European Spine Journal (2024) 33:3767–3775
https://doi.org/10.1007/s00586-024-08487-0

ORIGINAL ARTICLE

Puberty changes the natural history of idiopathic scoliosis: three prediction models for future radiographic curve severity from 1563 consecutive patients

Stefano Negrini[1,2] · Maryna Yaskina[3] · Sabrina Donzelli[4] · Alberto Negrini[5] · Giulia Rebagliati[5] · Claudio Cordani[2] · Fabio Zaina[5] · Eric C. Parent[6]

Received: 22 January 2024 / Revised: 29 July 2024 / Accepted: 3 September 2024 / Published online: 14 September 2024

Abstract

Purpose Understanding idiopathic scoliosis (IS) natural history during growth is essential for shared decision-making between patients and physicians. We developed a retrospective model with the largest available sample in the literature and we aimed to investigate if using three peri-pubertal growth periods provides better prediction than a unique model.

Methods Secondary analysis of a previous study on IS natural history data from radiographs before and at the first consult. Three groups: BEFORE (age 6–10), AT (age 11-Risser 2) and AFTER (from Risser 3) the pubertal growth spurt. Available predictors: Cobb angle, curve type, sex, observation time, and Risser score. We used linear mixed-effects models to predict future Cobb angles in each group. We internally validated prediction accuracy with over 100 patients per group (3 to 5-fold cross-validation).

Results We included 1563 participants (275 BEFORE, 316 AFTER, 782 females and 190 males AT). Curves increased over time mostly in AT, importantly in BEFORE, but also in AFTER. All models performed better than the general one. In BEFORE, 74.2% of the predictions were within ±5°, 71.8% in AFTER, 68.2% in AT females, and 60.4% in males. The predictors (baseline curve, observation time also squared and cubic, and Risser score) were similar in all the models, with sex influencing only AFTER.

Conclusion IS curve severities increase differently during growth with puberty stages. Model accuracy increases when tailored by growth spurt periods. Our models may help patients and clinicians share decisions, identify the risk of progression and inform treatment planning.

Keywords Scoliosis · Prognosis · Natural history · Puberty

Introduction

The natural history of a condition represents its evolution over time when untreated [1]. Understanding the natural history is essential to inform the prescription of any intervention aiming to modify the course of the condition while avoiding unnecessary treatments [2].

Scoliosis is a three-dimensional torsional deformity of the spine and trunk [3]. In 80% of cases, it is idiopathic (IS) because it is not attributable to a specific cause. Adolescent IS is the most common type (prevalence of 2–3%) [3–5]. IS progresses with growth, especially during puberty [6]. Not all patients show the same progression rate, and studies do not show consistent results [7–9]. A meta-analysis [10] confirmed the heterogeneity of natural history studies,

✉ Stefano Negrini
Stefano.negrini@unimi.it

1 Department of Biomedical, Surgical and Dental Sciences, University "La Statale", Milan, Italy

2 IRCCS Istituto Ortopedico Galeazzi, Milan, Italy

3 Women and Children's Health Research Institute, University of Alberta, Edmonton, Canada

4 Utrecht University Medical Center UMC Otropedic Department Utrecht, Utrecht, Netherlands

5 ISICO (Italian Scientific Spine Institute), Milan, Italy

6 Department of Physical Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, Canada

precluding providing reliable information on progression during growth.

The Bracing in Adolescent IS Trial (BrAIST) [11] was stopped early for ethical reasons, given the clear superiority of bracing over natural history in the BrAIST study [11]. This result also reduces the probability of observing other prospective untreated IS cohorts during growth. We proposed a novel approach to collecting natural history information based on retrospective observation of prospectively collected radiographs [12]. In a previous paper [12], we described a general progression model throughout all growth. We could predict 47% of the observed values with adequate precision (±5°). The prediction curve showed a progressively slower progression with time, consistent with our understanding of natural history after puberty but not with IS progression in other growth periods. To improve on this model, we considered the studies of Duval—Beaupère et al. [13], whose hypothesis is confirmed by clinical experience. They recognised 3 phases of scoliosis progression: a first slow period from discovery to the start of puberty (Point P), a second rapid period up to Risser 3, and a final period of gradual stabilisation until the end of growth. We hypothesised that splitting our data and developing models specifically for 3 phases similar to those described by Duval-Beaupère et al. [13] could improve the predictive validity of our earlier model. Therefore, this secondary analysis of the previous data [12] aims to describe the natural history of IS in different growth phases and whether a 3-stage model predicts the Cobb angle more accurately than the single-stage model we previously developed [12]. Due to the characthristics of the prospective database used for the study, we will study only the available data: frontal spine radiographic information (Risser, curve type, tri-radiate cartilage, Cobb angle) and demographic parameters (sex, age). These prediction models for future curve severity could have a significant clinical value in informing treatments and achieving shared-decision making with patients and families.

