IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 25, Issue 8 Ser. 3 (August. 2026), PP 46-50
www.iosrjournals.org

Artificial Intelligence in the Prediction of Skeletal Growth and the Optimization of Orthopedic and Orthodontic Treatment Timing: Closing the Texas Medicaid Access Gap — A Narrative Review

Natalia Cristina Gomes Temponi[1]
(Specialist in Orthodontics, FACOP, São Paulo, SP, Brazil) ORCID: 0009-0006-4396-1831

Michelle Petegrosso[2]
(Specialist in Pediatric Dentistry, Associação Paulista de Cirurgiões Dentistas - APCD, Santo André, SP, Brazil) ORCID: 0009-0003-1159-3171

Ingrid Graf Vianna Salles[3]
(Specialist in Orthodontics, Associação Brasileira de Ensino Odontológico – ABENO, São Paulo, SP, Brazil) ORCID: 0009-0005-8100-4904

Andreza Calazans Rodrigues[4]
(Specialist in Endodontics, INCO 25, Niterói, RJ, Brazil) ORCID: 0009-0008-1295-2312

André Luiz Girotto[5]
(Specialist in Functional Jaw Orthopedics, CFO – Resolution CFO 25/2002; , São Paulo, SP, Brazil) ORCID: 0009-0007-7427-8733

Paula Cristina da Costa Alcantara[6]
(PhD Student in Orthodontics, Rio de Janeiro State University - UERJ, Master`s Degree in Pediatric Dentistry, Federal University of Rio de Janeiro - UFRJ, Rio de Janeiro, RJ, Brazil) ORCID: 0009-0001-8288-6479

Rafael Rudimar Holler Peres[7]
(Master`s Degree in Orthodontics, University Positivo; Curitiba, PR, Brazil) ORCID: 0009-0008-8615-9835

Sabrina dos Santos Carvalho Soares[8]
(Specialist in Orthodontics, Faculdade Unidas do Norte de Minas – FUNORTE / Instituto de Ciências da Saúde – ICS, Manaus, Amazonas, Brazil) ORCID: 0009-0000-1026-6361

Abstract
Background: The success of dentofacial orthopedic and growth-modifying orthodontic interventions depends on identifying the correct biological window relative to the pubertal growth spurt (PGS), since chronological age is an unreliable proxy for skeletal maturity. Artificial intelligence (AI), particularly deep learning, has emerged over the past decade as a tool to automate and refine skeletal maturation assessment. Materials and Methods: A narrative review of the literature (1982-2026) was performed using PubMed, Web of Science, and Google Scholar, focusing on classic and AI-based methods for cervical vertebral maturation (CVM) and hand-wrist skeletal age estimation, sex-related biological and performance differences, and AI-assisted treatment-planning optimization. Results: Deep-learning models, especially convolutional neural networks, achieve pooled diagnostic accuracies around 80-95% for staging skeletal maturation from cephalometric and hand-wrist radiographs, frequently matching or exceeding inter-observer agreement among clinicians. Sex is both a well-established biological modifier of maturation timing and a significant performance modifier for AI models: girls consistently show earlier cervical vertebral maturation than boys, and several deep-learning systems show higher accuracy and lower error when assessing female patients, while also being able to classify sex directly from hand-wrist radiographs with an accuracy exceeding that of expert human observers. Reinforcement-learning platforms are now being explored to move beyond diagnosis toward individualized, multi-objective treatment-timing

optimization. A missed or mistimed growth window carries direct financial consequences in the U.S. fee-for-service dental economy, converting an interceptive-treatment candidate into a surgical candidate at several times the cost; this differential intersects with documented dental-access shortages in states such as Texas, where a large Medicaid-enrolled pediatric population coincides with extensive dental Health Professional Shortage Area designations. Conclusion: AI substantially improves the speed, reproducibility, and in some respects the accuracy of skeletal maturation assessment, and may offer particular economic value in access-constrained, Medicaid-dependent settings, but should currently be regarded as a decision-support tool that requires clinical supervision, particularly given persistent sex-related performance variability during the pubertal window most relevant to treatment timing.

