

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

INVESTIGACIÓN

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

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

Resumen

Objetivos: 1) determinar la concordancia diagnóstica de cinco indicadores cefalométricos de clase esquelética 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 esquelética (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 esquelética como del biotipo facial. Mientras que en la prueba ponderada de kappa de Cohen, la mayoría de las concordancias entre los indicadores de clase esquelética 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.

$$\begin{aligned} \text{kappa de Fleiss}_{\text{Clase esquelética}} &= -0.0714, z = -1.46, p\text{-value} = 0.143 \\ \text{kappa de Fleiss}_{\text{Biotipo facial}} &= -0.0909, z = -1.28, p\text{-value} = 0.201 \end{aligned}$$

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 esquelética 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.

AUTORÍA

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

CORRESPONDENCIA

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.

Recibido: 25/04/2025

Aceptado: 03/10/2025



Palabras clave

Cefalometría,
Imágenes diagnós-
ticas, Maloclusión,
Ortodoncia, Análisis
estadístico.

Abstract

Objectives: 1) to determine the diagnostic agreement of five skeletal class cephalometric indicators and four facial biotype cephalometric indicators through Cohen's weighted kappa test, and 2) to compare these results against those obtained after applying Cohen's weighted Chi-square test.

Material and methods: In a sample of 439 lateral telerradiographs, skeletal class indicators (ANB angle, APDI, FABA angle, Wits indicator and beta angle) and facial biotype indicators (FMA, skull base angle, Bjork Jarabak polygon and VERT index) were determined. Diagnostic concordance among them was assessed with Fleiss' kappa, Cohen's weighted kappa and Cohen's weighted Chi-square tests.

Results: Statistically significant differences were observed both in the diagnosis of skeletal class and facial biotype. While in Cohen's weighted kappa test, most of the concordances between skeletal class and facial biotype indicators were "Mild" or "Moderate", Cohen's weighted Chi-square test showed highly significant discordance in most comparisons, with negative values of the Fleiss' kappa test, indicating the systematic absence of concordance when considering all indicators simultaneously.

Fleiss' kappa_{skeletal class} = -0.0714, z = -1.46, p-value = 0.143

Fleiss' kappa_{facial biotype} = -0.0909, z = -1.28, p-value = 0.201

Conclusion: 1) When studying the diagnostic agreement of the five skeletal class indicators and the four facial biotype indicators, significant differences were found between them, a result consistently supported by the use of Cohen's weighted Chi2 statistical test, and 2) In relation to the tests to be used to know the level of statistical significance of the diagnostic concordance, it is recommended to take into account: (i) if the evaluators are only two, and they don't show an extreme discordance in their respective evaluations, use the "classical" Cohen's kappa, (ii) if the evaluators are only two, and they do show an extreme discordance in their respective evaluations, then use Cohen's weighted kappa, and, (iii) for more than two evaluators apply the Cohen's weighted Chi-square test.

Keywords

Cephalometry, Diagnostic imaging, Malocclusion, Orthodontics, Statistical analysis.

Resumo

Objetivos: 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, Análises estatística.

Introducción y antecedentes

La clase esquelética se define como la relación sagital entre el maxilar y la mandíbula,⁽¹⁾ se divide en tres categorías: Clase I, Clase II y Clase III. Además, el biotipo facial se define como el conjunto de características morfodiferenciales que determinan el componente vertical de la morfología craneofacial y que permite comprender los comportamientos funcionales y las necesidades terapéuticas específicas de cada biotipo.⁽²⁻⁴⁾ En la literatura se describen 3 tipos: braquifacial (o hipodivergente), mesofacial (o normodivergente) y dolicofacial (o hiperdivergente).⁽⁵⁻⁷⁾

En ambos casos, para su determinación se utiliza rutinariamente la cefalometría, una herramienta basada en la medición de distancias lineales y angulares en telerradiografías craneales. Aunque se ha propuesto una amplia variedad de indicadores cefalométricos para clasificar la clase esquelética y el biotipo facial, no existe un indicador universalmente aceptado o "gold standard".⁽⁸⁻¹⁴⁾ Por razones clínicas, se espera que estos indicadores tengan una alta concordancia diagnóstica entre ellos, de modo que los resultados obtenidos sean consistentes, independientemente del indicador utilizado. Otro aspecto a considerar es que a lo largo del tiempo se han propuesto nuevos indicadores para superar las limitaciones de los anteriores. Por ejemplo, en relación con la clase esquelética, Jacobson sugiere que el ángulo ANB puede ser malinterpretado dependiendo de la posición del punto cefalométrico Nasion,^(9,15) recomendando el uso de los puntos A y B de Downs en relación con el plano oclusal para la evaluación sagital de la relación maxilomandibular.⁽⁹⁾ Sin embargo, también se ha reportado una alta variabilidad y en ocasiones una difícil determinación de este último método debido al plano oclusal.⁽¹⁶⁾ Mientras tanto, los creadores del ángulo FABA⁽¹⁷⁾ encontraron en oclusiones normales una alta variabilidad en el indicador APDI y el indicador de Wits, cuestionando su validez y sugiriendo el ángulo FABA por encima de otros indicadores. Finalmente, según los autores del ángulo Beta, la ventaja de esta propuesta sería que no depende de referencias craneales, de la estabilidad de la oclusión o de la rotación mandibular.⁽⁸⁾ En cuanto al biotipo facial, se han propuesto los siguientes indicadores: la sumatoria del polígono de Björk-Jarabak (JBK),⁽¹²⁾ el ángulo del plano de Frankfurt a plano mandibular (FMA),⁽¹³⁾ ángulo de la base de cráneo (BSCR)⁽¹¹⁾ y plano mandibular (SN-GoGn)⁽¹⁴⁾ y el VERT de Ricketts.⁽¹⁸⁾

Debido a esta gran cantidad de propuestas, cada indicador se convierte en una solución parcial e independiente, alejándose de una concordancia clínica ideal.