Materials and methods

Design

We performed a retrospective analysis of the clinical records from tertiary care clinics in our Institute, which specialises in scoliosis. To obtain natural history information, we evaluated the radiographs children performed before accessing our clinics, provided they did not receive any treatment. The local Ethical Committee approved the study. All patients (or their parents, if minors) provided informed consent to a retrospective anonymised analysis of their clinical data. The 2018 Research Grant by the Scoliosis Research Society (SRS) provided the funding for the study. This paper presents a secondary analysis of the data reported in another study [12].

Participants

We selected consecutive eligible participants from all the children's data available in our Institute's clinical files. We checked whether they had complete treatment history and available radiographic images (collected since October 2008). To increase the number of participants, we recruited new patients prospectively throughout the research project from 2017 to April 2020. We also asked all children to bring the radiographs taken before coming to our Institute and systematically checked their treatment history.

We recruited patients with IS who had at least two standing full-spine coronal radiographs before any treatment: one at our first consultation and at least another one while still growing (up to European [EU] Risser 4) taken before coming to our Institute. At their first Institute visit, participants could be during growth or skeletally mature (when the curve is likely stable) because we were interested in predicting curve severity until the end of growth. We set an upper age limit at 25 years for the first consult to avoid including participants with curve progression after bone maturity.

To analyse data according to Duval-Beaupère et al.'s hypothesis [13], we had to identify the pubertal growth spurt, which was impossible in our retrospective radiographic sample. Consequently, we identified the best available proxy by looking at the skeletal maturity. We wanted to identify three groups (1) BEFORE: before the growth spurt; (2) AT: from the onset of the growth spurt to EU Risser 2 skeletal maturity (corresponding to US Risser 3 [14]) – due to the different age between sexes at pubertal growth spurt, we split this age into two groups: 2a) AT-females and 2b) AT-males; (3) AFTER: from EU Risser 3 to end of growth. We used the EU Risser score because it was the only bone maturity sign visible in the available radiographs. Radiographic data allow easy retrieving of the well-established EU Risser 3 boundary to identify AFTER puberty [11, 14, 15], while there is no radiographic threshold exists to determine BEFORE puberty. According to Duval-Beaupère et al. [13], the first stage of the EU Risser sign appears in the middle of the pubertal growth spurt, which usually starts a few months earlier. For this reason, and because females experience the growth spurt before males, we searched the age at which we could minimise the number of patients with EU Risser 1 in the female cohort. This age served as the boundary between BEFORE and AT. Of all participants in the previous study [12], we included only those with two radiographs within the timeframe of one of the three groups we defined.

Exclusion criteria were prior brace or Physiotherapeutic Scoliosis Specific Exercises (PSSE) treatment, previous injuries, surgery, or pathologies that could have affected the scoliosis appearance or progression between the considered radiographs. We did not exclude patients according to the Cobb angle, provided they had at least one x-ray showing a curve above 10°.

Quality check

Two expert clinicians verified the clinical records of all eligible patients to exclude all those previously treated before they accessed our Institute. We asked for information from the treating physician and then from the patient in case of any doubt. We excluded all patients with uncertain prior treatments. Two evaluators blindly re-measured a random sample of 100 radiographs, confirming adequate measurement accuracy. This result allowed the use of measurements from the clinical records.

Variables

We included all the pairs of eligible individuals' radiographs within each group's timeframe. We considered the oldest radiograph as the baseline. We used continuous independent variables: Cobb angle in degrees (at each observation) and observation time in years (also squared and cubic) (the variable "TIME") between the radiographs. We used the following independent baseline categorical variables: EU Risser score, sex, curve type (single thoracic, single lumbar/thoracolumbar, double and others), and triradiate cartilage (open/close). We also checked a continuous variable for curve severity, using the clinically significant 30° threshold [16], applying the Cobb degrees – 30°, with the lowest cut-off at 0.

