Facultad de Odontología

Odontoestomatología
DOI: 10.22592/ode2026n47e252

Assessing the diagnostic concordance of cephalometric indicators of skeletal class and facial biotype: a statistical approach

INVESTIGATION

Evaluación de la concordancia diagnóstica de los indicadores cefalométricos de la clase esqueletal y del biotipo facial: un enfoque estadístico

Avaliação da concordância diagnóstica dos indicadores cefalométricos da classe esquelética e do biótipo facial: uma abordagem estatística

Abstract

Objectives: 1) To determine the diagnostic agreement of five cephalometric indicators of skeletal class and four of facial biotype through Cohen's weighted kappa test, and 2) to compare these results with those obtained after applying Cohen's weighted Chi-square test.
Material and methods: In a sample of 439 lateral teleradiographs, the skeletal class indicators (ANB angle, APDI, FABA angle, Wits indicator, and Beta angle) and facial biotype indicators (FMA, skull base angle, Björk–Jarabak polygon, and VERT index) were determined. Diagnostic concordance between them was assessed using Fleiss' kappa, Cohen's weighted kappa, and Cohen's weighted Chi-square tests.
Results: Statistically significant differences were observed in both skeletal class and facial biotype diagnoses. While, according to Cohen's weighted kappa test, most agreements between skeletal class and facial biotype indicators were "Fair" or "Moderate," Cohen's weighted Chi-square test revealed highly significant discordance in most comparisons, with negative values in Fleiss' kappa, indicating a systematic absence of concordance when all indicators were considered simultaneously.

$$\text{Fleiss' kappa}_{\text{skeletal class}} = -0.0714,\ z = -1.46,\ p = 0.143$$
$$\text{Fleiss' kappa}_{\text{facial biotype}} = -0.0909,\ z = -1.28,\ p = 0.201$$

The clinical relevance of these findings lies in the fact that diagnosis forms the foundation of proper clinical practice.

Conclusion: 1) When analyzing diagnostic agreement of the five skeletal class indicators and the four facial biotype indicators, significant differences were found among them – a result consistently supported by Cohen's weighted Chi-square test. 2) Regarding the statistical tests to be used in assessing the significance level of diagnostic agreement, it is recommended to consider: (i) for two evaluators without extreme discrepancies, use the "classical" Cohen's kappa test; (ii) for two evaluators with extreme discrepancies, use Cohen's weighted kappa test; and (iii) for more than two evaluators, use Cohen's weighted Chi-square test.

AUTHORS

Juan Diego Idrovo [1]
Valentina Zaffiri [1]
Alejandro Díaz [2]
Germán Manríquez [3]

CORRESPONDENCE
Germán Manríquez,
gmanriquezs@odontologia.uchile.cl

1 Cirujano Dentista, Práctica Privada, Santiago, Chile.
2 DDS, Mg.Sc., Departamento del Niño y Ortopedia Dentomaxilar, Centro de Análisis Cuantitativo en Antropología Dental, Facultad de Odontología, Universidad de Chile. Servicio de Cirugía Maxilofacial, Hospital San Borja Arriarán, Santiago, Chile.
3 Biólogo, Mg.Sc., Ph.D., Instituto de Investigación en Ciencias Odontológicas, Centro de Análisis Cuantitativo en Antropología Dental, Facultad de Odontología, y Departamento de Antropología, Grupo de Bioantropología, Facultad de Ciencias Sociales, Universidad de Chile.

Received: April 25, 2025
Accepted: October 03, 2025

Keywords
Cephalometry, Diagnostic imaging, Malocclusion, Orthodontics, Statistical analyses.

Resumen

Objetivos: 1) determinar la concordancia diagnóstica de cinco indicadores cefalométricos de clase esqueletal y cuatro de biotipo facial utilizando la prueba de kappa ponderada de Cohen, y 2) comparar estos resultados con los obtenidos después de aplicar la prueba de Ji^2 ponderada de Cohen.
Material y métodos: En una muestra de 439 telerradiografías laterales, se determinaron los indicadores de clase esqueletal (ángulo ANB, APDI, ángulo FABA, indicador de Wits y ángulo beta) y biotipo facial (FMA, ángulo de la base de cráneo, polígono de Bjork Jarabak e índice VERT). La concordancia diagnóstica entre ellos se evaluó mediante las pruebas de kappa de Fleiss, kappa de Cohen ponderado y Ji^2 ponderada de Cohen.
Resultados: Se observaron diferencias estadísticamente significativas tanto en el diagnóstico de la clase esqueletal como del biotipo facial. Mientras que en la prueba ponderada de kappa de Cohen, la mayoría de las concordancias entre los indicadores de clase esqueletal y biotipo facial fueron "Leves" o "Moderadas", la prueba de Ji^2 ponderada de Cohen mostró una discordancia altamente significativa en la mayoría de las comparaciones, con valores negativos de la prueba de kappa de Fleiss, indicativos de la ausencia sistemática de concordancia al considerar todos los indicadores a la vez.

kappa de Fleiss$_{clase\ esqueletal}$= -0.0714, z = -1.46, p-value = 0.143
kappa de Fleiss$_{biotipo\ facial}$= -0.0909, z = -1.28, p-value = 0.201

La importancia clínica de estos resultados reside en que el diagnóstico está en la base de una correcta práctica clínica.
Conclusión: 1) Al estudiar la concordancia diagnóstica de los cinco indicadores de clase esqueletal y los cuatro de biotipo facial, se encontraron diferencias significativas entre ellos, resultado avalado consistentemente con uso de la prueba estadística Ji^2 ponderada de Cohen y 2) En relación con las pruebas que se deben utilizar para conocer el nivel de significancia estadística de la concordancia diagnóstica, se recomienda tener en cuenta: i) si se trata de dos evaluadores sin discordancias extremas, utilizar la prueba kappa de Cohen "clásica", ii) de dos evaluadores con discordancias extremas, la prueba kappa ponderada de Cohen, o, iii) de más de dos evaluadores, la prueba de Ji^2 ponderada de Cohen.