En cuanto a la evaluación de la concordancia diagnóstica entre diferentes evaluadores (indicadores), la prueba de kappa de Cohen⁽¹⁹⁾ ha sido tradicionalmente utilizada desde su creación hasta la fecha.⁽²⁰⁾ A pesar de su popularidad, esta prueba no tiene en cuenta las discordancias extremas, y su uso se recomienda solo cuando no se observan diferencias significativas entre evaluadores (indicadores). Teniendo en cuenta esto, Cohen propuso posteriormente una prueba estadística que consideró el peso de las discordancias extremas,⁽²¹⁾ así como una prueba de Ji^2 ponderada⁽²²⁾ en el caso de haber más de dos evaluadores (indicadores). Considerando que el nivel de concordancia entre los indicadores cefalométricos mencionados anteriormente es bajo,⁽²³⁻²⁵⁾ el presente estudio tiene como objetivos: 1) determinar la concordancia diagnóstica entre ellos utilizando la prueba de kappa ponderada,⁽²¹⁾ y 2) comparar estos resultados con los obtenidos después de aplicar la prueba de Ji^2 ponderada,⁽²²⁾ contribuyendo de esta manera a mejorar la práctica clínica en ortodoncia.

Materiales y Métodos

Se obtuvo una muestra compuesta por 439 telerradiografías laterales (275 mujeres y 164 hombres) de una base de datos anonimizada de pacientes mayores de 18 años, sin antecedentes de tratamiento ortodóncico, ortopédico ni quirúrgico maxilofacial, que presentaban una resolución adecuada para la identificación de los puntos cefalométricos. Para cada radiografía, se determinaron los siguientes indicadores de clase esquelética: ángulo ANB, APDI, ángulo FABA, ángulo Beta e indicador de Wits, este último agrupando la variable sexo debido a la ausencia de diferencias estadísticamente significativas entre mujeres y hombres (prueba de Ji^2 : $Ji^2 = 0.313$, p (sin asociación) = 0.855, $df = 2$) (**Tabla 1, Figura 1**), hecho que ha sido recientemente corroborado.⁽²⁶⁾ En cuanto al biotipo facial, se utilizaron los siguientes indicadores: la sumatoria del polígono de Björk-Jarabak (JBK),⁽¹²⁾ el ángulo del plano de Frankfurt y plano mandibular (FMA),⁽¹³⁾ el ángulo de la base de cráneo y plano mandibular (BSCR),⁽¹⁴⁾ y el VERT de Ricketts^(18,27) (**Tabla 2, Figura 2**). Todos estos indicadores se determinaron utilizando la plataforma WebCeph V1.5.0.⁽¹⁾

Tabla 1. Indicadores de clase esquelética usados en este estudio, su definición y sus respectivos valores clasificatorios (A= Punto A, B= Punto B, C= Centro del cóndilo).

Indicador	Definición	Clase I	Clase II	Clase III	Referencias
ANB	Ángulo formado por la intersección de las líneas Nasion - A y Nasion - B.	0°- 4°	> 4°	< 0°	(Steiner and Hills, 1953)
Indicador de Wits	Distancia entre líneas perpendiculares al plano oclusal proyectadas desde el punto A y el punto B.	De -3 mm a 1 mm (hombres), De -2 mm a 2 mm (mujeres)	> 1 mm. (hombres), > 2 mm. (mujeres)	< -3 mm (hombres), < -2 mm (mujeres)	(Jacobson, 1975)
APDI	Sumatoria de los ángulos facial (Fh/NPg), plano palatino /plano de Frankfurt (PP/Fh) y plano AB (AB/NPg).	De 78° a 84°	< 78°	> 84°	(Kim and Vietas, 1975)
FABA	Ángulo formado por la intersección del plano de Frankfurt y el plano AB.	De 78° a 83°	< 78°	> 83°	(Yang and Suhr, 1995)
Beta	Ángulo formado por la intersección de la línea perpendicular A - CB y el plano AB.	De 27° a 34°	< 27°	> 34°	(Baik and Ververidou, 2004)