Table 1 Characteristics of the subgroups for sex, European Risser, curve type and Cobb degrees

Groups	Before puberty	AT puberty	After puberty
	N (%)	*N* (%)	*N* (%)
Sex			
Female	246 (90)	782 (80)	318 (84)
Male	29 (10)	190 (20)	60 (16)
European Risser			
0	274 (100)	660 (84)	-
1	1 (0.4)	91 (12)	-
2	-	31 (4)	-
3	-	-	301 (80)
4	-	-	75 (20)
5	-	-	2 (0.5)
Curve Type			
No curve	-	2 (0.3)	-
Single thoracic curve	54 (20)	127 (16)	52 (14)
Single thoracolumbar or lumbar curve	55 (20)	194 (15)	121 (32)
Double curve	142 (52)	410 (52)	174 (46)
Upper thoracic, triple, double thoracic or quadruple curve	24 (9)	49 (6)	31 (8)
Cobb degrees classification			
≤ 30°	262 (95)	666 (85)	244 (65)
> 30°	13 (5)	116 (15)	134 (35)

Model derivation and validation

Considering the growth model characteristics of the curve severity data (slow progression, rapid progression and stabilization), we derived a prediction model of future Cobb angles for each puberty subgroup. We used linear mixed-effects models with random effects and a variance components structure. Independent variables were chosen based on their clinical and statistical significance. We decided on the final best model for each subgroup by evaluating the smallest Akaike (AIC) and Bayesian (BIC) Information Criterion.

Internal validation was done using cross-validation to test the prediction accuracy of each model. We aimed for a 5-fold cross-validation. If any group had less than 500 participants, we performed the maximum possible cross-validations with nearly 100 patients per run.

To assess prediction accuracy, we used standard prediction intervals. For clinical purposes, we also used new intervals of specified width centred at predicted values obtained from the model. We set the first of such intervals at the recognised radiographic measurement error (± 5°) threshold used to define a significant change [14]; we then progressively added ± 5° until we determined the interval with 95% accuracy during the cross-validations.

Results

Group selection and sample description

To determine the age boundary between BEFORE and AT, we found 77, 246 and 548 female participants at ages 9, 10 and 11, with an EU Risser 1 prevalence of 1,3%, 3,2% and 10,2%, respectively. Consequently, we included age 10 in BEFORE and 11 in AT subgroups. We finally had 275 participants (246 females) in BEFORE, 782 in AT females, 190 in AT males, and 378 (318 females) in AFTER. Table 1 reports the participants' characteristics of the three subgroups. The average time between radiographs was shorter in AT-females (1,1 ± 0,6 years) compared to the other groups (1,4 ± 0,8, 1,3 ± 0,7, and 1,5 ± 1,2 for BEFORE, AT-males and AFTER, respectively) (Table 2). Cobb angles increased in all subgroups from baseline to the end of observation (Table 2).

Table 2 Baseline and end of observation data of the main variables studied to develop the models

Groups	Before puberty (*N* = 275)		AT puberty				After puberty (*N* = 378)		P
			Females (*N* = 782)		Males (*N* = 190)				
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Time between radiographs (years)	1.4	0.8	1.0	0.6	1.3	0.7	1.5	1.2	< 0.01
Main curve severity (degree Cobb)									
Baseline	16.9	7.1	21.3	10.4	18.3	8.2	27.9	12.1	< 0.01
End of observation	23.7	10.7	28.8	13.1	26.9	12.2	31.2	13.5	< 0.01
Age (years)									
Baseline	8.6	1.2	12.1	0.8	12.8	1.2	14.5	1.4	< 0.01
End of observation	11.3	1.5	13.7	1.5	14.5	1.4	16.3	2.1	< 0.01

Table 3 Prediction models in the three groups

Groups	Before puberty			At puberty			After puberty		
	Coeff.	95CI	*P*	Coeff.	95CI	*P*	Coeff.	95CI	*P*
Intercept	1.64	-2.69; 5.98	0.46	-1.01	-3.64; 1.62	0.45	1.44	-0.37; 3.25	0.12
Baseline Curve (°Cobb)	1.09	0.98; 1.21	< 0.0001	1.12	1.07; 1.16	< 0.0001	1.03	0.99; 1.07	< 0.0001
Time (Years)	0.82	-5.05; 6.70	0.78	6.88	1.93; 11.83	0.007	1.78	1.07; 2.50	< 0.0001
$Time^2$	0.73	-1.89; 3.35	0.58	-1.81	-4.89; 1.26	0.25	-0.14	-0.24; -0.036	0.008
$Time^3$	-0.08	-0.41; 0.25	0.62	0.35	-0.20; 0.89	0.21			
European Risser 1				-1.85	-3.33; -0.37	0.01			
European Risser 2				-2.53	-5.01; -0.07	0.045			
European Risser 4							-2.04	-3.23; -0.86	0.0008
Female							-1.44	-2.73; -0.14	0.03

95CI: 95% Confidence Interval. Coeff.: Coefficient. Note. EU Risser 0 and 3 are not included because they represent the baseline values (= 0) for AT and AFTER groups, respectively

Considering individual patients, we found 48%, 62%, and 49% worsened 5° Cobb or more in groups BEFORE, AT and AFTER, respectively. We also found patients improved 5° Cobb or more: 2% in BEFORE and AFTER and 1% in AT, with 1 improved 10° Cobb or more in BEFORE and AFTER, and 4 in AT.