Palabras clave
Cefalometría, Imágenes diagnósticas, Maloclusión, Ortodoncia, Análisis estadístico.

Resumo

Objectivos: 1) determinar a concordância diagnóstica de cinco indicadores cefalométricos de classe esquelética e quatro de biótipo facial, através do teste kappa ponderado de Cohen, e 2) comparar estes resultados com os obtidos após a aplicação do teste Qui-quadrado ponderado de Cohen.
Material e métodos: Numa amostra de 439 telerradiografias laterais, foram determinados os indicadores da classe esquelética (ângulo ANB, APDI, ângulo FABA, indicador Wits e ângulo beta) e do biótipo facial (FMA, ângulo da base do crânio, polígono Bjork Jarabak e índice VERT). A concordância diagnóstica entre eles foi avaliada através dos testes kappa de Fleiss, kappa ponderado de Cohen e qui-quadrado ponderado de Cohen.
Resultados: Foram observadas diferenças estatisticamente significativas no diagnóstico da classe esquelética e do biótipo facial. Enquanto no teste kappa ponderado de Cohen, a maioria das concordâncias entre os indicadores da classe esquelética e do biótipo facial foi "Ligeira" ou "Moderada", o teste Qui-quadrado ponderado de Cohen mostrou discordâncias altamente significativas na maioria das comparações, com valores negativos do teste kappa de Fleiss, indicando a ausência sistemática de concordância ao considerar todos os indicadores simultaneamente.

kappa de Fleiss $_{classe\ esquelética}$= -0.0714, z = -1.46, p-value = 0.143
kappa de Fleiss $_{biótipo\ facial}$= -0.0909, z = -1.28, p-value = 0.201

Conclusões: 1) Ao estudar a concordância diagnóstica de os cinco indicadores de classe esquelética e os quatro indicadores de biótipo facial, foram encontradas diferenças significativas entre eles, um resultado consistentemente apoiado pelo uso do teste estatístico qui-quadrado ponderado de Cohen, e 2) Em relação aos testes a serem utilizados para conhecer o nível de significância estatística da concordância diagnóstica, recomenda-se levar em conta: (i) para dois avaliadores sem discordância extrema, utilizar o teste kappa de Cohen "clássico", (ii) para dois avaliadores com discordância extrema, o teste kappa de Cohen ponderado, ou, (iii) para mais de dois avaliadores, o teste Qui-quadrado de Cohen ponderado.

Palavras-chave
Cefalometria, Imagens diagnósticas, Má oclusão, Ortodontia, Análise estatística.

Introduction and background

Skeletal class is defined as the sagittal relationship between the maxilla and the mandible[1] and is divided into three categories: Class I, Class II, and Class III. In addition, facial biotype is defined as the set of morphodifferential characteristics that determine the vertical component of craniofacial morphology and allow for an understanding of the functional behaviors and specific therapeutic needs of each biotype.[2–4] Three types are described in the literature: brachyfacial (or hypodivergent), mesofacial (or normodivergent), and dolichofacial (or hyperdivergent).[5–7]

In both cases (i.e. skeletal class and facial biotype), cephalometrics –a tool based on the measurement of linear and angular distances in cranial teleradiographs– is routinely used for their determination. Although a wide variety of cephalometric indicators have been proposed to classify skeletal class and facial biotype, there is no universally accepted or "gold standard" indicator.[8–14] For clinical purposes, these indicators are expected to show a high level of diagnostic concordance, so that the results obtained remain consistent regardless of the indicator used. Another aspect to consider is that, over time, new indicators have been proposed to overcome the limitations of previous ones. For example, regarding skeletal class, Jacobson suggested that the ANB angle may be misinterpreted depending on the position of the cephalometric Nasion point,[9,15] recommending the use of Downs points A and B in relation to the occlusal plane for sagittal evaluation of the maxillomandibular relationship.[9] However, high variability and occasional difficulty in determining this latter method due to the occlusal plane have also been reported.[16] Meanwhile, the creators of the FABA angle[17] found high variability in the APDI and Wits indicators among normal occlusions, questioning their validity and suggesting the FABA angle as preferable to other indicators. Finally, according to the authors of the Beta angle, the advantage of their proposal is that it does not depend on cranialreferences, occlusal stability, or mandibular rotation.[8]With respect to facial biotype, the following indicators have been proposed: the summation of the Björk–Jarabak polygon (JBK),[12] the Frankfurt–mandibular plane angle (FMA),[13] the skull base angle (BSCR)[11] and mandibular plane (SN-GoGn),[14] and the Ricketts VERT.[18]

Due to this large number of proposals, each indicator represents a partial and independent solution, diverging from an ideal clinical concordance.

Regarding the evaluation of diagnostic agreement between different evaluators (indicators), Cohen's kappa test[19] has been traditionally used since its development.[20] Despite its popularity, this test does not account for extreme discrepancies and is recommended only when no significant differences between assessors (indicators) are observed. Taking this into consideration, Cohen subsequently proposed a statistical test that accounted for the weight of extreme discordances,[21] as well as a weighted Chi-square test[22] for cases involving more than two assessors (indicators). Given that the level of concordance between the aforementioned cephalometric indicators is low,[23–25] the present study aims to: 1) determine the diagnostic concordance between them using the weighted kappa test,[21] and 2) compare these results with those obtained after applying the weighted Chi-square test,[22] thereby contributing to the improvement of clinical practice in orthodontics.