Key Word: *Artificial intelligence; skeletal maturation; cervical vertebral maturation; bone age; deep learning; pubertal growth spurt; orthodontics; dentofacial orthopedics; sexual dimorphism.*

Date of Submission: 09-08-2026 Date of Acceptance: 21-08-2026

I. Introduction

The clinical success of functional and orthopedic appliances used to correct sagittal and vertical skeletal discrepancies is strongly time-dependent. Patients treated before, during, or after the pubertal growth spurt (PGS) respond very differently to the same biomechanical protocol, which is why identifying the optimal treatment window has long been considered one of the central problems of dentofacial orthopedics.

Chronological age is a poor surrogate for biological maturity because of wide inter-individual variability in the onset, intensity, and duration of the growth spurt. To overcome this limitation, clinicians have historically relied on skeletal maturation indicators derived from hand-wrist radiographs [1] or, more recently, from cervical vertebral maturation (CVM) staging on a single lateral cephalogram [2], the latter avoiding the additional radiation exposure required by a dedicated hand-wrist film.

Over the last decade, artificial intelligence (AI) - and deep learning in particular - has been increasingly applied to automate, standardize, and potentially improve the accuracy of skeletal maturation assessment [3]. This narrative review summarizes current evidence on (i) the classic methods used to assess skeletal maturation, (ii) the performance of AI models in predicting skeletal age and the pubertal growth peak, (iii) the biological and AI-detectable differences between male and female patients, and (iv) emerging applications of AI in the optimization of individualized orthopedic and orthodontic treatment planning.

II. Skeletal Maturation Assessment: From Hand-Wrist Radiographs to Cervical Vertebral Maturation

The hand-wrist radiograph remains a historical reference standard for skeletal age assessment. Fishman's Skeletal Maturity Indicators (SMI), based on six anatomical sites in the thumb, third finger, fifth finger, and radius, established a clinically oriented, four-stage maturation system that correlates closely with the pubertal growth spurt [1].

Because hand-wrist films require a separate radiographic exposure in addition to the cephalogram routinely obtained in orthodontic diagnosis, cervical vertebral maturation (CVM) staging was developed as a single-cephalogram alternative. Baccetti, Franchi and McNamara refined the CVM method into six stages (CS1-CS6) based on the morphology of the second, third, and fourth cervical vertebrae, demonstrating that the mandibular growth peak occurs between CS3 and CS4 [2]. Grave and Townsend, studying a longitudinal cohort, similarly related CVM ossification events to peak stature and mandibular growth velocity, concluding that CVM stage 2 is generally an appropriate time to begin orthopedic treatment in boys, whereas an earlier, more cautious approach is warranted in girls [10]. This method has since become the most widely used clinical tool for timing orthopedic intervention, despite documented difficulty in reproducibility even among experienced examiners [3].

III. Artificial Intelligence Models for Skeletal Age and Pubertal Growth Peak Prediction

Convolutional neural networks (CNNs) and other deep-learning architectures have been trained to automate both CVM staging and hand-wrist skeletal age estimation, generally reducing analysis time and inter-observer variability compared to manual assessment [3].

A 2026 systematic review and meta-analysis pooling 21 studies (2020-2025) reported a pooled accuracy of 0.83 (95% CI 0.76-0.89) for AI models identifying the pubertal growth spurt from cephalometric and hand-wrist radiographs, with hand-wrist-based models showing higher sensitivity and cephalometric models showing higher specificity [4]. A separate AI system directly correlating cervical vertebral morphology with mandibular growth slope, rather than relying on conventional stage naming, was developed and validated using data from the American Association of Orthodontists Foundation growth centers [5].

Comparative work evaluating five different hand-wrist maturation methods with deep learning found that model choice and reference method both materially affect diagnostic reliability, underscoring that AI performance cannot be generalized across all maturation indices [6]. In orthopedic surgery, a ConvNeXt-based deep-learning model trained on more than 20,000 hand radiographs across three public datasets achieved a mean absolute error of 3.68 months on held-out test data, outperforming previously published models by over 12% [7]. Earlier work specifically automating Fishman's SMI achieved a prediction accuracy of 0.772 using a hybrid system validated on 711 hand-wrist radiographs from an external institution [8].

More recently, multimodal large language models (LLMs) have been tested directly against the Gilsanz-Ratib atlas: ChatGPT-5 achieved a mean absolute error of 1.46 years and was judged clinically useful, while Gemini 2.5 Pro, Grok-3, and Claude 4 Sonnet showed substantially higher error rates, indicating that general-purpose LLMs are not yet interchangeable with purpose-built diagnostic CNNs for this task [9].

