European Journal of Orthodontics, 2020, 200–205
doi:10.1093/ejo/cjz031
Advance Access publication 8 May 2019

OXFORD

Original article

Development and validation of a prediction model for long-term unsuccess of early treatment of Class III malocclusion

Bernardo Quiroga Souki[1], Michele Nieri[2], Chiara Pavoni[3], Helena Maria Pavan Barros[1], Tarcisio Junqueira Pereira[1], Veronica Giuntini[2], Paola Cozza[3] and Lorenzo Franchi[2]

[1]Graduate Program in Orthodontics, Pontifical Catholic University of Minas Gerais, Belo Horizonte, Brazil, [2]Department of Experimental and Clinical Medicine, University of Florence and [3]Department of Clinical Sciences and Translational Medicine, University of Rome Tor Vergata, Italy

Correspondence to: Lorenzo Franchi Department of Experimental and Clinical Medicine, Section of Dentistry, Orthodontics, The University of Florence, Via del Ponte di Mezzo, 46-48 50127 Firenze, Italy. E-mail: lorenzo.franchi@unifi.it

Summary

Aim: To develop and validate a prediction model to forecast long-term stability of early treatment with rapid maxillary expansion (RME) and facemask (FM) in a large sample of Class III growing patients.

Methods: The Brazilian Group (BG) consisted of 73 consecutively treated Caucasian Class III patients (41 females and 32 males). Mean age at T0 (before treatment) was 7.1 ± 1.6 years, while mean age at T1 (long-term follow-up) was 21.8 ± 3.2 years. The Italian Group (IG, validation cohort) comprised 28 consecutively treated Caucasian Class III patients (14 females and 14 males, mean age at T0 9.0 ± 1.3 years and mean age at T1 18.2 ± 1.4 years). Cephalometric analysis was performed on lateral cephalograms at T0. Gender and cephalometric variables, chronologic age, and dentition phase at T0 were used as predictors for long-term unsuccessful treatment at T1. All predictors for unsuccessful treatment in the BG were subjected to bivariate logistic regression. Only those statistically significant predictors in the bivariate logistic regression entered mixed stepwise logistic regression with $P = 0.05$ to enter and to leave. The validity of the prediction model derived from the BG was then tested on the IG.

Results: The prediction model consisted of only one cephalometric variable: the angle between the Condylar Axis and the Mandibular Plane (CondAx–MP) (odds ratio: 1.52, 95% confidence interval: 1.25–1.85, $P < 0.0001$). Unsuccessful treatment at T1 was predicted for values of CondAx–MP at T0 greater than the cut-off value of 147.8 degrees. BG patients predicted incorrectly were 3 out of 22 for the unsuccessful cases and 1 out of 51 for the successful cases. Therefore, accuracy was 0.95, sensitivity 0.86, specificity 0.98, and positive and negative predictive values were 0.95 and 0.94. When the predictive model was applied on IG, all five unsuccessful cases were predicted correctly, while only 1 out of 23 successful patients was predicted incorrectly.

Conclusion: CondAx-ML was identified as a reliable predictor for long-term stability of early Class III treatment with RME and FM.

Introduction

The combination of rapid maxillary expansion (RME) followed by maxillary protraction with facemask (FM) is one of the most popular protocols for Class III malocclusion (1, 2). The primary goal of RME/FM protocol is to eliminate occlusal interferences from the anterior crossbite and to provide maxillary advancement and mandibular growth control (1), reducing the number of cases with the need for orthognatic surgery (3). The long-term success rate for RME/FM treatment varies between studies (4–13). It is expected, however, that from one-quarter to one-third of the patients treated with a protraction FM will not maintain a positive overjet and require orthodontic camouflage or orthognathic surgery at the end of the growth phase (3–13). Therefore, efforts have been made to find useful cephalometric predictors and/or predictive models for long-term unsuccess of early treatment of Class III malocclusion (5–10, 12, 13). Prediction models are developed to aid health care providers in the estimation of the probability that a specific event will occur in the future, helping in the decision-making process. There is evidence that, in general, the quality of reporting of prediction model studies is inadequate (14, 15). Moreover, questions have been raised if it is really possible to perform a reliable forecast of morphogenetic Class III pattern based simply on baseline cephalometric data. A systematic review on the successful prediction of early treatment of Class III malocclusion, with different appliances (10), concluded that the quality of the papers was rated as low to medium due to different flaws, as the heterogeneity of the samples, and of the treatment methods. Additionally, as most studies on prediction of Class III malocclusion outcome were based on Asian individuals, the identification of predictors for the long-term outcomes of early Class III treatment in Caucasians with FM might have been biased.