Prediction models

The predictors of future Cobb angle included in the best predictive models for the different groups were similar (Table 3). They included the baseline Cobb angle, the observation time between baseline and prediction (linear, squared, and cubic) (viariable "Time"), and the EU Risser score (in AT and AFTER). Sex was a predictor only in AFTER. The predictors with a negative association with predicted curve severity were higher EU Risser score (in AT and AFTER), squared "Time" (in AT and AFTER), cubic "Time" (in BEFORE) and female sex (in AFTER). The variable "Time" showed a larger effect in AT (6,88 versus 0,82 and 1,78 for BEFORE and AFTER, respectively). Neither curve type nor triradiate cartilage were included in the groups' final prediction models.

Prediction accuracy in internal cross-validations

Subgroup size allowed performing three cross-validations in BEFORE and AFTER and 5 in AT-females. We cross-validated AT males in the female's final model. The median prediction error varied between 2,9 and 4,0 for the models developed with females but reached 14,7 when applying the female model AT puberty in males (Table 4). Considering the classical radiographic measurement error of ± 5°, the accuracy for the final models (measured as the proportion of the predictions within 5° of the observed values) was 74% (BEFORE), 72% (AFTER), 68% (AT-females), and 60% (AT-males). We reached the target of almost 95% accuracy of the observed values within ± 15° from the prediction for the groups with the most rapid progression (BEFORE and AT, females and males) and within ± 10° in the slowest one (AFTER).

Discussion

Understanding the natural history of idiopathic scoliosis is essential for selecting the intervention to modify the disease's course and avoid unnecessary treatments. There

Table 4 Accuracy of future Cobb angle prediction using the models developed in the three puberty groups. 25th = 25th percentile; 75th = 75th percentile

Groups	Error			Prediction accuracy			
	*M*edian	25th	75th	*W*ithin interval	*W*ithin ± 5°	*W*ithin ± 10°	*W*ithin ± 15°
Before Puberty	3.8	1.8	7.3	25%	74%	90%	97%
At Puberty	4.0	2.4	7.1	15%	68%	90%	97%
Females	14.7	4.4	51.3	15%	60%	84%	95%
Males							
After Puberty	2.9	1.4	5.3	23%	72%	95%	-
Overall Model [12]	5.5	2.7	9.9	9.3%	47%	80%	91%

is general agreement that the progression risk changes according to the growth phase. Duval-Beaupère et al. [13] hypothesised three different curve progression slopes during growth, as confirmed by clinical experience. We aimed to verify if, following this assumption, we could build IS curve progression prediction models that were more precise than the overall model we developed previously [12]. Our new models proved more accurate. The prediction accuracy within the accepted radiographic measurement error of ± 5° improved from 47% OVERALL to 74% (BEFORE), 72% (AFTER), 68% (AT-females), and 60% (AT-males). These results also confirm that an important portion of the variance in the progression of scoliosis remains individual and based on other non-radiographic factors.

As expected, we verified that AT puberty is the period with the largest curve severity increases. In AFTER, we found that EU Risser 3 did not correspond to a stop of progression as usually proposed [15]: this result could be particularly relevant in severe curves in clinics. It is important to note that we studied AFTER patients from EU Risser 3 [17], which corresponds to the US Risser 4: consequently, progression is possible also after what is usually considered skeletal maturity [15]. While related to progression [9], we did not retain the topographical curve classification in the final models. A sex difference was present only in AFTER, with females progressing less than males. These results need further studies to be confirmed. Nevertheless, considering the higher prevalence of females with IS, they could explain the general assumption that the progression of scoliosis stops at EU Risser 3. In AT, we tested the model's accuracy developed for females using male data. We showed that pubertal progression is only delayed in males but keeps a similar predictable behaviour. The accuracy decrease suggests caution in the implementation, and future studies should focus on the evolution in males AT puberty.

An interesting and surprising result concerns patients' radiographic improvements in time above the classical measurement error threshold of 5°. When there were not yet ethical concerns about repeating radiographs within a short time, old publications showed results consistent with ours [18]. This stresses the need for caution in considering radiographic measures and complementing them with clinical repeatable data collection to allow proper decision-making.