IV. Sex-Related Differences in Skeletal Maturation and AI Performance

Sexual dimorphism in skeletal maturation is well established clinically. A systematic review and meta-analysis of 41 studies (9867 participants) using the Baccetti CVM method found a statistically significant difference between sexes at every CVM stage, confirming that girls consistently present earlier skeletal maturation than boys across all six stages [11]. This biological asymmetry appears to be both detectable and consequential for AI model performance.

In a landmark study, a deep convolutional neural network trained to predict patient sex directly from hand-wrist radiographs achieved 95.9% accuracy, compared with only 58% and 46% accuracy for two experienced human radiologists, indicating that hand and wrist bones contain sex-discriminating information imperceptible to the human eye [12]. The model's class activation maps localized attention to the second and third metacarpal base and thumb sesamoid in females, and to the distal radioulnar joint and third metacarpophalangeal joint in males [12]. A more recent two-stage forensic deep-learning system, combining anatomical pose estimation with a multi-task classifier trained on over 13,000 hand-wrist radiographs, likewise achieved simultaneous age and sex estimation from the hand skeleton alone, without relying on regions of more pronounced sexual dimorphism such as the pelvis or skull [13].

Beyond sex classification, several bone-age regression models show sex-stratified accuracy differences. A multi-modal deep-learning model integrating hand radiographs with gender and chronological age reported markedly higher accuracy in females (95.0% in patients <=11 years; 92.4% in patients >11 years) than in males (79.7% in patients <=13 years; 84.6% in patients >13 years) [14]. Similarly, an evaluation of multimodal LLM-based bone age estimation found excellent performance in pre-pubertal patients of both sexes and in mid-pubertal females, but a clear decline in mid-pubertal males, which the authors attributed to sex-specific differences in the timing of pubertal skeletal change [9].

Taken together, these findings indicate that AI models trained without explicit sex stratification may systematically under- or over-estimate skeletal maturity in male patients during the pubertal window - precisely the period most relevant to orthopedic treatment timing - reinforcing the recommendation that sex be incorporated as an explicit input variable in AI-assisted growth prediction tools [14].

V. AI-Driven Optimization of Treatment Planning and Timing

Beyond diagnostic staging, AI is increasingly being positioned as a tool for optimizing the treatment plan itself rather than only estimating skeletal age. A multi-task reinforcement learning platform integrated with explainable AI has been proposed for orthodontic-orthognathic decision support, treating treatment planning as a sequential decision-making process that simultaneously optimizes multiple clinical objectives, such as skeletal correction, dental alignment, and facial esthetics, rather than a single static endpoint; a multicenter prospective validation involving 200 patients was planned for 2025-2026 [15].

This shift - from AI as a diagnostic classifier toward AI as a planning and simulation partner - mirrors the broader trajectory of the field: once skeletal age and growth-peak timing can be estimated reliably and sex-appropriately [4,11,14], the same underlying data (cephalometric landmarks, hand-wrist maturation stage, chronological age, sex) can, in principle, feed multi-objective optimization models that recommend not only whether to treat, but which appliance, at which stage, and within which sex-specific biological window is likely to maximize the orthopedic response [15].

VI. Limitations and Clinical Considerations

Despite consistently promising accuracy metrics, several limitations temper the immediate clinical substitution of AI for expert judgment. Model performance depends heavily on training-dataset size, labeling consistency, and population representativeness, and most published models remain validated primarily on single-institution or single-ethnicity cohorts [3]. Because most AI systems are trained using expert-assigned labels as the reference standard, they inherit - rather than eliminate - the inter-observer variability present in the original clinical judgments used to train them [3].

The sex-related performance asymmetry discussed in Section IV represents a specific and clinically important source of potential error, particularly for male patients evaluated near the pubertal growth peak, when treatment-timing decisions carry the greatest consequence [9,14]. Reinforcement-learning platforms for treatment-plan optimization, while conceptually powerful, remain at an early validation stage, with prospective multicenter outcome data still pending [15]. Current consensus in the orthodontic AI literature is therefore that these tools should be used as decision-support aids under expert supervision rather than as autonomous diagnostic or planning systems [3].