The adhesion to a guideline on how to collect and to report clearly all aspects of a prediction model might contribute to the enhancing of the quality and of the transparency of health research. Recently, the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) initiative developed a set of recommendations for the reporting of studies developing, validating, or updating a prediction model for diagnostic or prognostic purposes (14).

Thus, the aim of the current investigation was to develop and to validate a prediction model to forecast long-term unsuccess of early orthopaedic treatment with RME and FM in Class III growing patients.

Materials and methods

We followed the TRIPOD statement (14) for transparent reporting of a multivariable prediction model for individual prognosis. This study was conceived as a Type 2b analysis (14). One group of Class III patients (Brazilian Group, BG) served to develop the prediction model, while another group of Class III patients (Italian Group, IG) was used to evaluate its predictive performance. This study was approved by the Institutional Review Board of Pontifical Catholic University of Minas Gerais (number 057992/2016). The BG included all patients with Class III malocclusion who were treated consecutively with RME/FM protocol from 1992 to 2009 at the Pontifical Catholic University of Minas Gerais (Belo Horizonte, Brazil). The IG consisted of all patients with Class III malocclusion who were treated consecutively with RME/FM protocol from 1992 to 2009 at the University of Rome tor Vergata (Roem, Italy) and at the University of Florence (Florence, Italy). The following inclusion criteria were applied: 1. European–Caucasian ancestry (white individuals); 2. Class III malocclusion, characterized by an anterior crossbite or edge-to-edge incisal relationship and a Wits appraisal of −2.5 mm or less; 3. no orthopaedic/orthodontic treatment prior to T0; 4. cephalograms of adequate quality available before treatment (T0); 5. facial and intraoral photos and/or lateral cephalograms at a long-term observation at least 5 years after FM treatment (T1); 6. minimal age of 17 years at T1; and 7. no craniofacial syndrome or deformities.

Orthopaedic treatment of Class III malocclusion was carried out with the same RME/FM protocol. The mean active treatment period was 1 year. The patients were instructed to wear the FM for 12–14 hours per day (night-time included). All patients were treated at least to a positive dental overjet before discontinuing treatment, with most patients who were overcorrected towards Class II occlusal relationships. As occurs in studies involving any removable device, compliance with the instructions of the orthodontist and staff varied among the patients. Extra-oral elastics delivered 400–500 g of force per side, with about 30 degrees of downward inclination relative to the occlusal plane. All patients received a retention protocol after RME and FM. In particular, in the BG, the retention device was a chin cup that was worn for at least 1 year. In the IG, the retention after FM consisted of either the removable mandibular retractor or the Class III Bionator. When required, a period of treatment with fixed appliances was undertaken to refine and detail the occlusion after the permanent teeth (with the exception of the third molars) had erupted.

Unsuccessful treatment was defined on the basis of the evaluation of both occlusion and profile as derived from the facial and intraoral photos and/or from the lateral cephalograms at T1. For the occlusion, a four-level rating scale was developed for the definition of failure of RME/FM treatment based on plus and minus signals (‘++’, definitively good; ‘+’, good; ‘−’, not good but acceptable; ‘− −’, definitively not acceptable). A ‘definitively not acceptable’ occlusal outcome was defined as the concomitant presence of Class III permanent molar relationship and negative overjet of at least one incisor. Patients who presented minor Class III molar relationship concomitant with either positive overbite/overjet or an edge-to-edge incisors relationship were classified as ‘not good but acceptable’ occlusion. Minor Class III was defined as the mesiobuccal cusp tip of the maxillary first molars lying posteriorly to the mesiobuccal groove of the mandibular groove up to halfway between Class I and full Class III relationship. The classification as a ‘good’ occlusion was given for patients with minor problems with tooth intercuspation, but with positive overbite/overjet, and Class I molar relationship. Patients whose occlusion at T1 was within the principles of a good occlusion (good tooth intercuspation, positive overbite/overjet, and Class I molar relationship) were classified as ‘definitively good’. Regarding the facial component, the facial angle Glabella–Subnasale–Pogonion (G–Sn–Pg) was measured either on the picture of the facial profile or on the lateral cephalograms. An acceptable profile (‘+’) was defined when the facial angle was equal to or smaller than 174 degrees, otherwise the profile was considered unacceptable (‘−’) (16). This threshold value of 174 degrees was calculated as the mean value plus 1 SD for the G–Sn–Pg angle derived from Anić-Milosević *et al.* (16). Patients were classified as unsuccessful when rated with two minus (‘− −’) or three minus (‘− − −’) after the combination of occlusal and facial gradings.