Comparison with previous publications

Di Felice [10] performed a meta-analysis of the current literature covering 16 papers with 4083 patients reporting on the natural history of idiopathic scoliosis until 2018. The authors found that infantile IS between 15 and 35°, with 3 to 12 years of follow-up, progressed in 49% of cases (95% Confidence Interval (95CI) 1–97%). They did not find studies on juvenile IS alone but mixed with adolescent IS: curves below 25° progressed in 49% of cases (95CI: 19–79%) by 2,2° to 9,6° per year over 1 to 5,5 years. Finally, adolescents with IS between 11,1 and 13,8 years and curves from 19° to 30° progressed in 42% of cases (95CI: 11–73%) over follow-ups of 7 months to 4 years. The high heterogeneity suggests caution and the need for additional studies. We obtained our results from a large sample of 1563 untreated patients, a group as large as 40% of all patients previously studied in the literature. Compared to what Di Felice reported [10], we (1) did not have data to study infantile IS; (2) described for the first time juvenile IS with enough detail (even if we included age 10 in this group); and (3) split adolescent IS into two growth phases, with entirely different progression rates. Of note, instead of 10 years as the starting age for AIS as classically defined [3], we used age 11 after a specific check. This decision allowed us to increase the numbers in BEFORE with an improved validation process in this age group.

Lonstein and Carlson [9] reported one of the most widely cited models. They retrospectively observed 727 cases with juvenile and adolescent IS aged 12,5 years (< 10 to 19) for 25 months (12–88) and found progression in 23,2%. They followed the patients until progression or the end of growth occurred. Unfortunately, their work was never internally or externally validated. They also produced a regression equation for patients likely to progress by more than 5°, including the most important parameters, Cobb angle, EU Risser score, and age. These predictors are consistent with ours,

even if we also found a correlation with sex for AFTER only.

Many studies have reported that single lumbar (and thoracolumbar) curves had a lower risk of progression [10]. Further, another study described the expected curve progression in decreasing order by types: triple 9,1°/yr, thoracic 8,7°/yr, double 6,8°/yr, lumbar 3,8°/yr, and thoracic-lumbar 3,7°/yr [13]. Similarly, we found that single lumbar and thoracolumbar curves progressed less than the others, even though the final models did not retain curve type as a predictor because it did not improve our results.

Recently, Dolan [19] developed a model on untreated adolescents from the BrAIST study (curves 20–40°, age 10–12, EU Risser 0–2 at the start) in a sub-group of 115. The authors used a mixed group to validate the model, including some BrAIST patients braced less than 6 h per day and two samples recruited at other times from Institutes that participated in the BrAIST study. They used logistic regression analysis to predict a different outcome: curve progression to 45° or surgery before EU Risser 4. Yet, their final model included the Cobb angle, curve type, and skeletal maturity evaluated according to Sanders [20]. They assessed but did not retain age, sex, and triradiate cartilage. Dolan followed up an untreated group until the end of growth. Consequently, they could consider the probability of reaching the surgical threshold as one of the primary treatment outcomes [15]. Our prediction models, with the Cobb angle as the outcome measure, address a different but highly relevant need [14]: to determine the risk of progression at a specific growth phase.

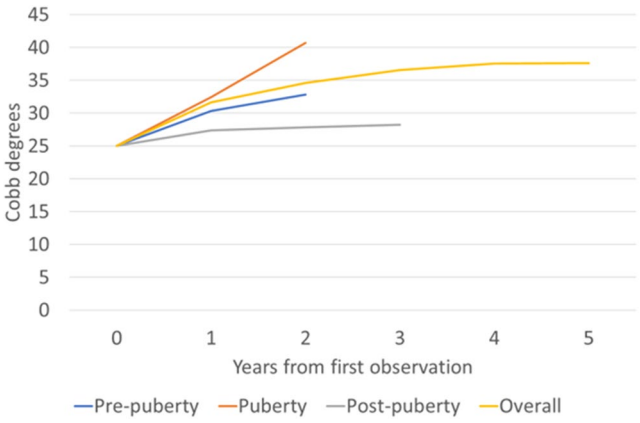

Fig. 1 Comparison between the three curve severity prediction models developed in this study from data from specific puberty periods and the model developed combining patients from the whole growth spectrum from the same research project. Hypothetical case with a starting point at 25 degrees. The length of the prediction lines for each model corresponds to the mean interval plus one standard deviation between the included radiographs. The general model slightly underestimates pubertal progression and overestimates all the other ages – the higher prevalence of pubertal cases could explain this result. Nevertheless, the overall model gives a more extended prediction due to the longer intervals between included radiographs

In our previous study [12], we derived and validated a prediction model from the whole sample of 2317 patients, compared to 1563 in this secondary analysis. The difference is due to the exclusion of subjects with pairs of radiographs in different puberty groups. While reducing the heterogeneity, the subgrouping procedure also reduced the sample size. This could affect the reliability of the analysis, considering we kept the same number of predictors. The increased accuracy of the current models could be due to the decrease in heterogeneity or the reduced number of cross-validations with the risk of overfitting [12]. Moreover, our age-dependent models cover a shorter prediction timespan than the overall model: 2–3 years for puberty subgroups instead of 5 years for the OVERALL model (Fig. 1). Therefore, although this strategy enhanced the accuracy and clinical usefulness, we are also aware of the potential risk of overfitting. External validation will help determine the clinical value of the equations.