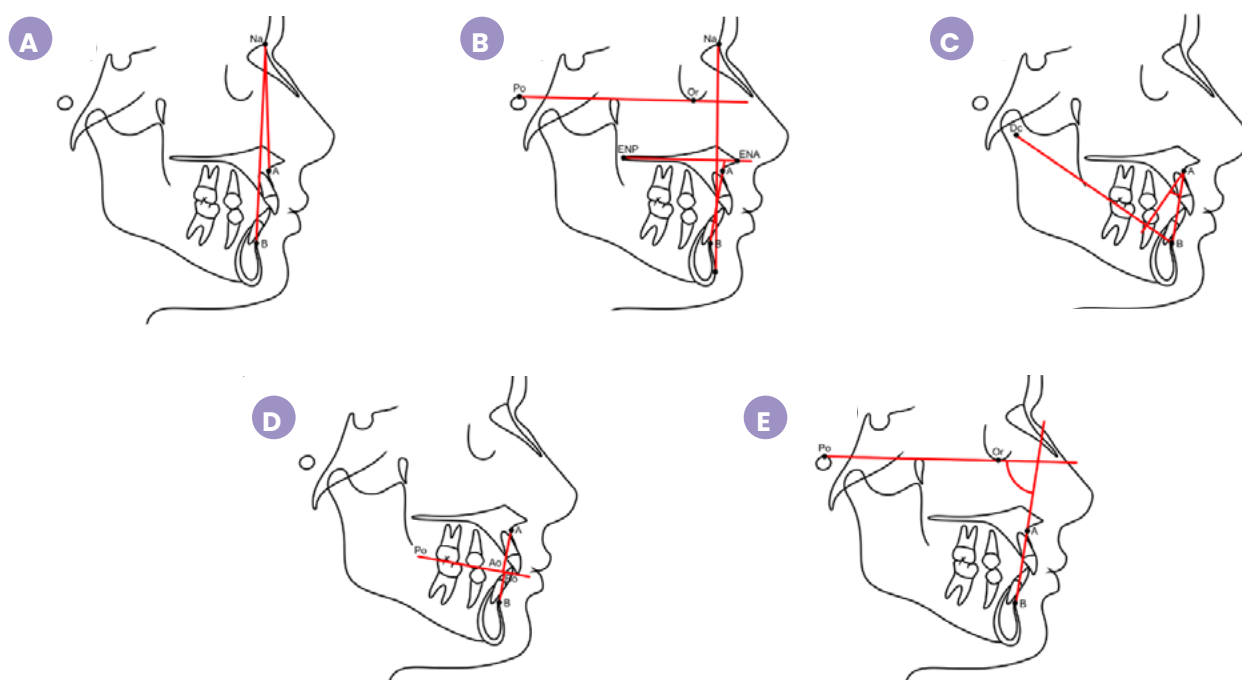


Figura 1. Indicadores usados en el presente estudio para determinar la clase esquelética.

A) ANB B) APDI C) Beta D) Wits E) FABA

Tabla 2. Definición de los indicadores de biotipo facial utilizados en el presente estudio y sus respectivos valores clasificatorios.

Indicador Cefalométrico	Definición	Braquifacial	Mesofacial	Dolicofacial	Referencias
Ángulo FMA	Intersección entre el plano de Frankfurt (Po-Or) y el plano mandibular (Go-Me).	$\leq 20^\circ$	Entre 20° y 28°	$> 28^\circ$	(Tweed, 1946)
Ángulo Base de cráneo (BSCR)	Ángulo formado por el plano S-N y el plano Go-Gn	$\leq 26^\circ$	Entre 26° y 38°	$> 38^\circ$	(Steiner and Hills, 1953)
Sumatoria del polígono de Björk-Jarabak (JRBK)	Sumatoria de los ángulos conformados por los puntos: Na-S-Ar-Go-Me.	$\leq 390^\circ$	Entre 390° y 402°	$> 402^\circ$	(Rodríguez y cols, 2014)
VERT	Incluye: eje facial (Ba-Na; Pt-Gn), profundidad facial (Po-Or; Na-Pg), plano mandibular (Po-Or; Go-Gn), altura facial inferior (ENA-Xi-Pm) y arco mandibular (DC-Xi-Pm)	$\geq 0,5$	Entre -0,5 y 0,5	$\leq -0,5$	(Ricketts, 1982)

Como medida de concordancia entre los diferentes indicadores, en este estudio se utilizaron la prueba de kappa de Fleiss (1971)⁽²⁸⁾ y la prueba de kappa ponderada de Cohen.⁽²¹⁾ Estos análisis se realizaron en la interfase RStudio del ambiente computacional integrado R,⁽²⁹⁾ utilizando los paquetes “irr”⁽³⁰⁾ y “vcd”,⁽³¹⁾ respectivamente. El nivel de concordancia se categorizó según la nomenclatura establecida por Landis y Koch:⁽³²⁾ < 0 = Débil; $0-0.2$ = Ligera; $0.21-0.40$ = Leve; $0.41-0.60$ = Moderada; $0.61-0.80$ = Sustancial; $0.81-1.00$ = Casi perfecta. Además, considerando la naturaleza multivariante de nuestros datos, a partir de la propuesta de Cohen del año 1972,⁽²²⁾ se empleó una prueba de Ji^2 ponderada, proporcionando un valor de significancia específico para cada unidad de análisis (es decir, las clases esqueléticas I, II y III, así como los biotipos faciales dolicofacial, mesofacial y braquifacial, en ambos casos con sus respectivos valores de p). La prueba kappa de Fleiss es un complemento de la prueba de Ji^2 ponderada de Cohen que entrega el grado de concordancia o acuerdo entre todos los evaluadores (o indicadores) a la vez, en lugar de hacerlo por separado para cada uno de los respectivos pares de evaluadores (o indicadores). Cuando el valor del kappa de Fleiss es igual a 1, significa que hay una total concordancia, cuando es igual a cero se trata de

una concordancia equivalente a lo que se esperaría por simple azar y cuando el valor del kappa de Fleiss es negativo, corresponde a una total discordancia entre los evaluadores (o indicadores). Este enfoque supera las limitaciones tanto de las pruebas de kappa de Cohen no ponderada como de la ponderada,^(19,21) que están restringidas a un solo par de evaluadores (indicadores). En todos los análisis estadísticos como punto de corte para el valor de p se utilizó un alfa del 5%.