The assessment of unsuccessful treatment was performed by one examiner (B.Q.S) with more than 20 years of orthodontic experience. The inter-rater agreement with another examiner (L.F.) with more than 20 years of orthodontic experience using Kappa statistics was performed on 30 randomly selected patients at T1 before the analysis.

The following variables were evaluated for each patient at T0: chronologic age, gender, and dentition phase (deciduous, early mixed, intermediate mixed, and permanent) (17).

Cephalometric analysis was performed by a single examiner (B.Q.S.) on lateral cephalograms obtained at T0. Lateral cephalograms were de-identified before digitization. Therefore, the examiner who performed the cephalometric analysis was blinded to the outcome when measuring the predictors. All cephalograms were digitized and measured with a cephalometric software (Viewbox version 3.0, dHAL Software, Kifissia, Greece). Correction was made for radiographic enlargement and all measurements were converted to life size. The following 15 landmarks were identified: Basion (Ba), Nasion (N), Anterior Nasal Spine (ANS), Posterior Nasal Spine (PNS), A point A (A), B point (B), Pogonion (Pg), Gnathion (Gn), Menton (Me), Gonion (Go), Gonion intersection (Goi), Articulare (Ar), Articulare anterior (Ara, intersection between the anterior surface of the condyle and the inferior cranial base surface), Condylion (Co), and Center of the condyle (Cc). Point Cc was defined as the midpoint between the two Articulare points Ar and Ara (Articulare anterior). The following three lines or planes were constructed to define the angular variables: Mandibular Plane (MP), Goi–Me; Palatal Plane (PP), ANS–PNS, and Condylar axis (CondAx) passing through points Co and Cc. The cephalometric analysis comprised 24 variables (12 angular, 10 linear, and 2 ratios). Measurements for cranial base angulation and dimension: NSBa, S–N, and S–Ar. Angular and linear measurements for sagittal relationships: SNA, SNB, ANB, AB–MP, and Wits appraisal. Angular and linear measurements for vertical skeletal relationships: SN–PP, SN–MP, PP–MP, CoGoMe, CoGoN, NGoMe, N–Me, and ANS–Me. Angular measurement for condylar inclination: CondAx–MP. Linear measurements for maxillary and mandibular dimensions: Co–A, PNS–A, Co–Gn, Co–Go, and Go–Gn. Two ratios were calculated to measure the lower anterior face height (LAFH) and lower posterior face height (LPFH). LAFH was given by ANS–Me/N–Me and LPFH was given by Co–Go/S–Ar.

All patients that fulfilled the inclusion criteria for both BG and IG were included in the prediction analysis.

Statistical analysis methods

Inter-rater agreement on the assessment of unsuccessful treatment was performed on 30 randomly selected patients at T1 using Kappa statistics. Intra-rater reproducibility assessment for the cephalometric variables was carried out on 25 randomly selected patients after 2 week washout period with Intraclass Correlation Coefficient (ICC, two-way mixed with absolute agreement). The random error was assessed with Springate's method of moments estimator (MME) (18).

Descriptive statistics for continuous predictors (mean and standard deviation) and for categorical predictors (frequency and percentage) at T0 was calculated for both BG and IG. Bivariate logistic regression was performed on all predictors for unsuccessful outcome in the BG. In order to identify a prediction model for the BG, only the statistically significant variables of the bivariate logistic regression were subjected to mixed stepwise logistic regression with $P = 0.05$ to enter and to leave. Accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and the area under the receiver operating characteristic (ROC) curve were calculated for the selected prediction model (internal validation). The validity of the prediction model derived from the BG was then tested on the IG (validation cohort). Accuracy, sensitivity, specificity, PPV, NPV, and the area under the ROC curve were also calculated in the validation cohort to evaluate the predictive performance of the prediction model. JMP® version 13.0.0 2016 (SAS Institute Inc., Cary, NC, USA) and MedCalc Statistical Software version 12.7.8 (MedCalc Software bvba, Ostend, Belgium) were used for the statistical computations.