VII. Economic and Financial Significance in the U.S. Context: The Case of Texas

The timing accuracy discussed in the preceding sections is not only a biomechanical variable but a substantial cost variable within the U.S. fee-for-service dental economy. Market-reported U.S. price ranges place Phase I (interceptive) orthodontic treatment at approximately $2,000-4,500 and comprehensive adolescent orthodontic treatment at approximately $5,000-7,000, whereas combined orthodontic-orthognathic surgical correction - the outcome most likely when the pubertal growth window is missed and a purely dentoalveolar or growth-modifying approach is no longer viable - typically ranges from $20,000-40,000. A mistimed or undetected growth window can therefore convert a patient from an interceptive-treatment candidate into a surgical candidate at a cost multiple of five- to fifteen-fold, a differential that falls directly on families, employers, and public payers rather than on the treating clinician alone.

Texas illustrates how this cost differential intersects with documented access constraints. Texas has one of the largest state Medicaid-enrolled pediatric dental populations in the country while simultaneously ranking among the states with the most dental Health Professional Shortage Area (HPSA) designations in the nation; recent HRSA-based data indicate that roughly one-third of Texas residents live in a county with a geographic dental shortage designation. A national analysis of the congruence between county-level dental HPSA designations and the CDC Social Vulnerability Index specifically identified the Texas-Mexico border region, together with Alaska, as areas where whole-county dental shortage designations and high social vulnerability most consistently overlap [16], meaning that some of the state's most Medicaid-dependent pediatric populations also have the least access to the serial orthodontic monitoring visits traditionally used to identify the CVM window. An AI system capable of estimating growth-stage timing from a single cephalometric radiograph already obtained for routine orthodontic diagnosis - rather than requiring additional specialist visits - could therefore offer disproportionate value precisely in the shortage counties where repeated in-person monitoring is least feasible.

Beyond timing itself, Texas Medicaid - like most state Medicaid dental programs - authorizes orthodontic coverage using the Handicapping Labio-Lingual Deviation (HLD) Index, an objective point-based scoring system requiring a minimum of 26 points for approval. However, studies evaluating HLD-based Medicaid coverage decisions in another large state program (Illinois) found that managed care organization coverage decisions did not consistently agree with HLD scores independently calculated from digital models or visual inspection [17], and that a 2017 change in HLD implementation was associated with a significant decrease in orthodontic access for Medicaid-enrolled children [18]. To the extent that AI-based, reproducible skeletal and occlusal staging could be incorporated into the documentation submitted for Medicaid prior authorization, it could reduce this scoring inconsistency, potentially lowering the administrative cost of denials, appeals, and resubmissions for both providers and state Medicaid programs, while improving the consistency of coverage decisions for Medicaid-enrolled children.

Finally, because several deep-learning growth-stage models reviewed in Section IV show measurably lower accuracy in male patients during the pubertal window, the financial risk of a missed or mistimed growth window - and the resulting cost escalation toward orthognathic surgery - may fall disproportionately on boys unless sex-specific calibration is built into any AI system deployed in Medicaid-funded or safety-net dental settings. Given the size of Texas's Medicaid-enrolled pediatric population and its concentration in dental HPSA counties, this consideration has direct relevance to how such tools should be validated and deployed within the state before broader adoption.

VIII. Conclusion

Artificial intelligence has demonstrated substantial value in automating and standardizing the assessment of skeletal maturation, whether through cervical vertebral maturation staging on cephalograms or hand-wrist radiograph analysis, frequently matching or exceeding human inter-observer agreement. A clinically important and recurring finding across independent studies, consistent with decades of classic auxological literature, is that skeletal maturation timing - and consequently AI model accuracy - is not uniform between sexes, with girls maturing earlier and several deep-learning systems performing measurably better in female than in male patients during the pubertal window most relevant to orthopedic treatment timing. Emerging reinforcement-learning approaches point toward a future in which AI contributes not only to diagnosing the growth stage but to actively optimizing the individualized treatment plan and its timing. Until larger, sex-stratified, and prospectively validated datasets are available, AI-based growth prediction should be integrated into clinical practice as a supportive, not a substitutive, decision-making tool.