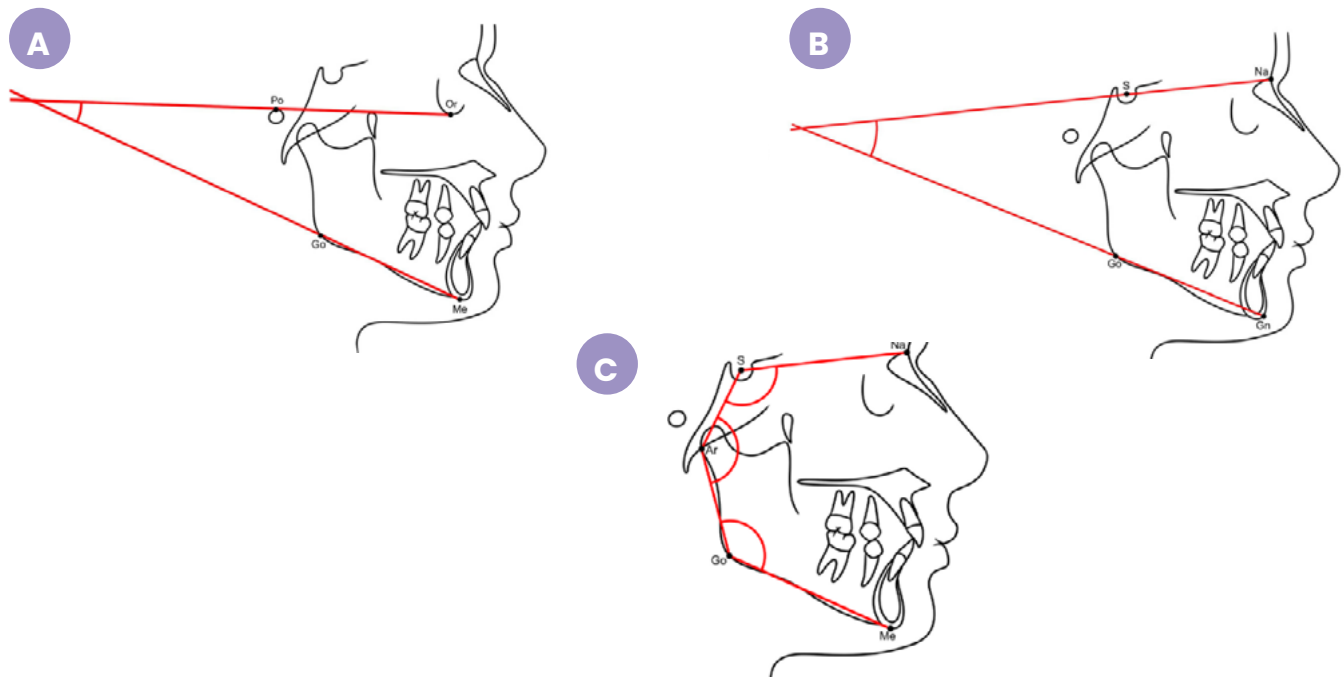
Resultados

Después de aplicar los indicadores cefalométricos correspondientes, se encontraron diferencias significativas en el diagnóstico tanto para la clase esquelética como para el biotipo facial.

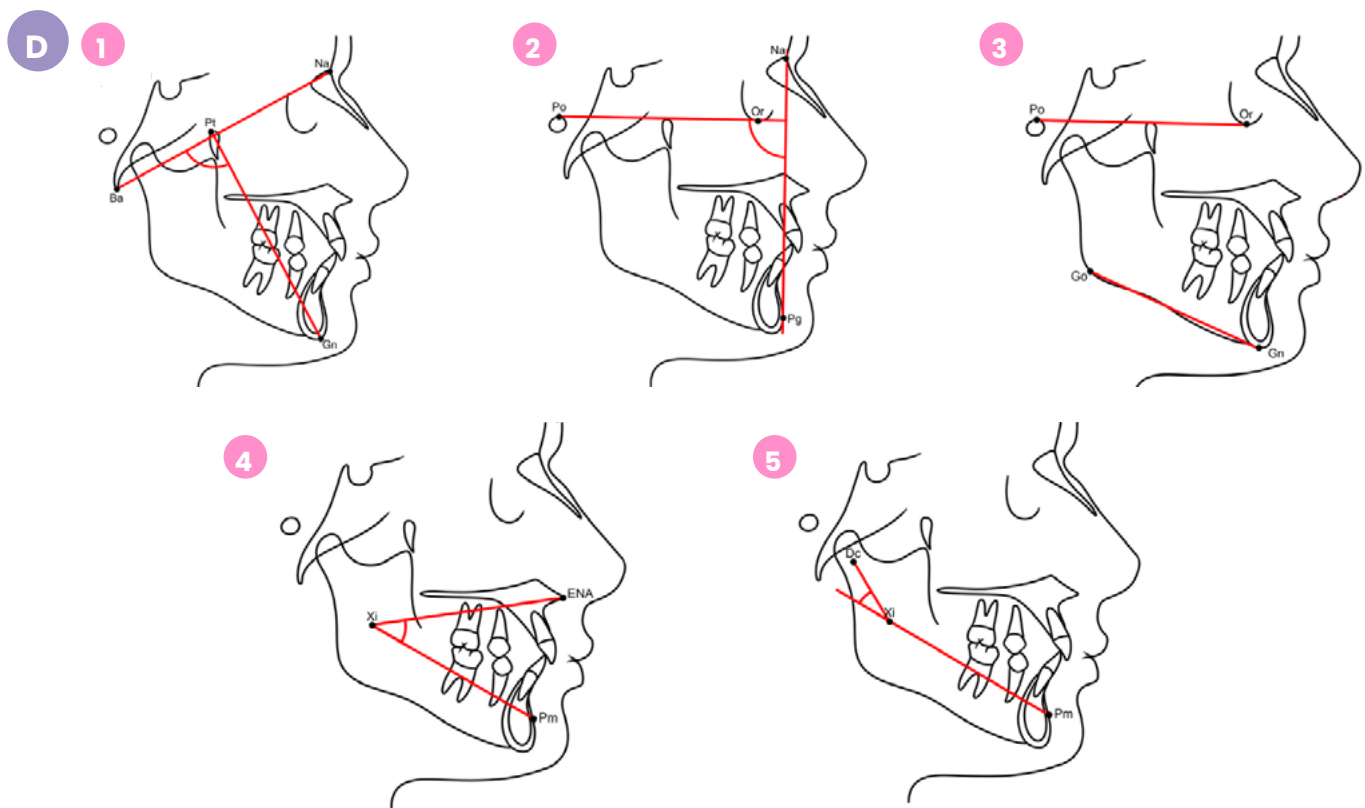
CLASE ESQUELETAL

El ángulo ANB, APDI y el indicador de Wits mostraron una mayor frecuencia de individuos de Clase II, mientras que el ángulo FABA reveló una mayor frecuencia de individuos de Clase III. Por el contrario, el ángulo Beta resultó en un mayor número de individuos de Clase I (**Figura 3**).