Results

The inter-rater agreement on the evaluation of unsuccessful treatment at T1 was excellent (Kappa 0.89, 95% confidence interval 0.68–1.0) (19). The values for ICCs varied from 0.87 to 0.99, indicating excellent intra-rater agreement (19). The MME random error measurements ranged from 0.4 to 1.3 degrees for the angular variables and from 0.3 and 0.9 mm for the linear measurements.

The BG group consisted of 73 Class III patients (41 females and 32 males). The mean age at T0 (before treatment) was 7.1 ± 1.6 years, the mean age at T1 (long-term follow-up) was 21.8 ± 3.2 years, and the T0–T1 interval was 14.7 ± 3.3 years. The IG group (validation cohort) comprised 28 Class III patients (14 females and 14 males). The mean age at T0 (before treatment) was 9.0 ± 1.3 years, the mean

Table 1. Descriptive statistics. Mean (standard deviation) [minimum; maximum] for quantitative variables. Frequency (percentage) for qualitative variables

	Brazilian group N = 73	Italian group N = 28
Age T0 (years)	7.1 (1.6) [4.4; 11.1]	9.0 (1.3) [6.0; 11.2]
Age T1 (years)	21.8 (3.2) [17.0; 30.5]	18.2 (1.4) [17.0; 23.9]
T0–T1 interval (years)	14.7 (3.3) [8.4; 25.3]	9.3 (1.9) [5.9; 14.3]
Gender (F)	41 (56%)	14 (50%)
Gender (M)	32 (44%)	14 (50%)
Unsuccess	22 (30%)	5 (18%)
Success	51 (70%)	23 (82%)
Deciduous	26 (36%)	0 (0%)
Early mixed	25 (34%)	10 (36%)
Intermediate mixed	16 (22%)	10 (36%)
Late mixed	6 (8%)	6 (21%)
Permanent	0 (0%)	2 (7%)
NSBa (degree)	130.7 (4.4)	130.5 (7.2)
S–N (mm)	61.1 (2.8)	63.1 (2.2)
S–Ar (mm)	27.0 (3.2)	29.5 (2.9)
SNA (degree)	80.0 (4.0)	79.6 (4.5)
SNB (degree)	79.3 (3.5)	78.4 (3.9)
ANB (degree)	0.7 (2.2)	1.2 (2.3)
AB-MP (mm)	64.7 (3.6)	65.6 (4.4)
Wits (mm)	−5.5 (1.9)	−5.7 (2.0)
SN–PP (degree)	8.1 (3.1)	8.8 (3.3)
SN–MP (degree)	37.8 (4.2)	38.7 (4.4)
PP–MP (degree)	29.7 (4.2)	29.9 (3.9)
CoGoMe (degree)	130.1 (4.9)	129.1 (3.3)
CoGoN (degree)	53.5 (3.3)	50.9 (3.5)
NGoMe (degree)	76.6 (3.7)	78.1 (3.3)
ANS–Me (mm)	55.0 (4.5)	59.7 (3.3)
N–Me (mm)	96.8 (6.7)	105.6 (4.3)
CondAx–MP (degree)	142.8 (8.0)	141.9 (7.3)
Co–A (mm)	72.5 (4.1)	76.3 (4.4)
PNS–A (mm)	38.8 (3.1)	41.0 (2.7)
Co–Gn (mm)	95.4 (6.1)	102.5 (5.6)
Co–Go (mm)	43.6 (4.1)	47.8 (3.5)
Go–Gn (mm)	62.2 (4.8)	66.4 (4.3)
AnsMe/NMe (ratio)	0.6 (0.02)	0.6 (0.02)
CoGo/SAr (ratio)	1.6 (0.2)	1.6 (0.2)