Materials and Methods

A sample consisting of 439 lateral teleradiographs (275 females and 164 males) was obtained from an anonymized database of patients over 18 years of age, with no history of orthodontic, orthopedic, or maxillofacial surgical treatment, and with adequate image resolution for the identification of cephalometric points. For each radiograph, the following skeletal class indicators were determined: ANB angle, APDI, FABA angle, Beta angle, and Wits indicator; the latter grouping the sex variable due to the absence of statistically significant differences between females and males (Chi-square test: $\chi^2 = 0.313$, p (no association) = 0.855, df = 2) (**Table 1**, **Figure 1**), a finding recently corroborated by other authors.[26] Regarding facial biotype, the following indicators were used: the summation of the Björk–Jarabak polygon (JBK),[12] the Frankfurt–mandibular plane angle (FMA),[13] the skull base angle and mandibular plane (BSCR),[14] and the Ricketts VERT[18,27] (**Table 2**, **Figure 2**). All these indicators were determined using the WebCeph V1.5.0 platform.[10]

Table 1. Definition of skeletal class indicators used in this study and their respective classification values. (A = Point A, B = Point B, C = Center of the condyle)

Indicator	Definition	Class I	Class II	Class III	References
ANB	Angle formed by the intersection of the lines Nasion - A and Nasion - B.	0° - 4°	> 4°	< 0°	(Steiner and Hills, 1953)
Wits indicator	Distance between the perpendicular lines to the occlusal plane projected from Point A and Point B.	−3 mm to 1 mm (males), −2 mm to 2 mm (females)	>1 mm (males), >2 mm (females)	<−3 mm (males), <−2 mm (females)	(Jacobson, 1975)
APDI	Sum of the facial (Fh/NPg), palatal plane/Frankfurt plane (PP/Fh), and AB plane (AB/NPg) angles.	78°–84°	< 78°	> 84°	(Kim and Vietas, 1975)
FABA	Angle formed by the intersection of the Frankfurt plane and the AB plane.	78°–83°	< 78°	> 83°	(Yang and Suhr, 1995)
Beta	Angle formed by the intersection of the perpendicular line A–CB and the AB plane.	27°- 34°	< 27°	> 34°	(Baik and Ververidou, 2004)

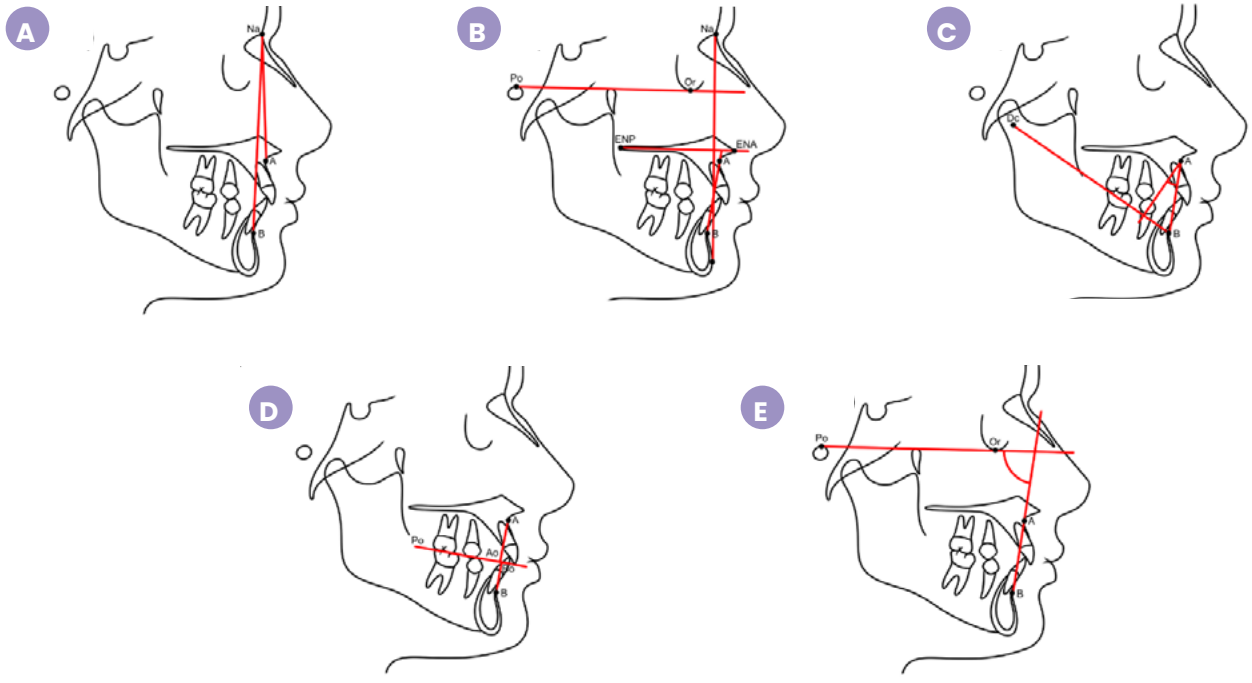


Figure 1. Indicators used in this study to determine skeletal class.
A) ANB **B)** APDI **C)** Beta **D)** Wits **E)** FABA

Table 2. Definition of the facial biotype indicators used in this study and their respective classification values.

Cephalometric Indicator	Definition	Brachyfacial	Mesofacial	Dolichofacial	References
FMA Angle	Intersection between the Frankfurt plane (Po–Or) and the mandibular plane (Go–Me).	≤ 20°	Between 20° and 28°	> 28°	(Tweed, 1946)
Skull base angle (BSCR)	Angle formed by the S–N plane and the Go–Gn plane.	≤ 26°	Between 26° and 38°	> 38°	(Steiner and Hills, 1953)
Sum of the Björk–Jarabak polygon (JBK)	Sum of the angles formed by the points Na–S–Ar–Go–Me.	≤ 390°	Between 390° and 402°	> 402°	(Rodríguez et al., 2014)
VERT	Includes: facial axis (Ba–Na; Pt–Gn), facial depth (Po–Or; Na–Pg), mandibular plane (Po–Or; Go–Gn), lower facial height (ENA–Xi–Pm), and mandibular arch (DC–Xi–Pm).	≥ 0.5	Between −0.5 and 0.5	≤ -0.5	(Ricketts, 1982)