Figura 2. Indicadores cefalométricos utilizados en este estudio para la determinación del biotipo facial.



A) FMA **B)** Ángulo base de cráneo **C)** Polígono de Björk-Jarabak



D) VERT, compuesto por:

(1) Eje facial **(2)** Profundidad facial **(3)** Plano Mandibular **(4)** Altura facial inferior **(5)** Arco mandibular.

En cuanto a la concordancia diagnóstica obtenida mediante la prueba de kappa de Cohen ponderada (Tabla 3, diagonal superior), la mayoría de las concordancias fueron “Leves” o “Moderadas”. La mayor concordancia fue entre el indicador APDI y el ángulo FAB A (0.55 = Moderada), mientras que la más baja fue entre el ángulo ANB y el ángulo FAB A (0.19 = Ligera). En contraste,

la prueba de Ji^2 mostró diferencias altamente significativas en la mayoría de las comparaciones entre los indicadores (Tabla 3, diagonal inferior). Este último resultado fue corroborado por la prueba kappa de Fleiss ($kappa = -0.0714$, $z = -1.46$, p (ausencia de concordancia) = 0.143).

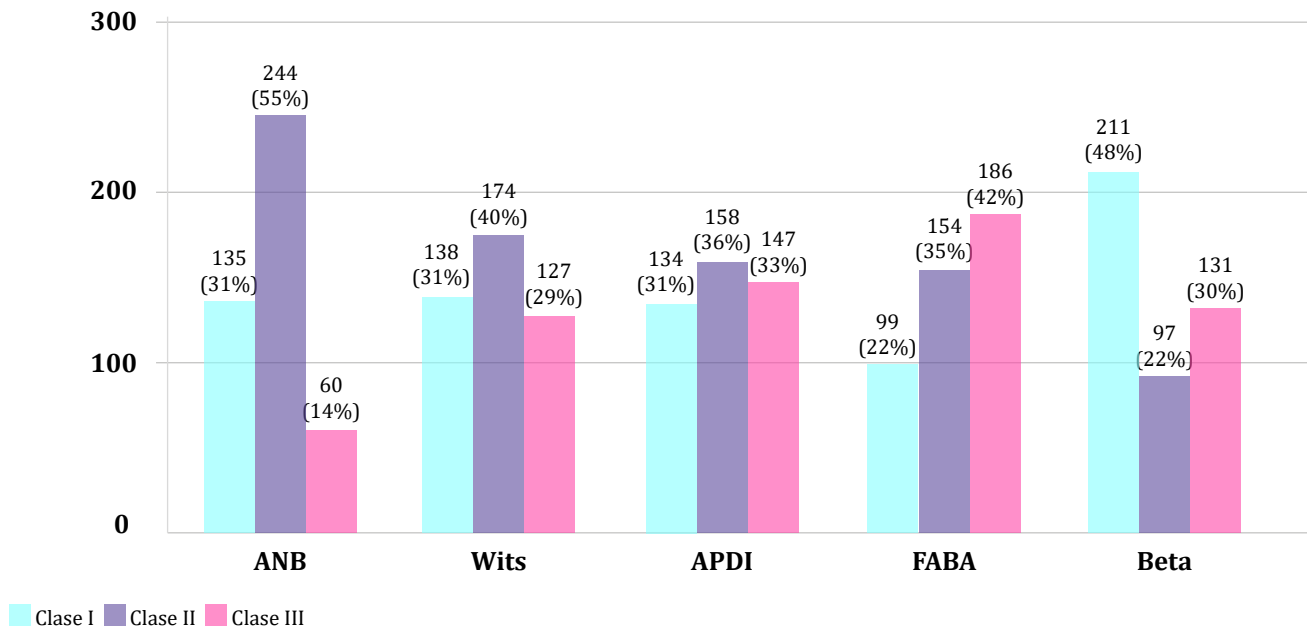


Figura 3. Frecuencia de clase esquelética de acuerdo con cada indicador de clase esquelética por separado (definiciones en el texto).

Tabla 3. Nivel de concordancia entre los indicadores de clase esquelética usando la prueba kappa de Cohen ponderada⁽²¹⁾ (diagonal superior) y la prueba Ji^2 ponderada⁽²²⁾ (diagonal inferior) ($p < 1E-02$, en **negrita**).

Indicador	Clase Esquelética	ANB	Wits	APDI	FAB A	Beta
ANB	I		0.34	0.29	0.19	0.28
	II	*	Leve	Leve	Ligera	Leve
	III					
Wits	I	8.27E-01		0.47	0.35	0.46
	II	2.24E-06	*	Moderada	Leve	Moderada
	III	3.34E-08				
APDI	I	9.41E-01	7.70E-01		0.55	0.5
	II	5.70E-09	2.65E-01	*	Moderada	Moderada
	III	4.61E-12	1.45E-01			
FAB A	I	6.00E-03	3.03E-03	7.47E-03		0.43
	II	1.05E-09	1.63E-01	7.78E-01	*	Moderada
	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	

El asterisco (*) indica total concordancia (ausencia de diferencias), cuando cada indicador se compara consigo mismo.