References

[1]. Fishman LS. Radiographic evaluation of skeletal maturation: a clinically oriented method based on hand-wrist films. Angle Orthod. 1982;52(2):88-112.
[2]. Baccetti T, Franchi L, McNamara JA Jr. The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics. Semin Orthod. 2005;11(3):119-129.
[3]. Farzanegan P, Zarabadi M, Najary S, Tahmasbi MA, Behnaz M. The role of artificial intelligence in orthodontics for determining skeletal age based on cervical vertebra maturation degree: a comprehensive review. Health Sci Rep. 2025;8(11):e71487.
[4]. Chaves J, Ruffini M, Souza A, Moriguchi C, Ribeiro DH, Moura D, Maranhão G, Gomes G, Leal ME, Duarte J. Artificial intelligence for predicting the pubertal growth spurt using cephalometric and hand-wrist radiographs: a systematic review and meta-analysis. BMC Oral Health. 2026;26(1):1116.
[5]. Alipour Shoari S, Sadrolashrafi SV, Sohrabi A, Afrouzian R, Ebrahimi P, Kouhsoltani M, Kouh Soltani M. Estimating mandibular growth stage based on cervical vertebral maturation in lateral cephalometric radiographs using artificial intelligence. Prog Orthod. 2024;25(1):18.
[6]. Tentaş S, Özden S. Deep learning based evaluation of skeletal maturation: a comparative analysis of five hand-wrist methods. Orthod Craniofac Res. 2025. doi:10.1111/ocr.70008.
[7]. Bram JT, Pareek A, Beber SA, Jones RH, Shariatnia MM, Daliliyazdi A, Tracey OC, Green DW, Fabricant PD. Determination of skeletal age from hand radiographs using deep learning. Am J Sports Med. 2025;53(11):2715-2725.
[8]. Kim H, Kim CS, Lee JM, Lee JJ, Lee J, Kim JS, Choi SH. Prediction of Fishman's skeletal maturity indicators using artificial intelligence. Sci Rep. 2023;13(1):5870.
[9]. Ozkan E, Koyun M. Atlas-assisted bone age estimation from hand-wrist radiographs using multimodal large language models: a comparative study. Diagnostics (Basel). 2026;16(3):487.
[10]. Grave K, Townsend G. Cervical vertebral maturation as a predictor of the adolescent growth spurt. Aust Orthod J. 2003;19(1):25-32.
[11]. Magalhães MI, Machado V, Mascarenhas P, Botelho J, Mendes JJ, Delgado AS. Chronological age range estimation of cervical vertebral maturation using Baccetti method: a systematic review and meta-analysis. Eur J Orthod. 2022;44(5):548-555.
[12]. Yune S, Lee H, Kim M, Tajmir SH, Gee MS, Do S. Beyond human perception: sexual dimorphism in hand and wrist radiographs is discernible by a deep learning model. J Digit Imaging. 2019;32(4):665-671.
[13]. Çatak C, Ocak M, Çolak D. A two-stage deep learning approach for forensic age and sex determination from hand-wrist radiographs: an anatomical pose estimation and multi-task model. Int J Legal Med. 2026.
[14]. Kim JK, Park D, Chang MC. Assessment of bone age based on hand radiographs using regression-based multi-modal deep learning. Life (Basel). 2024;14(6):774.
[15]. Li Z, Wang L. Multi-task reinforcement learning and explainable AI-driven platform for personalized planning and clinical decision support in orthodontic-orthognathic treatment. Sci Rep. 2025;15:24502.
[16]. Benavidez GA, Crouch E, Nelson J, Martin A. Congruence between county dental health provider shortage area designations and the social vulnerability index. Prev Chronic Dis. 2024;21:230315.
[17]. Makki A, Elnagar MH, Sanchez F, Caplin J, Viana G, Hasan Z, Obrez A, Kusnoto B. Assessment of Handicapping Labio-Lingual Deviation index scoring methods and their effect on orthodontic treatment coverage by Medicaid. J Public Health Dent. 2022;82(4):478-483.
[18]. Caplin J, Rozgony A, Farhi M, Viana GC, Al-Grouz M, Belavsky BZ. The implementation of the Handicapping Labio-Lingual Deviation index on access to orthodontic care. J Public Health Dent. 2020;80(2):107-113.