The differences among the three growth periods correspond to clinical experience and previous research. Scoliosis can manifest in various curve types, curve magnitude and clinical characteristics. The present design could consider only radiographic parameters, age, and sex. The high variance we found demonstrated the complexity of scoliosis. We highlighted the need for further studies in this field with larger samples to account for additional potential features predicting future scoliosis severity. Future models developed in larger cohorts with internal and external validation are desirable.

Clinical and educational use of the models

The proposed models can be helpful clinically and for educational purposes. Clinicians can estimate the progression anticipated for each individual with IS. For example, for a girl, 12 years of age, with EU Risser 0 and a 32° primary curve, the expected Cobb angles after 1 and 2 years would be:

- **AT regression equation**: -1,01+(°Cobb×1,12)+(years between observations×6,88)-((years between observations)^2×1,81)+((years between observations)^3×0,35)-1,85(if EU Risser 1)-2,53(if EU Risser 2).
- **Prediction at 1 year**: -1,01+(32×1,12)+(1×6,88)-(1^2×1,81)+(1^3×0,35)-1,85×0–2,53×0= -1,01+35,84+6,88−1,81+0,35=**40,25°**.
- **Prediction at 2 years**: -1,01+(32×1,12)+(2×6,88)-(2^2×1,81)+(2^3×0,35)-1,85×0–2,53×0 = -1,01+35,84+13,76−7,24+2,8=**44,15°**.

The probability that the prediction is correct depends on the accuracy reported in Table 4 for each age group. Table 5

Table 5 Probability of and the expected curve severity after 1 and 2 years for a girl, 12 years of age, European Risser 0, with a 32° primary curve

Progression		
Probability	°Cobb at 1 year	°Cobb at 2 years
68%	35°-45°	39°-49°
90%	30°-50°	34°-54°
97%	25°-55°	29°-59°

reports the likelihood of progression for this hypothetical case: this is also an example of how these data could be presented to the patient and their family to make them understand the risks of progression in their specific case.

Clinicians could base the discussion with the patient and parents on what is reported in Table 5 in the appendix. In this specific case, both predicted future severities are important and would help the clinician in counselling the family to consider effective, aggressive, conservative treatments [3, 21]. Moreover, clinicians could consider combining this prediction model with the BrAIST one [19] to support counselling and final informed decisions as they focus on different outcomes. Clinicians should limit individual predictions based on our models to two years for BEFORE and AT and three years for AFTER. There are two main reasons for this recommended limitation: the average observation time between the pairs of radiographs of the participants included the development of each model (Table 2); the longer times would bring the prediction outside the puberty phases of the models.

Another important use of our results is developing educational tools to help patients, families, and students understand scoliosis severity over time. One such graph shows the potential evolution throughout all growth of three curves identified at age seven and for three cases starting with the same Cobb angle but at different ages (Fig. 2). We used a series of predictions to develop these graphs where the end of one prediction became the starting point for the subsequent prediction. According to the results, we used the projections over two years for BEFORE and AFTER and over one year for AT. Logically, this approach amplifies the prediction error and decreases the confidence in results as the prediction interval increases. Consequently, such graphs should not be used to predict single-patient results. Still, they can serve as an educational tool to explain the natural history of IS and help our overall understanding.

Strengths and limitations

This study represents, the largest cohort for prediction within specific age subgroups. We performed a very accurate case selection process with quality checks. We proposed a new approach to studying IS natural history because it would be unethical to complete a prospective follow-up of untreated patients since we would have to withhold treatment now proven effective even when it would become indicated [3, 21]. This type of study would benefit from collaborative data collection in the future, i.e., involving multiple centres in collecting radiographs of patients who arrive at their observation without previous treatment. Other possibilities could include patients refusing treatments (but acknowledging they may have characteristics leading to poorer outcomes) or using populations without access to treatment.