As a measure of agreement between the different indicators, Fleiss' kappa test (1971)[28] and Cohen's weighted kappa test[21] were used in this study. These analyses were performed in the RStudio interface of the RStudio Integrated Development Environment for R[29], using the packages "irr"[30] and "vcd"[31], respectively. The level of agreement was categorized according to the nomenclature established by Landis and Koch:[32]
<0 = Weak; 0–0.20 = Slight; 0.21–0.40 = Fair; 0.41–0.60 = Moderate; 0.61–0.80 = Substantial; 0.81–1.00 = Almost perfect. Furthermore, considering the multivariate nature of our data, based on Cohen's 1972 proposal,[22] a weighted Chi-square test was applied, providing a specific significance value for each unit of analysis (i.e., skeletal classes I, II, and III, as well as the facial biotypes dolichofacial, mesofacial, and brachyfacial, each with their respective *p* values). Fleiss' kappa test complements Cohen's weighted Chi-square test, as it determines the degree of agreement among all evaluators (or indicators) simultaneously, rather than separately for each pair of evaluators (or indicators). When the Fleiss kappa value equals 1, total agreement is present; when it equals 0, agreement corresponds to that expected by chance; and when negative, it indicates complete disagreement among evaluators (or indicators). This approach overcomes the limitations of both the unweighted and weighted Cohen's kappa tests,[19,21] which are restricted to a single pair of evaluators (indicators). In all statistical analyses, an alpha of 5% was used as the cutoff for the *p* value.

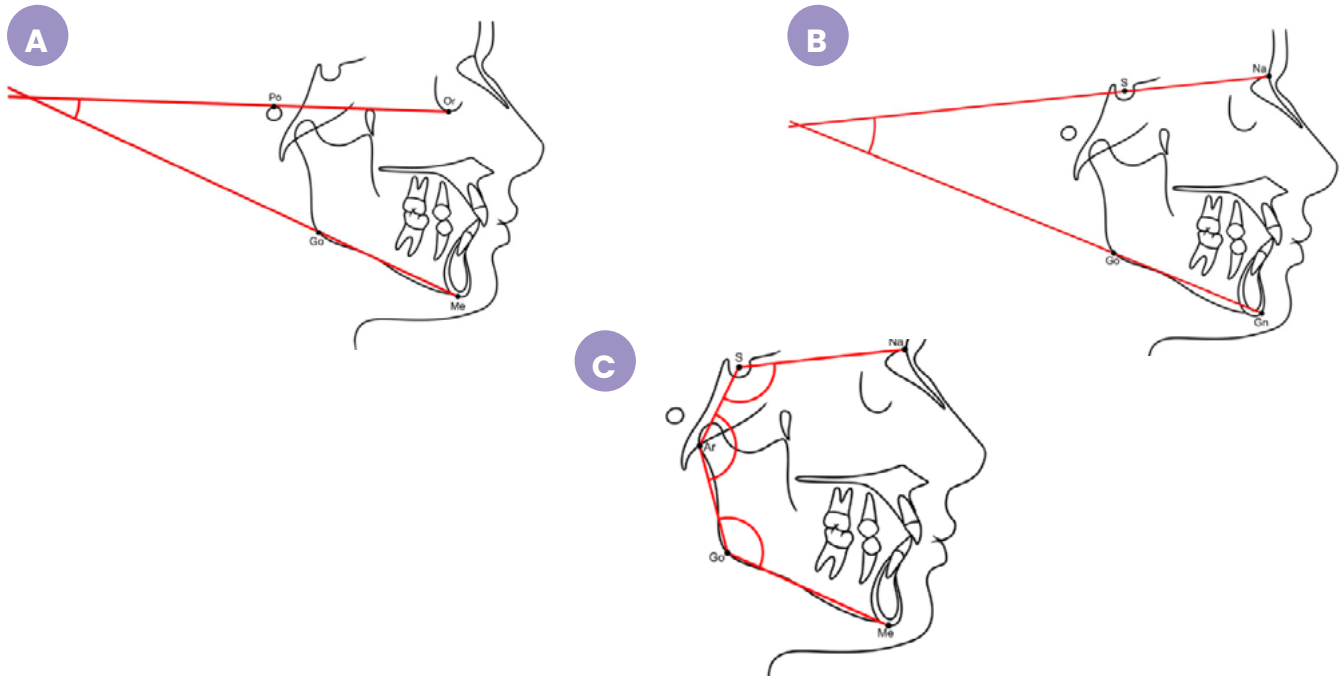
Results

After applying the corresponding cephalometric indicators, significant diagnostic differences were found for both skeletal class and facial biotype.

SKELETAL CLASS

The ANB angle, APDI, and Wits indicator showed a higher frequency of Class II individuals, while the FABA angle revealed a higher frequency of Class III individuals. In contrast, the Beta angle showed a higher number of Class I individuals (**Figure 3**).

Figure 2. Cephalometric indicators used in this study to determine facial biotype.



A) FMA **B)** Skull base angle **C)** Björk–Jarabak polygon

D) VERT, composed of:
(1) Facial axis **(2)** Facial depth **(3)** Mandibular plane **(4)** Lower facial height, and **(5)** Mandibular arch.

Regarding the diagnostic concordance obtained using Cohen's weighted kappa test (**Table 2**, upper diagonal), most concordances were fair or moderate. The highest concordance was between the APDI indicator and the FABA angle (0.55 = Moderate), whereas the lowest was between the ANB angle and the FABA angle (0.19 = Slight). In contrast, the weighted Chi-square test showed highly significant differences in most comparisons between the indicators (**Table 3**, lower diagonal). This result was corroborated by Fleiss' kappa test (Kappa = −0.0714, z = −1.46, *p* (no concordance) = 0.143).