age at T1 (long-term follow-up) was 18.2 ± 1.4 years, and the T0–T1 interval was 9.3 ± 1.9 years. The number of unsuccessful cases was 22 in the BG and 5 in the IG. In Table 1, we reported the descriptive statistics for all the variables that were assessed in both BG and IG. Table 2 describes the results of the prediction of unsuccess for all the predictors after application of the bivariate logistic regression in the BG. Significant predictors for unsuccess were: age at T0, non-deciduous dentition, gender (male), CondAx–MP, Go–Gn, Co–Gn, Co–A, PNS–A, N–Me, and, inversely, AB–MP. To identify a prediction model, only the statistically significant predictors in the bivariate logistic regression were subjected to mixed stepwise logistic regression with $P = 0.05$ to enter and to leave. The only predictor that was selected in the predictive model was CondAx–MP (odds ratio: 1.52, 95% confidence interval: 1.25–1.85, $P < 0.0001$). Table 3 reports the prediction model. The area under the curve was 0.95. The probability of unsuccess (P) can be calculated as:

$$P = \frac{1}{1 + e^{-(-62.029+0.41973\,\mathrm{CondaxMP})}}$$

Multiplying by 100 converts the probability into a percent risk. Figure 1 depicts the prediction profiler plot for CondAx–MP with the percentage of unsuccess on the Y-axis and the values of the predictive variable on the X-axis. Unsuccessful treatment at T1 was predicted for values of CondAx–MP at T0 greater than the cut-off value of 147.8 degrees. When considering this cut-off value, BG patients predicted incorrectly were 3 out of 22 for the unsuccessful cases and 1 out of 51 for the successful cases. Therefore, for the predictive model CondAx–MP in the BG, the accuracy was 0.95, the sensitivity was 0.86, the specificity was 0.98, the PPV was 0.95, and the NPV was 0.94. When the prediction model was applied on the validation cohort IG, the results were even better as all unsuccessful cases were predicted correctly, while only 1 out of 23 successful patients was predicted incorrectly. Therefore, for the predictive model CondAx–MP in the IG, the accuracy was 0.96, the sensitivity was 1.00, the specificity was 0.96, the PPV was 0.83, and the NPV was 1.00.

Table 2. Prediction of unsuccess for each variable in the Brazilian group tested with bivariate logistic regression. $N = 73$, unsuccess = 22 (30%)

	OR	95% OR	*P*-value
Age T0	1.45	1.05; 1.99	0.0191
Gender (M)	3.21	1.13; 9.09	0.0250
Deciduous	0.19	0.05; 0.73	0.0068
NSBa	0.98	0.87; 1.10	0.7198
S–N	1.17	0.97; 1.41	0.0930
S–Ar	1.06	0.90; 1.24	0.4881
SNA	1.06	0.93; 1.20	0.3764
SNB	1.06	0.92; 1.23	0.3889
ANB	1.02	0.81; 1.30	0.8383
AB–MP	0.85	0.73; 1.00	0.0401
Wits	0.96	0.74; 1.25	0.7699
SN–PP	1.05	0.89; 1.24	0.9515
SN–MP	1.06	0.94; 1.20	0.3228
PP–MP	1.03	0.92; 1.16	0.5867
CoGoMe	1.11	0.99; 1.24	0.0629
CoGoN	1.06	0.90; 1.24	0.4833
NGoMe	1.13	0.99; 1.30	0.0741
ANS–Me	1.11	0.99; 1.25	0.0699
N–Me	1.10	1.02; 1.19	0.0152
CondAx–MP	1.52	1.25; 1.85	<0.0001
Co–A	1.18	1.03; 1.35	0.0143
PNS–A	1.25	1.03; 1.51	0.0151
Co–Gn	1.14	1.03; 1.25	0.0044
Co–Go	1.06	0.94; 1.20	0.3430
Go–Gn	1.14	1.01; 1.27	0.0204
Ratio AnsMe/NMe	0.0006	3.3×10^{-15}; 1.0×10^{8}	0.5707
Ratio CoGo/SAr	1.13	0.06; 21.5	0.9358

OR, odds ratio; CI, confidence interval.

Discussion

The effectiveness of an orthodontic treatment can be evaluated in several ways. Long-term stability, however, is one of the most ambitious goals for the practitioner (20). When dentofacial orthopaedic treatment is performed during the early developmental phases, facial residual growth might challenge the maintenance of successful outcomes. Active facial growth can be an ally of orthodontists during early treatment, but it can be a problem during the retention phase, especially in patients with Class III malocclusion (4, 11, 21). As the absolute effectiveness of early Class III treatment has to be evaluated after the end of active mandibular growth (that might extend over

Table 3. Prediction model for unsuccess in the Brazilian group ($R^2 = 0.59$)

Term	Estimate	Std error	OR	95% CI	*P*-value
Intercept	−62.029	14.859			
CondAx-ML	0.420	0.101	1.52	1.25; 1.85	<0.0001

OR, odds ratio; CI, confidence interval.

