726
L1- NPg	mm	0.4	2.3	1.4	2.2	0.016

- *Significantly different after adjustment for multiple comparisons.

Table 21. Model performance of the unpruned & pruned trees for both the training and testing data sets.

	Unpruned tree		**Pruned tree**	
Data / Statistics	Training	Test	Training	Test
Accuracy %	87.9%	84.0%	85.7%	84.0%
Kappa	0.75	0.68	0.70	0.68
Sensitivity	0.88	0.83	0.94	0.83
Specificity	0.87	0.85	0.75	0.85
Positive predictive value	0.90	0.83	0.83	0.83
Negative predictive value	0.85	0.85	0.91	0.85