BIOTIPO FACIAL

Los cuatro indicadores cefalométricos de biotipo facial analizados en el presente estudio mostraron una mayor frecuencia de individuos mesofaciales en comparación con los individuos braquifaciales y dolicofaciales (**Figura 4**). Estas diferencias fueron estadísticamente significativas (prueba de Ji^2 con valores p corregidos por Bonferroni en las proporciones observadas frente a las esperadas de cada patrón facial). Para el indicador JBK (todos los patrones faciales:

$$p_{\text{Braqui}} = 3.02\text{E-}03, p_{\text{Meso}} = 1.28\text{E-}16, p_{\text{Dolico}} = 2.63\text{E-}09),$$

y BSCR (todos los patrones faciales:

$$p_{\text{Braqui}} = 6.89\text{E-}13, p_{\text{Meso}} = 2.06\text{E-}18, p_{\text{Dolico}} = 1.59\text{E-}02).$$

En el caso de FMA, dos de los tres patrones faciales mostraron diferencias estadísticamente significativas

$$(p_{\text{Braqui}} = 2.04\text{E-}11, p_{\text{Meso}} = 7.36\text{E-}05, p_{\text{Dolico}} = 5.78\text{E-}01),$$

mientras que el indicador VERT no mostró tales diferencias en ninguno de los tres tipos de patrones faciales

$$(p_{\text{Braqui}} = 3.72\text{E-}01, p_{\text{Meso}} = 5.38\text{E-}01, p_{\text{Dolico}} = 7.94\text{E-}01).$$

En cuanto a la concordancia diagnóstica obtenida mediante la prueba kappa de Cohen ponderada (**Tabla 4**, diagonal superior), las concordancias fueron “Leves” y “Moderadas”. La mayor concordancia fue entre el indicador JBK y BSCR (0.59 = Moderada), mientras que la más baja fue entre el ángulo BSCR y el VERT (0.36 = Leve). En contraste, la prueba de Ji^2 ponderada mostró diferencias altamente significativas en la mayoría de las comparaciones entre los indicadores (**Tabla 4**, diagonal inferior). Este último resultado fue corroborado por la prueba kappa de Fleiss (kappa= -0.0909, z= -1.28, p (ausencia de concordancia)= 0.201).

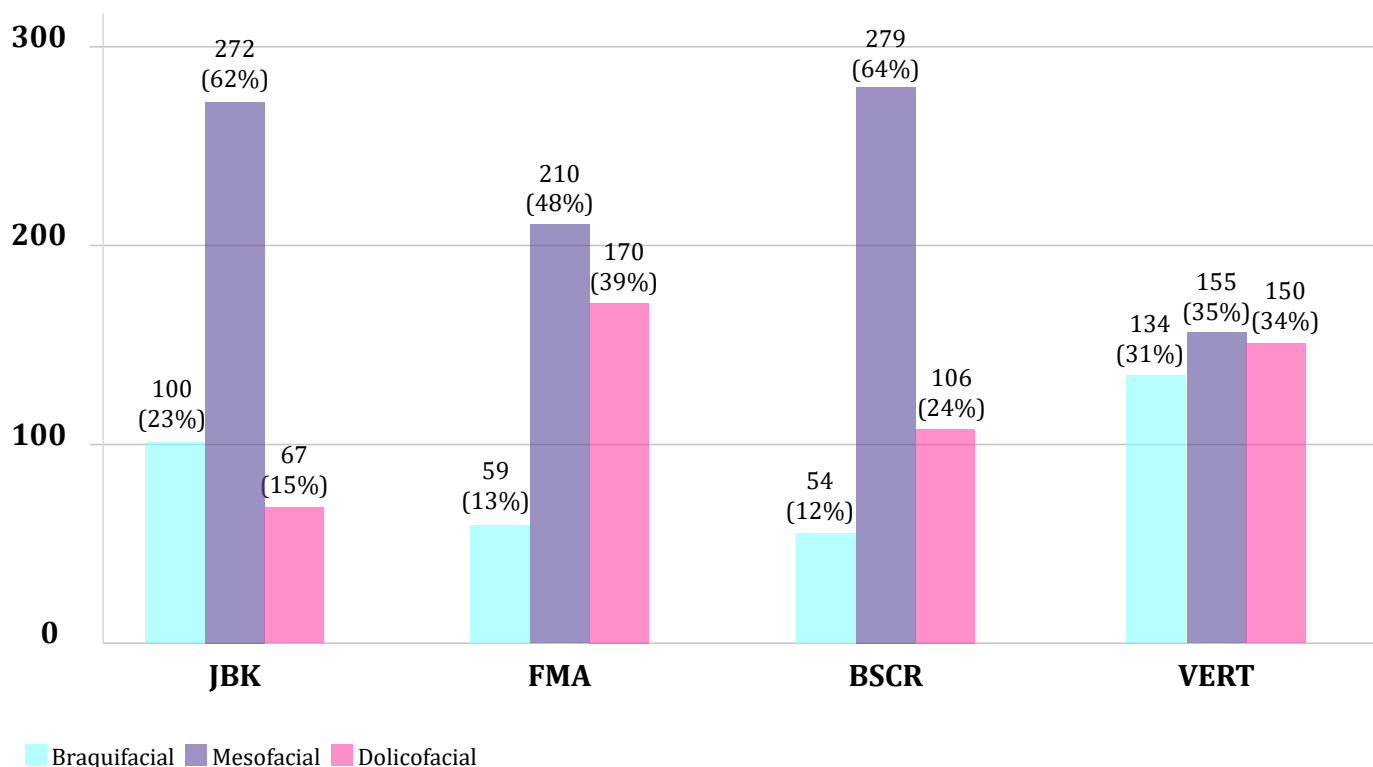


Figura 4. Frecuencia de biotipo facial de acuerdo con cada indicador de biotipo facial por separado (definiciones en el texto).