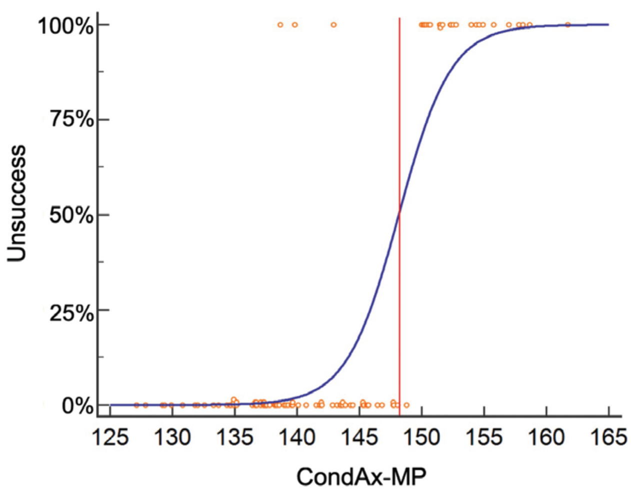

Figure 1. Prediction profile plot for CondAx–MP estimated for the Brazilian Group in the logistic regression model. The X-axis represents the values of the predictice variable (CondAx–MP), while the Y-axis represents the probability of unsuccess. The blue curve derived from the estimate of the logistic regression model shows the probability of unsuccess on the values of CondAx–MP. For the cut-off value of 147.8 degree (red vertical line) of CondAx–MP, the estimated probability of unsuccess is 50%. Each dot represents one patient. Dots located on 0% value on the Y-axis refer to successful cases, while dots located on 100% value on the Y-axis refer to unsuccessful cases. The position of the dots on the X-axis correspond to their CondAx–MP values. The dots that are left to the cut-off value of 147.8 degree indicated by the red vertical line are patients predicted as successful while dots right to the red vertical line are patients predicted as unsuccessful. The single dot right to the red line and located to 0% on the Y-axis was wrongly predicted as unsuccessful while the three dots left to the red line and located to 100% on the Y-axis were wrongly predicted as successful.

the second decade of life) (22), both clinicians and patients could benefit from reliable predictive parameters of long-term successful and/or unsuccessful outcomes before early treatment begins.

Thus, in the past two decades, several investigations have been reported in the search of a prediction model for baseline cephalometric features that can be used to forecast the long-term outcome of the early approach to Class III malocclusion with the RME/FM protocol. A systematic review on the successful prediction of early treatment of Class III malocclusion, with different appliances (10), concluded that heterogeneity of the samples, and of the treatment methods, prevented a valid analysis among the seven previous papers that have dealt with the predictability of RME/FM treatment. The quality of the papers was rated as low to medium (10). Due to different flaws, the identification of predictors for the long-term outcomes of early Class III treatment with FM might have been biased. The following problems were identified: 1. the method of statistical analysis (discriminant analysis instead of logistic regression analysis); 2. the relatively small sample size (varying between 26 to 64 individuals, which might impair the multivariate analyses); 3. the immature age of the patients at the final evaluation (some were not given, but in most papers the final age was before 15 years old); 4. ethnical ancestry (Asian population, in three out of seven papers, including the two with larger sample size); 5. lack of validation of the prediction model.

Nowadays, it is recognized that a complete and accurate report of prediction model investigations is mandatory to critically appraise, externally validate, and eventually use prediction models in the daily clinical practice. Therefore, the aim of this study was to develop and validate a prediction model to forecast long-term unsuccess of early orthopaedic treatment with RME and FM in Class III growing patients. In particular, the prediction model was derived from a BG of 73 patients. Patients had been treated with RME/FM at the age of 7.1 years and re-examined 14.8 years later at 21 years of age. The prediction model consisted of only one variable (CondAx–MP). The results of the current study showed that the inclination of the condylar axis to the mandibular plane is a powerful predictor for long-term unsuccessful outcomes of early Class III treatment with RME and FM. Unsuccessful treatment at T1 was predicted for pre-treatment values of CondAx–MP greater than the cut-off value of 147.8 degrees. For this model, the accuracy, sensitivity, and specificity were high (95, 86, and 98 per cent, respectively). Positive predictive value (probability that patients with CondAx–MP greater than 147.8 degrees at T0 are truly unsuccessful cases) and negative predictive value (probability that patients with CondAx–MP smaller than 147.8 degrees at T0 are truly successful cases) were also high (95 and 94 per cent, respectively).