There are some limitations. The observational retrospective data collection did not allow gathering information other than radiographic ones. The missing data about patients' real growth spurt and height when radiographs were taken is particularly important: the differing typical growth velocity in the three considered phases would have added justification for using the three skeletal maturation stages we employed. The study only examined radiographic predictors, while other essential predictors [9] were unavailable in our database. Many patients had short observation times in each model, which was unavoidable because of the design. Splitting the populations according to the growth period reduced the number of participants in each model, thus increasing the risk of overfitting. The models have a short observation time, implying that the prediction also has specific time limitations. Because we included only patients who came at a second observation, we may overestimate the probability of increases in curve severity. Nevertheless, this is also a clinical reality since scoliosis is frequently not diagnosed until it reaches a visible threshold. Moreover, not

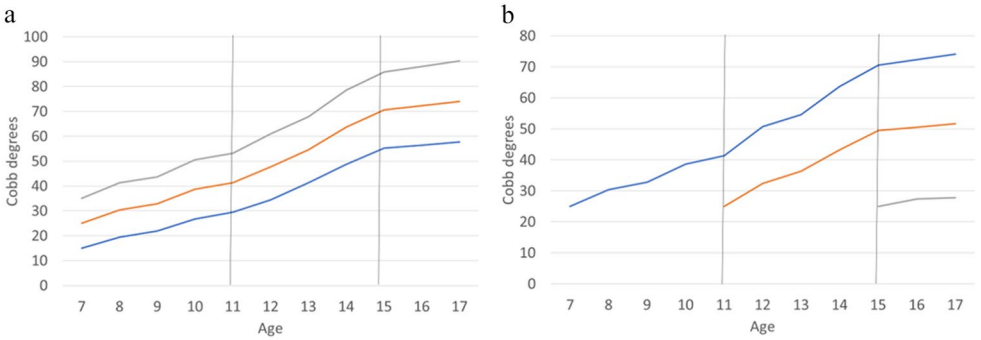

Fig. 2 Two hypothetical overall curve severity evolution models for idiopathic scoliosis until the end of growth. A: three patients with scoliosis identified at age seven with curves of 15°, 25° or 35°. B: three other patients with a scoliosis curve of 30° identified at age 6, 11 (EU Risser 0) or 14 (EU Risser 3 - female)

all the patients we observed progressed during the observation period.

Conclusion

Understanding the prognosis of scoliosis is essential for clinical decision-making. Our study produced accurate curve severity prediction models tailored to growth spurt periods, easy to use in clinics, and useful for education and research. Future collaborative efforts, possibly combining clinical and radiographic data, could allow the derivation and validation of more accurate clinical prediction rules to improve prediction accuracy further and explain a larger portion of the variance.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00586-024-08487-0.

Acknowledgements This research has been facilitated by the Women and Children's Health Research Institute through the Stollery Children's Hospital Foundation and was supported and founded by the Italian Ministry of Health – Ricerca Corrente. We acknowledge the work of all ISICO physicians for their routine systematic data collection and their contribution to reviewing their patients' charts for this research project. We wish to thank Michela Ciminello, Matteo Matrantonio, Silvana Parzini, Martina Poggio, Valentina Premoli, Marco Rampi, Sara Rossi Raccagni, Luca Selmi, and Massimiliano Vanossi for their help in checking and reviewing all data in the patients' charts. We also thank Francesca Di Felice for helping in the funding acquisition.

Author contributions Conceptualization: Stefano Negrini, Eric C. Parent, Sabrina Donzelli, Alberto Negrini, Fabio Zaina; Methodology: Stefano Negrini, Eric C. Parent, Sabrina Donzelli, Alberto Negrini, Fabio Zaina, Maryna Yaskina; Data check: Giulia Rebagliati, Sabrina Donzelli; Formal analysis and investigation: Maryna Yaskina; Writing - original draft preparation: Stefano Negrini; Writing - review and editing: Sabrina Donzelli, Marya Yaskina, Alberto Negrini, Giulia Rebagliati, Claudio Cordani, Fabio Zaina, Eric Parent; Funding acquisition: Stefano Negrini, Sabrina Donzelli, Alberto Negrini, Fabio Zaina, Eric Parent. Final approval of the text: all authors.

Funding Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement.

Declarations

Ethical approval The local Ethics Committee for the authors in Italy (Parere 801_2015 bis) and Ethic Review Board for the authors in Canada (Pro00108737) approved the study.

Conflict of interest This project was funded by a Standard Research grant from the Scoliosis Research Society. SN and AN own stock of ISICO.