Tabla 4. Nivel de concordancia entre los indicadores cefalométricos del patrón facial utilizando la prueba de kappa ponderada de Cohen⁽²¹⁾ (diagonal superior) y la prueba de Ji^2 ponderada⁽²²⁾ (diagonal inferior) ($p < 1E-02$, en **negrita**).

Indicador	Biotipo Facial	JBK	FMA	BSCR	VERT
JBK	Braquifacial				
	Mesofacial	*	0.35 Leve	0.59 Moderada	0.45 Moderada
	Dolicofacial				
FMA	Braquifacial	3.27E-04			
	Mesofacial	2.61E-05	*	0.40 Moderada	0.46 Moderada
	Dolicofacial	4.87E-15			
BSCR	Braquifacial	4.46E-05	6.14E-01		
	Mesofacial	6.25E-01	2.76E-06	*	0.36 Leve
	Dolicofacial	9.37E-04	3.28E-06		
VERT	Braquifacial	9.45E-03	9.83E-10	4.65E-11	
	Mesofacial	2.79E-15	1.66E-04	5.76E-17	*
	Dolicofacial	8.37E-11	1.61E-01	1.08E-03	

El asterisco (*) indica total concordancia (ausencia de diferencias), cuando cada indicador se compara consigo mismo.

Discusión

El presente estudio evaluó la concordancia diagnóstica entre cinco indicadores cefalométricos de clase esquelética y cuatro de biotipo facial. La muestra, consistente en 439 telerradiografías laterales de cráneo, abarcó un amplio espectro de variabilidad morfológica craneofacial, permitiendo un análisis representativo de la población en estudio. Nuestros resultados evidenciaron una variabilidad significativa en los diagnósticos obtenidos para la clase esquelética y el biotipo al aplicar los distintos indicadores cefalométricos. Este hallazgo coincide con estudios previos que han documentado discrepancias en la clasificación diagnóstica al utilizar diferentes metodologías cefalométricas.⁽³³⁻³⁵⁾

Para evaluar la consistencia entre los indicadores cefalométricos, tanto de clase esquelética como de biotipo facial, el presente trabajo utilizó dos estrategias. Por una parte, se empleó la prueba kappa ponderada de Cohen,⁽²¹⁾ que mostró niveles de concordancia mayoritariamente entre los rangos «Leve» y «Moderada», en contraste con la alta discordancia observada en los diagnósticos cefalométricos (**Figura 3** y **Figura 4**). Cabe destacar que la prueba de Ji^2 ponderada⁽²²⁾ reflejó fielmente dicha discordancia diagnóstica para ambos tipos de análisis. Esta discrepancia entre ambas pruebas estadísticas resalta la importancia de complementar los análisis de

concordancia con metodologías que tengan la precisión suficiente para detectar la variabilidad existente en los sistemas de clasificación cefalométrica de los pacientes.

De este modo, los resultados del presente estudio sugieren que los indicadores cefalométricos comúnmente utilizados pueden ser insuficientes para caracterizar de manera precisa la variabilidad morfológica craneofacial, no existiendo un estándar diagnóstico único que permita clasificar de manera inequívoca, ni la clase esquelética, ni el biotipo. Pensamos que esta limitación se debe a que la cefalometría tradicional, dada la naturaleza lineal de sus datos (puntos y ángulos), no captura del todo la complejidad del macizo craneofacial, teniendo como consecuencia una pérdida significativa de información espacial.^(36,37) En este contexto, nuestros resultados serían relevantes no solo desde el punto de vista metodológico, sino que también en el sentido clínico.

La importancia clínica de nuestros hallazgos radica en que, en el diagnóstico ortodóncico, los indicadores cefalométricos comúnmente utilizados son insuficientes para describir la variación morfológica observada en los pacientes. En otras palabras, no existe un indicador óptimo (“gold standard”) capaz de clasificar correctamente,⁽³⁸⁾ lo que puede deberse a la pérdida de información geométrica ocasionada por la reducción del análisis craneofacial en un conjunto de puntos cefalométricos utilizados en cada indicador.^(36,37) Por este motivo, en las

últimas décadas se han introducido en el área odontológica las herramientas de la morfometría geométrica, definida como el análisis estadístico de la forma y su covariación con otras variables.⁽³⁹⁾ Este enfoque ha sido propuesto como una opción complementaria para el diagnóstico dentoesquelético,^(33,35,40) proporcionando una alternativa a los diversos indicadores cefalométricos comúnmente utilizados para diagnosticar la clase esquelética mediante la captura del patrón intrínseco de variación morfológica craneofacial.

Así, desde una perspectiva clínica, la correcta identificación de la morfología craneofacial es un pilar fundamental en ortodoncia, especialmente en pacientes en crecimiento, donde un diagnóstico preciso permite intervenir tempranamente y guiar el desarrollo esquelético de manera óptima. La falta de concordancia diagnóstica entre los distintos indicadores representa un desafío significativo en la planificación terapéutica, aumentando el riesgo de interpretaciones erróneas que puedan afectar la toma de decisiones clínicas.

En segundo lugar, la variabilidad en la clasificación diagnóstica encontrada en el presente estudio, tanto para clase esquelética como biotipo facial, tiene implicancias metodológicas en la investigación ortodóncica. Numerosos estudios comparan distintas características de la morfología craneofacial entre grupos clasificados según la clase esquelética o el biotipo, asumiendo que los indicadores empleados son fiables y reproducibles.⁽⁴¹⁻⁴³⁾ No obstante, si los criterios de segmentación de estos grupos se basan en indicadores de baja concordancia, las conclusiones derivadas de dichos estudios pueden presentar limitaciones importantes.