For the external validation testing, a sample composed by the 28 Italian Caucasian patients was collected. The accuracy, sensitivity, specificity, positive predictive, and negative predictive values also were high (96, 100, 96, 83, and 100 per cent, respectively). This is a very important information, proving that the model runs well with other groups of patients. Unfortunately, the validation of predictive models with an external sample of similar individuals is a methodological step that previous studies on Class III malocclusion had not performed. Future studies using the Condylar axis to mandibular plane angle in the prediction of long-term effectiveness are necessary but with different groups of Caucasian individuals, and even with other ethnical groups, which can corroborate or not the current findings.

The results of this study are in agreement with those reported by the systematic review by Fudalej *et al.* (10), who reported that a large gonial angle was a predictor for unsuccessful results with different treatment modalities for Class III malocclusion. Also Nardoni *et al.* (12) found that CondAx–MP was a predictor for unsuccessful treatment with RME and FM. However, unsuccessful treatment was predicted for small values of CondAx–MP at the start of treatment. This different outcome can be explained by the fact that the values for CondAx–MP had to be combined with those of another predictor (lower anterior face height).

A limitation of this study was that the prediction model derived from a Caucasian population, and thus it is applicable for this ethnical group of individuals. Another limitation is the sample size of the validation cohort (IG). It was smaller than the BG from which the prediction model was developed. To test this predictive model in a larger sample of Caucasian patients treated with the RME/FM protocol is a future project. It is also a limitation of the study design to have analysed only the data from baseline cephalometric pattern, not including all other information regarding the Class III familiar tracts of the patients, as well as the combination of facial type and sagittal inter-arch relationship. However, the objective of this study was to identify risk factors based on pre-treatment cephalometric pattern. Another limitation of this study was that the validation cohort (IG) did not match perfectly the group from which the prediction model was derived (BG). In fact, there were between-group differences in chronologic ages at T0, T1, in the T0–T1 interval, in dentition stages, and in a few of the cephalometric variables. These differences, however, did not affect the generalizability and accuracy of the prediction model. Finally, a limitation of this study was that 12 patients in the BG and 13 patients in the IG were re-evaluated between 17.0 and 17.9 years of age. Baccetti *et al.* (22) showed that mandibular growth, especially in males, may continue after the age of 18 years.

In conclusion, it was found that the inclination of the condylar axis relative to the mandibular plane should be analysed and measured before treatment by the orthodontist, identifying those Class III patients whose angle is greater than 148 degrees. Those patients present a poor long-term prognosis. This information adds another tool in diagnosis rationale of young Class III patients. However, depending on the cultural, economic, and institutional peculiarities of different countries and services, individuals with poor prognosis will receive a personalized orientation on the best approach in the short and long term.

Conclusion

CondAx-ML was identified as a reliable predictor for long-term stability of early Class III treatment with RME and FM.

Conflict of Interest

None to declare.

References

1. Ngan, P. and Moon, W. (2015) Evolution of Class III treatment in orthodontics. *American Journal of Orthodontics and Dentofacial Orthopedics,* 148, 22–36.
2. Cordasco, G., Matarese, G., Rustico, L., Fastuca, S., Caprioglio, A., Lindauer, S.J. and Nucera, R. (2014) Efficacy of orthopedic treatment with protraction facemask on skeletal Class III malocclusion: a systematic review and meta-analysis. *Orthodontics & Craniofacial Research,* 17, 133–143.
3. Mandall, N., *et al.* (2016) Early class III protraction facemask treatment reduces the need for orthognathic surgery: a multi-centre, two-arm parallel randomized, controlled trial. *Journal of Orthodontics,* 43, 164–175.