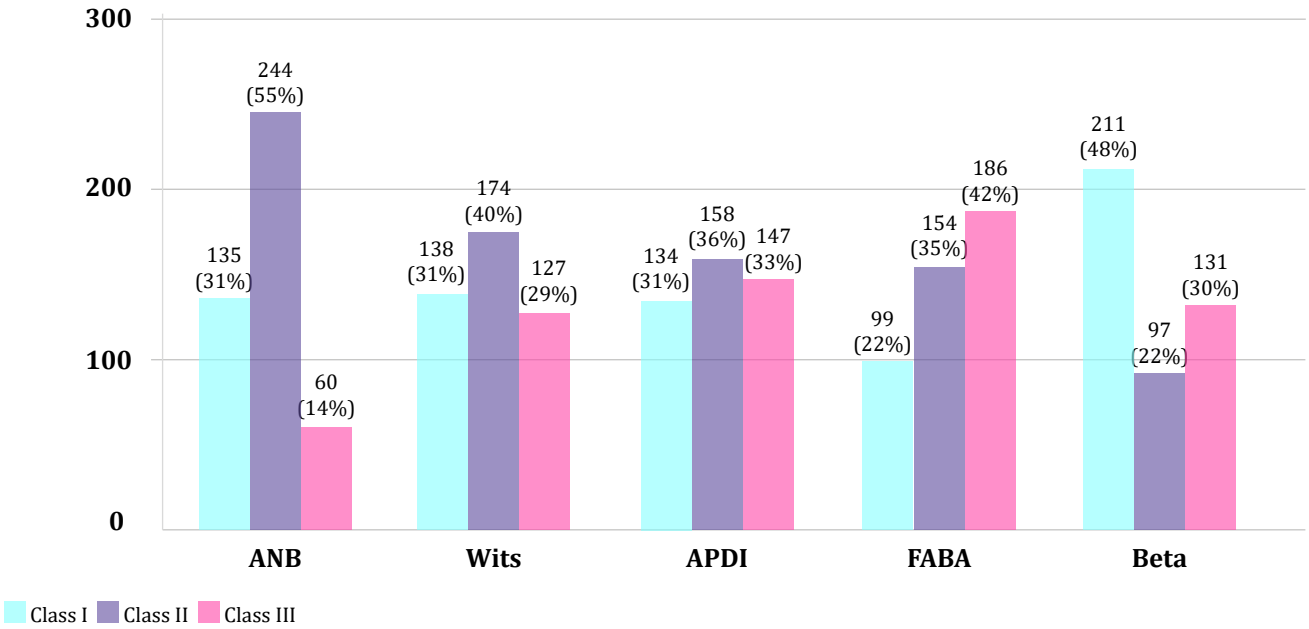


Figure 3. Frequency of skeletal class according to each skeletal class indicator separately (definitions in text).

Table 3. Level of agreement between skeletal class indicators using Cohen's weighted kappa test[21] (upper diagonal) and weighted Chi-square test[22] (lower diagonal) (*p* < 1E−02, in **bold**).

Indicator	Skeletal class	ANB	Wits	APDI	FABA	Beta
ANB	I	*	0.34 Fair	0.29 Fair	0.19 Slight	0.28 Fair
	II					
	III					
Wits	I	8.27E-01	*	0.47 Moderate	0.35 Fair	0.46 Moderate
	II	**2.24E-06**				
	III	**3.34E-08**				
APDI	I	9.41E-01	7.70E-01	*	0.55 Moderate	0.5 Moderate
	II	**5.70E-09**	2.65E-01			
	III	**4.61E-12**	1.45E-01			
FABA	I	**6.00E-03**	**3.03E-03**	**7.47E-03**	*	0.43 Moderate
	II	**1.05E-09**	1.63E-01	7.78E-01		
	III	**2.83E-21**	**3.22E-05**	**6.68E-03**		
Beta	I	**1.53E-07**	**4.80E-07**	**1.03E-07**	**2.60E-15**	*
	II	**2.46E-24**	**1.85E-08**	**5.76E-06**	**2.07E-05**	
	III	**6.33E-09**	7.67E-01	2.46E-01	**1.11E-04**	

The asterisk (*) indicates total agreement (absence of differences) when each indicator is compared with itself.

FACIAL BIOTYPE

The four cephalometric indicators of facial biotype analyzed in this study showed a higher frequency of mesofacial individuals compared with brachyfacial and dolichofacial individuals (**Figure 4**). These differences were statistically significant (weighted Chi-square test with Bonferroni-corrected *p*-values in the observed vs. expected proportions of each facial pattern). For the JBK indicators (all facial patterns:

p_{Brachy} = 3.02E−03, p_{Meso} = 1.28E−16, $p_{Dolicho}$ = 2.63E−09)

and BSCR (all facial patterns:

p_{Brachy} = 6.89E−13, p_{Meso} = 2.06E−18, $p_{Dolicho}$ = 1.59E−02).

In the case of FMA, two of the three facial patterns showed statistically significant differences

(p_{Brachy} = 2.04E−11, p_{Meso} = 7.36E−05, $p_{Dolicho}$ = 5.78E−01),

whereas the VERT indicator did not show such differences in any of the three types of facial patterns

(pBrachy = 3.72E−01, p_{Meso} = 5.38E−01, pDolicho = 7.94E−01).

Regarding the diagnostic concordance obtained using Cohen's weighted kappa test (**Table 4**, upper diagonal), the concordances were fair and moderate. The highest agreement was between the JBK and BSCR indicators (0.59 = Moderate), while the lowest was between the BSCR angle and the VERT indicator (0.36 = Fair). In contrast, the weighted Chi-square test showed highly significant differences in most comparisons between indicators (**Table 4**, lower diagonal). This result was corroborated by Fleiss' kappa test

(Kappa = −0.0909, z = −1.28, p (no concordance) = 0.201).