References

1. Turabian J (2017) A narrative review of natural history of diseases and Continuity of Care in Family Medicine. Archives Community Med Public Health 041–047. https://doi.org/10.17352/2455-5479.000023
2. The value of the study of natural history in genetic disorders and congenital anomaly syndromes. - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1050520/. Accessed 5 Sep 2022
3. Negrini S, Donzelli S, Aulisa AG et al (2018) 2016 SOSORT guidelines: orthopaedic and rehabilitation treatment of idiopathic scoliosis during growth. Scoliosis Spinal Disord 13. https://doi.org/10.1186/s13013-017-0145-8
4. Burwell RG, James NJ, Johnson F et al (1983) Standardised trunk asymmetry scores. A study of back contour in healthy school children. J Bone Joint Surg Br 65:452–463. https://doi.org/10.1302/0301-620X.65B4.6874719
5. Brooks HL, Azen SP, Gerberg E et al (1975) Scoliosis: a prospective epidemiological study. J Bone Joint Surg Am 57:968–972
6. Grivas TB, Vasiliadis E, Mouzakis V et al (2006) Association between adolescent idiopathic scoliosis prevalence and age at menarche in different geographic latitudes. Scoliosis 1:9. https://doi.org/10.1186/1748-7161-1-9
7. Bunnell WP (1986) The natural history of idiopathic scoliosis before skeletal maturity. Spine (Phila Pa 1976) 11:773–776. https://doi.org/10.1097/00007632-198610000-00003
8. Nachemson AL, Peterson LE (1995) Effectiveness of treatment with a brace in girls who have adolescent idiopathic scoliosis. A prospective, controlled study based on data from the Brace Study of the Scoliosis Research Society. J Bone Joint Surg Am 77:815–822. https://doi.org/10.2106/00004623-199506000-00001
9. Lonstein JE, Carlson JM (1984) The prediction of curve progression in untreated idiopathic scoliosis during growth. J Bone Joint Surg Am 66:1061–1071
10. Di Felice F, Zaina F, Donzelli S, Negrini S (2018) The natural history of idiopathic scoliosis during growth: a Meta-analysis. Am J Phys Med Rehabil 97:346–356. https://doi.org/10.1097/PHM.0000000000000861
11. Weinstein SL, Dolan LA, Wright JG, Dobbs MB (2013) Effects of bracing in adolescents with idiopathic scoliosis. N Engl J Med 369:1512–1521. https://doi.org/10.1056/NEJMoa1307337
12. Parent EC, Donzelli S, Yaskina M et al (2023) Prediction of future curve angle using prior radiographs in previously untreated idiopathic scoliosis: natural history from age 6 to after the end of growth (SOSORT 2022 award winner). Eur Spine J 32:2171–2184. https://doi.org/10.1007/s00586-023-07681-w
13. Duval-Beaupere G (1996) Threshold values for supine and standing Cobb angles and rib hump measurements: prognostic factors for scoliosis. Eur Spine J 5:79–84. https://doi.org/10.1007/BF00298385
14. Negrini S, Hresko TM, O'Brien JP et al (2015) Recommendations for research studies on treatment of idiopathic scoliosis: Consensus 2014 between SOSORT and SRS non-operative

management committee. Scoliosis 10:8. https://doi.org/10.1186/s13013-014-0025-4
15. Richards BS, Bernstein RM, D'Amato CR, Thompson GH (2005) Standardization of criteria for adolescent idiopathic scoliosis brace studies: SRS Committee on Bracing and Nonoperative Management. Spine (Phila Pa 1976) 30:2068–2075 discussion 2076–2077. https://doi.org/10.1097/01.brs.0000178819.90239.d0
16. Ponseti IV, Pedrini V, Wynne-Davies R, Duval-Beaupere G (1976) Pathogenesis of scoliosis. Clin Orthop Relat Res 268–280
17. Troy M, Miller P, Price N et al (2019) The Risser+ grade: a new grading system to classify skeletal maturity in idiopathic scoliosis. Eur Spine J 28. https://doi.org/10.1007/s00586-018-5821-8
18. Negrini S, Negrini A, Santambtogio GC, Sibilla P Relation between static angles of the spine and a dynamic event like posture: Approach to the Problem. Stud Health Technol Inform 209–214
19. Dolan LA, Weinstein SL, Abel MF et al (2019) Bracing in adolescent idiopathic scoliosis trial (BrAIST): development and validation of a Prognostic Model in untreated adolescent idiopathic scoliosis using the simplified skeletal Maturity System. Spine Deform 7:890–898e4. https://doi.org/10.1016/j.jspd.2019.01.011
20. Sanders JO, Khoury JG, Kishan S et al (2008) Predicting scoliosis progression from skeletal maturity: a simplified classification during adolescence. J Bone Joint Surg Am 90:540–553. https://doi.org/10.2106/JBJS.G.00004
21. Negrini S, Minozzi S, Bettany-Saltikov J et al (2015) Braces for idiopathic scoliosis in adolescents. Cochrane Database Syst Rev CD006850. https://doi.org/10.1002/14651858.CD006850.pub3