4. Westwood, P.V., McNamara, J.A. Jr, Baccetti, T., Franchi, L. and Sarver, D.M. (2003) Long-term effects of Class III treatment with rapid maxillary expansion and facemask therapy followed by fixed appliances. *American Journal of Orthodontics and Dentofacial Orthopedics,* 123, 306–320.
5. Baccetti, T., Franchi, L. and McNamara, J.A. Jr (2004) Cephalometric variables predicting the long-term success or failure of combined rapid maxillary expansion and facial mask therapy. *American Journal of Orthodontics and Dentofacial Orthopedics,* 126, 16–22.
6. Moon, Y.M., Ahn, S.J. and Chang, Y.I. (2005) Cephalometric predictors of long-term stability in the early treatment of Class III malocclusion. *The Angle Orthodontist,* 75, 747–753.
7. Ghiz, M.A., Ngan, P. and Gunel, E. (2005) Cephalometric variables to predict future success of early orthopedic Class III treatment. *American Journal of Orthodontics and Dentofacial Orthopedics,* 127, 301–306.
8. Wells, A.P., Sarver, D.M. and Proffit, W.R. (2006) Long-term efficacy of reverse pull headgear therapy. *The Angle Orthodontist,* 76, 915–922.
9. Yoshida, I., Yamaguchi, N. and Mizoguchi, I. (2006) Prediction of post-treatment outcome after combined treatment with maxillary protraction and chincap appliances. *European Journal of Orthodontics,* 28, 89–96.
10. Fudalej, P., Dragan, M. and Wedrychowska-Szulc, B. (2011) Prediction of the outcome of orthodontic treatment of Class III malocclusions–a systematic review. *European Journal of Orthodontics,* 33, 190–197.
11. Masucci, C., Franchi, L., Defraia, E., Mucedero, M., Cozza, P. and Baccetti, T. (2011) Stability of rapid maxillary expansion and facemask therapy: a long-term controlled study. *American Journal of Orthodontics and Dentofacial Orthopedics,* 140, 493–500.
12. Nardoni, D.N., Siqueira, D.F., Cardoso, M.d.e.A. and Capelozza Filho, L. (2015) Cephalometric variables used to predict the success of interceptive treatment with rapid maxillary expansion and face mask. A longitudinal study. *Dental Press Journal of Orthodontics,* 20, 85–96.
13. Choi, Y.J., Chang, J.E., Chung, C.J., Tahk, J.H. and Kim, K.H. (2017) Prediction of long-term success of orthopedic treatment in skeletal Class III malocclusions. *American Journal of Orthodontics and Dentofacial Orthopedics,* 152, 193–203.
14. Moons, K.G., Altman, D.G., Reitsma, J.B., Ioannidis, J.P., Macaskill, P., Steyerberg, E.W., Vickers, A.J., Ransohoff, D.F. and Collins, G.S. (2015) Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of Internal Medicine,* 162, W1–73.
15. Heus, P., Damen, J.A.A.G., Pajouheshnia, R., Scholten, R.J.P.M., Reitsma, J.B., Collins, G.S., Altman, D.G., Moons, K.G.M. and Hooft, L. (2018) Poor reporting of multivariable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement. *BMC Medicine,* 16, 120.
16. Anić-Milosević, S., Lapter-Varga, M. and Slaj, M. (2008) Analysis of the soft tissue facial profile by means of angular measurements. *European Journal of Orthodontics,* 30, 135–140.
17. Franchi, L., Baccetti, T., De Toffol, L., Polimeni, A. and Cozza, P. (2008) Phases of the dentition for the assessment of skeletal maturity: a diagnostic performance study. *American Journal of Orthodontics and Dentofacial Orthopedics,* 133, 395–400; quiz 476.e1.
18. Springate, S.D. (2012) The effect of sample size and bias on the reliability of estimates of error: a comparative study of Dahlberg's formula. *European Journal of Orthodontics,* 34, 158–163.
19. Fleiss, J. (1986) *The Design and Analysis of Clinical Experiments*. John Wiley and Sons, New York, NY.
20. Al Yami, E.A., Kuijpers-Jagtman, A.M. and van 't Hof, M.A. (1999) Stability of orthodontic treatment outcome: follow-up until 10 years postretention. *American Journal of Orthodontics and Dentofacial Orthopedics,* 115, 300–304.
21. Franchi, L., Baccetti, T. and Tollaro, I. (1997) Predictive variables for the outcome of early functional treatment of Class III malocclusion. *American Journal of Orthodontics and Dentofacial Orthopedics,* 112, 80–86.
22. Baccetti, T., Franchi, L., McNamara, J.A., Jr (2007) Growth in the untreated Class III subject. *Seminars in Orthodontics,* 13, 130–42.