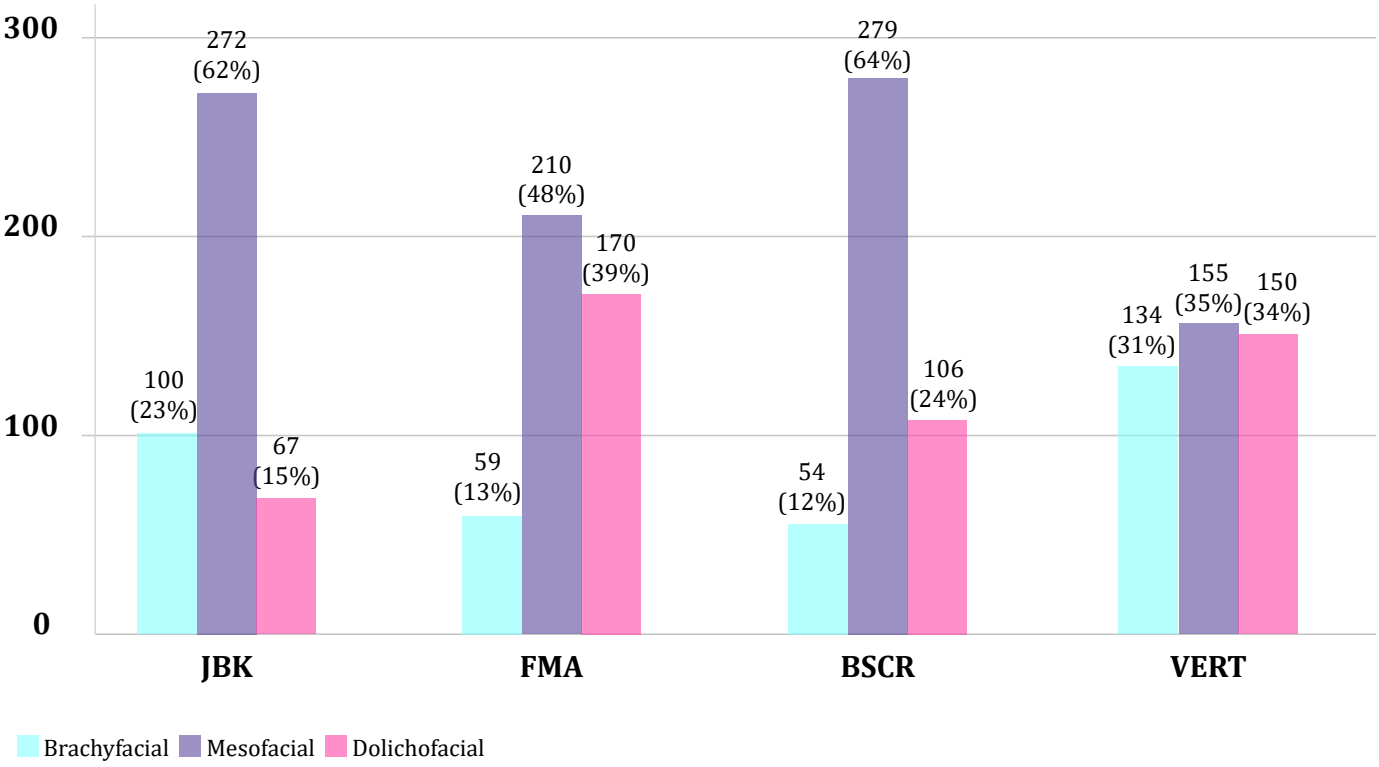


Figure 4. Frequency of facial biotype according to each cephalometric indicator separately (definitions in text).

Table 4. Level of agreement between cephalometric indicators of facial pattern using Cohen's weighted kappa test[21] (upper diagonal) and weighted Chi-square test[22] (lower diagonal) (p < 1E−02, in **bold**).

Indicator	Facial biotype	JBK	FMA	BSCR	VERT
JBK	Brachifacial	*	0.35 Slight	0.59 Moderate	0.45 Moderate
	Mesofacial				
	Dolichofacial				
FMA	Brachifacial	**3.27E-04**	*	0.40 Moderate	0.46 Moderate
	Mesofacial	**2.61E-05**			
	Dolichofacial	**4.87E-15**			
BSCR	Brachifacial	**4.46E-05**	6.14E-01	*	0.36 Slight
	Mesofacial	6.25E-01	**2.76E-06**		
	Dolichofacial	**9.37E-04**	**3.28E-06**		
VERT	Brachifacial	**9.45E-03**	**9.83E-10**	**4.65E-11**	*
	Mesofacial	**2.79E-15**	**1.66E-04**	**5.76E-17**	
	Dolichofacial	**8.37E-11**	1.61E-01	**1.08E-03**	

The asterisk (*) indicates total agreement (absence of differences) when each indicator is compared against itself.

Discussion

This study evaluated the diagnostic agreement between five cephalometric skeletal class indicators and four facial biotype indicators. The sample, consisting of 439 lateral skull teleradiographs, encompassed a wide spectrum of craniofacial morphological variability, allowing a representative analysis of the study population. Our results showed significant variability in the diagnoses obtained for both skeletal class and facial biotype when applying the different cephalometric indicators. This finding is consistent with previous studies that have reported discrepancies in diagnostic classification when using different cephalometric methodologies.[33–35]

To assess the consistency between cephalometric indicators of skeletal class and facial biotype, two statistical strategies were used. First, Cohen's weighted kappa test[21] was applied, which showed concordance levels mainly between the fair and moderate ranges, in contrast to the high discordance observed in the cephalometric diagnoses (**Figure 3** and **Figure 4**). It is noteworthy that the weighted Chi-square test[22] accurately reflected this diagnostic discordance in both analyses. The discrepancy between these two statistical tests underscores the importance of complementing concordance analyses with methodologies of sufficient precision to detect the variability inherent in patients' cephalometric classification systems.

Accordingly, the results of this study suggest that the cephalometric indicators commonly used in clinical practice may be insufficient to accurately characterize craniofacial morphological variability, as there is no single diagnostic standard that allows unequivocal classification of either skeletal class or facial biotype. We believe that this limitation arises because traditional cephalometry, due to the linear nature of its data (points and angles), does not fully capture the complexity of the craniofacial complex, resulting in a significant loss of spatial information [36,37]. In this context, our findings are relevant not only from a methodological standpoint but also from a clinical perspective.

Clinically, the relevance of our findings lies in the fact that, in orthodontic diagnosis, the cephalometric indicators commonly used are insufficient to describe the morphological variation observed in patients. In other words, there is no optimal ("gold standard") indicator capable of achieving accurate classification,[38] which may be attributed to the loss of geometric information caused by reducing craniofacial analysis to a limited set of cephalometric points used in each indicator.[36,37] For this reason, in recent decades, geometric morphometry tools, defined as the statistical analysis of shape and its covariation with other variables, have been introduced into the dental field.[39] This approach has been proposed as a complementary tool for dento-skeletal di-

agnosis,[33,35,40] offering an alternative to the traditional cephalometric indicators used to diagnose skeletal class by capturing the intrinsic pattern of craniofacial morphological variation.

From a clinical standpoint, the correct identification of craniofacial morphology is a fundamental pillar of orthodontics, especially in growing patients, where an accurate diagnosis enables early intervention and optimal guidance of skeletal development. The lack of diagnostic concordance among the different indicators represents a major challenge in therapeutic planning, increasing the risk of misinterpretation that may adversely affect clinical decision-making.

Second, the variability in diagnostic classification found in this study, for both skeletal class and facial biotype, has methodological implications for orthodontic research. Numerous studies compare different characteristics of craniofacial morphology between groups classified according to skeletal class or facial biotype, assuming that the indicators used are reliable and reproducible.[41–43] However, if the criteria used to segment these groups are based on indicators with low concordance, the conclusions drawn from such studies may present important limitations.

Our results highlight that classifying patients with heterogeneous skeletal characteristics within the same study group represents an important problem in orthodontic research. For example, grouping Class II dolichofacial patients with Class II brachyfacial patients could introduce bias in the interpretation of results, since craniofacial morphology is determined by a simultaneous interaction between skeletal class and facial biotype.[5]

In this sense, any analysis that does not consider both variables together could compromise the internal validity of studies and limit the applicability of their findings in clinical practice.

Given this scenario, it is imperative to develop more precise and discriminant diagnostic tools, such as geometric morphometry.[35] This methodology, defined as the statistical analysis of the components of shape and size and their covariation with other variables, allows a more comprehensive evaluation of craniofacial morphology, avoiding the limitations imposed by traditional cephalometric approaches.[33]

In summary, the results of this study reveal the diagnostic discrepancies that exist among the different cephalometric indicators used to classify skeletal class and facial biotype. This lack of concordance represents a problem both in clinical practice and in orthodontic research, underscoring the need to explore new analytical tools to improve accuracy and consistency in craniofacial diagnosis.

Conclusions

1) When assessing diagnostic concordance of the five skeletal class indicators and the four facial biotype indicators, significant differences were found in the diagnoses obtained when applying them in cephalometric analysis, and 2) Regarding the tests that should be used to determine the level of statistical significance of diagnostic concordance, the results suggest the following: (i) for two evaluators without extreme discordance, use the “classical” Cohen’s kappa; (ii) for two evaluators with extreme discordance, use Cohen’s weighted kappa; or (iii) for more than two evaluators, use Cohen’s weighted Chi-square test.

REFERENCES

1. Sassouni V. A classification of skeletal facial types. Am J Orthod. 55(2):109-23.
2. Cerda-Peralta B, Schulz-Rosales R, López-Garrido J, Romo-Ormazabal F. Parámetros cefalométricos para determinar biotipo facial en adultos chilenos. Rev Clínica Periodoncia Implantol Rehabil Oral. 2019;12(1):8-11.
3. Chite-Quispe L, Sánchez-Tito M. Analysis of the association between facial biotype, overbite and overjet in the permanent dentition. J Clin Exp Dent. 1 de mayo de 2023;15(5):e376-81.
4. Costea MC, Bondor CI, Muntean A, Badea ME, Mesaroş AŞ, Kuijpers-Jagtman AM. Proximity of the roots of posterior teeth to the maxillary sinus in different facial biotypes. Am J Orthod Dentofac Orthop Off Publ Am Assoc Orthod Its Const Soc Am Board Orthod. 2018;154(3):346-55.
5. Mizoguchi I, Toriya N, Nakao Y. Growth of the mandible and biological characteristics of the mandibular condylar cartilage. Jpn Dent Sci Rev. 2013;49(4):139-50.
6. Riddle PC, Nickel JC, Liu Y, Gonzalez YM, Gallo LM, Conley RS, et al. Mechanobehavior and mandibular ramus length in different facial phenotypes. Angle Orthod. 2020;90(6):866-72.
7. Saadeh M, Fayyad-Kazan H, Haddad R, Ayoub F. Facial soft tissue thickness differences among different vertical facial patterns. Forensic Sci Int. 2020; 317:110468.
8. Baik CY, Ververidou M. A new approach of assessing sagittal discrepancies: The Beta angle. Am J Orthod Dentofacial Orthop. 2004;126(1):100-5.
9. Jacobson A. The "Wits" appraisal of jaw disharmony. Am J Orthod. 1975;67(2):125-38.
10. Kim Y. WebCeph: Artificial Intelligence Orthodontics and Cloud-Based Orthognatics Platform. https://www.webceph.com
11. Steiner CC, Hills B. Cephalometrics for you and me. Am Journal Orthod. 1953;39(10):729-55.
12. Rodriguez-Cardenas YA, Arriola-Guillen LE, Flores-Mir C. Björk-Jarabak cephalometric analysis on CBCT synthesized cephalograms with different dentofacial sagittal skeletal patterns. Dent Press J Orthod. 2014;19(6):46-53.
13. Tweed CH. The Frankfort-mandibular plane angle in orthodontic diagnosis, classification, treatment planning, and prognosis. Am J Orthod Oral Surg. 1946;32:175-230.
14. Riedel R. The relation of maxillary structures to cranium in malocclusion and in normal occlusion. Angle Orthod. 1952;22:142-5.
15. Paddenberg E, Proff P, Kirschneck C. Floating norms for individualising the ANB angle and the WITS appraisal in orthodontic cephalometric analysis based on guiding variables. J Orofac Orthop. 2021.
16. Nagar S, Nagar R, Raghav P. Why WITS? Why not a way beyond? Contemp Clin Dent. 2014;5(4):518.
17. Yang SD, Suhr CH. F-H to AB plane angle (FABA) for assessment of anteroposterior jaw relationships. Angle Orthod. 1995;65(3):223-31.
18. Gregoret J. Instrumentación de la cefalometría estática lateral. En: Ortodoncia y cirugía ortognática : diagnóstico y planificación. Barcelona: Espaxs; 1998. p. 227-32. https://books.google.com/books/about/Ortodoncia_y_cirugia_ortognática.html

19. Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas. 1960;20(1):37-46.

20. Vach W, Gerke O. Gwet's AC1 is not a substitute for Cohen's kappa – A comparison of basic properties. MethodsX. 2023;10:102212.

21. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull. 1968;70(4):213-20.

22. Cohen J. Weighted Chi Square: an Extension of the Kappa Method. Educ Psychol Meas. 1972;32(1):61-74.

23. Ahmed M, Shaikh A, Fida M. Diagnostic performance of various cephalometric parameters for the assessment of vertical growth pattern. Dent Press J Orthod. 2016;21(4):41-9.

24. Ahmed M, Shaikh A, Fida M. Diagnostic validity of different cephalometric analyses for assessment of the sagittal skeletal pattern. Dent Press J Orthod. 2018;23(5):75-81.

25. Benedicto EDN, Kairalla SA, Oliveira GMS, Junior LRM, Rosário HD, Paranhos LR. Determination of vertical characteristics with different cephalometric measurements. Eur J Dent. 2016;10(01):116-20.

26. Zaffiri VZ, Idrovo JDI, Manríquez G, Diaz A. Determinación de la clase esqueletal mediante el indicador de Wits: ¿es correcto diferenciar la norma entre sexos? Odontoestomatología. 2024;26(44). Disponible en: https://odon.edu.uy/ojs/index.php/ode/article/view/702

27. Ricketts R. Orthodontic Diagnosis and Planning: Their Roles in Preventive and Rehabilitative Dentistry. En: Rocky Mountain/Orthodontics; 1982.

28. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76(5):378-82.

29. RStudio Team. RStudio: Integrated Development Enviroment for R [Internet].

30. Matthias Gamer, Jim Lemon, Ian Fellows, Puspendra Singh. Irr: Various Coefficients of Interrater Reliability and Agreement [Internet]. 2005. p. 0.84.1. Disponible en: https://CRAN.R-project.org/package=irr

31. Meyer D, Zeileis A, Hornik K, Friendly M. VCD: Visualizing Categorical Data 2002. p. 1.4-13.

32. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159.

33. Halazonetis DJ. Morphometrics for cephalometric diagnosis. Am J Orthod Dentofacial Orthop. 2004;125(5):571-81.

34. Chang HP, Liu PH, Chang CH. Midfacial and mandibular morphometry of children with Class II and Class III malocclusions. Journal of Oral Rehabilitation; 2005 p. 642-7.

35. Díaz A, Manríquez G. Skeletodental Diagnosis Using a Geometric Morphometric Approach. Int J Odontostomatol. 2014;8(1):05-11.

36. Moyers RE, Bookstein FL. The inappropriateness of conventional cephalometrics. Am J Orthod. 1979;75(6):599-617.

37. McIntyre GT, Mossey PA. Size and shape measurement in contemporary cephalometrics. Eur J Orthod. 2003;25(3):231-42.

38. Castro Arenas MV, Hurtado Peralta M, Oyonarte Weldt R. Performance of cephalometric assessment for sagital intermaxillary diagnosis. A narrative review. Rev Clínica Periodoncia Implantol Rehabil Oral. 2013;6(2):96-101.

39. James Rohlf F, Marcus LF. A revolution in morphometrics. Trends Ecol Evol. 1993;8(4):129-32.

40. Freudenthaler J, Čelar A, Ritt C, Mitteröcker P. Geometric morphometrics of different malocclusions in lateral skull radiographs. J Orofac Orthop. 2017;78(1):11-20.

41. Wilhelm BM, Beck FM, Lidral AC, Vig KW. A comparison of cranial base growth in Class I and Class II skeletal patterns. Am J Orthod Dentofac Orthop Off Publ Am Assoc Orthod Its Const Soc Am Board Orthod. 2001;119(4):401-5.

42. Jacob HB, Buschang PH. Mandibular growth comparisons of Class I and Class II division 1 skeletofacial patterns. Angle Orthod. 2014;84(5):755-61.

43. Polat OO, Kaya B. Changes in cranial base morphology in different malocclusions. Orthod Craniofac Res. 2007;10(4):216-21.

Data Availability

The dataset supporting the results of this study is available at the Center for Quantitative Analysis in Dental Anthropology, Faculty of Dentistry, University of Chile.

Conflict of Interest Statement

The authors declare no conflict of interest.

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or nonprofit sectors.

Authors' Contributions

AUTHORS' FULL NAME	ACADEMIC COLLABORATION													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Juan Diego Idrovo			x	x	x	x	x	x		x				x
Valentina Zaffiri				x		x	x			x				x
Germán Manríquez	x		x	x	x	x		x	x	x		x	x	x
Alejandro Díaz	x		x	x		x		x	x	x		x	x	x

1. Project Administration
2. Funding Acquisition
3. Formal Analysis
4. Conceptualization
5. Data Curation
6. Writing - Review and Editing
7. Research
8. Methodology
9. Resources
10. Writing - Original Draft Preparation
11. Software
12. Supervision
13. Validation
14. Visualization

Acceptance note

This article was approved by the journal editor, Dr. Natalia Tancredi Cueto, MSc.