Nuestros resultados ponen de manifiesto que clasificar pacientes con características esqueléticas heterogéneas dentro de un mismo grupo de estudio, constituye un problema importante en la investigación ortodóncica. Por ejemplo, agrupar pacientes Clase II dolicofaciales con pacientes Clase II braquifaciales podría inducir sesgos en la interpretación de los resultados, ya que la morfología craneofacial está determinada por una interacción simultánea entre la clase esquelética y el biotipo.⁽⁵⁾ En este sentido, cualquier análisis que no contemple ambas variables de manera conjunta podría comprometer la validez interna de los estudios y limitar la aplicabilidad de sus hallazgos en la práctica clínica.

Dado el panorama descrito, resulta imperativo el desarrollo de herramientas diagnósticas más precisas y discriminantes, como la morfometría geométrica.⁽³⁵⁾ Esta metodología, definida como el análisis estadístico de los componentes de la forma y el tamaño, y su covariación con otras variables, permite evaluar la morfología craneofacial de manera más integral, evitando las limitaciones impuestas por los enfoques cefalométricos tradicionales.⁽³³⁾

En resumen, los resultados del presente estudio ponen de manifiesto las discrepancias diagnósticas existentes entre los distintos indicadores cefalométricos empleados para la clasificación de la clase esquelética y el biotipo. Esta falta de concordancia representa un problema tanto en la práctica clínica como en la investigación ortodóncica, evidenciando la necesidad de explorar nuevas herramientas analíticas que permitan mejorar la precisión y coherencia en el diagnóstico craneofacial.

Conclusiones

1) Al estudiar la concordancia diagnóstica de los cinco indicadores de clase esquelética y los cuatro de biotipo facial, se encontraron diferencias significativas en el diagnóstico obtenido al aplicarlos en el análisis cefalométrico, y 2) En relación con las pruebas que se deben utilizar para conocer el nivel de significancia estadística de la concordancia diagnóstica, los resultados sugieren que para evaluar la concordancia se debe considerar: i) si se trata de dos evaluadores sin discordancias extremas, utilizar el kappa de Cohen "clásico", ii) de dos evaluadores con discordancias extremas (kappa ponderado de Cohen), o, iii) de más de dos evaluadores (J^2 ponderado de Cohen).

REFERENCIAS

1. Sassouni V. A classification of skeletal facial types. *Am J Orthod.* febrero de 1969;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.* septiembre de 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.* noviembre de 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.* noviembre de 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.* 1 de diciembre de 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.* 1 de febrero de 1975;67(2):125-38.
10. Kim Y. WebCeph: Artificial Intelligence Orthodontics and Cloud-Based Orthognatics Platform. [Internet]. [citado 26 de septiembre de 2024]. Disponible en: <https://www.webceph.com>
11. Steiner CC, Hills B. Cephalometrics for you and me. *Am Journal Orthod.* 1953;39(10):729-55.
12. Rodríguez-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.* diciembre de 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.* abril de 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.* 1 de junio de 1995;65(3):223-31.
18. Gregoret J. Instrumentación de la cefalometría estática lateral. En: *Ortodoncia y cirugía ortognatica : diagnóstico y planificación* [Internet]. Barcelona: Espaxs; 1998. p. 227-32.

19. Cohen J. A Coefficient of Agreement for Nominal Scales. *Educ Psychol Meas.* abril de 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.* octubre de 1968;70(4):213-20.
22. Cohen J. Weighted Chi Square: an Extension of the Kappa Method. *Educ Psychol Meas.* abril de 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.* agosto de 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.* enero de 2016;10(01):116-20.
26. Zaffiri VZ, Idrovo JDI, Manríquez G, Díaz A. Determinación de la clase esquelética mediante el indicador de Wits: ¿es correcto diferenciar la norma entre sexos? *Odontoestomatología* [Internet]. 18 de diciembre de 2024 [citado 8 de abril de 2025];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: *Orthodontic Diagnosis and Planning: Their Roles in Preventive and Rehabilitative Dentistry.* Rocky Mountain/Orthodontics; 1982.
28. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull.* noviembre de 1971;76(5):378-82.
29. RStudio Team. RStudio: Integrated Development Environment for R [Internet]. Disponible en: <https://www.rstudio.com/>
30. Matthias Gamer, Jim Lemon, Ian Fellows, Puspendra Singh. Irr: Various Coefficients of Interrater Reliability and Agreement [Internet]. 2005 [citado 15 de agosto de 2025]. 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 [Internet]. 2002 [citado 26 de septiembre de 2024]. p. 1.4-13. Disponible en: <https://CRAN.R-project.org/package=vcd>
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 sagittal intermaxillary diagnosis. A narrative review. *Rev Clínica Periodoncia Implantol Rehabil Oral*. agosto de 2013;6(2):96-101.
39. James Rohlf F, Marcus LF. A revolution 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*. 1 de enero de 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*. abril de 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*. septiembre de 2014;84(5):755-61.
43. Polat OO, Kaya B. Changes in cranial base morphology in different malocclusions. *Orthod Craniofac Res*. noviembre de 2007;10(4):216-21.

Disponibilidad de datos

El conjunto de datos que apoya los resultados de este estudio se encuentra disponible en el Centro de Análisis Cuantitativo en Antropología Dental de la Facultad de Odontología de la Universidad de Chile.

Declaración de conflicto de intereses

Los autores declaran no tener conflictos de interés.

Fuente de financiamiento

Esta investigación no recibió ninguna subvención específica de agencias de financiación en los sectores público, comercial o sin fines de lucro.

Contribución de autoría

NOMBRE Y APELLIDO	COLABORACIÓN ACADÉMICA													
	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. Administración del proyecto | 8. Metodología |
| 2. Adquisición de fondos | 9. Recursos |
| 3. Análisis formal | 10. Redacción - borrador original |
| 4. Conceptualización | 11. Software |
| 5. Curaduría de datos | 12. Supervisión |
| 6. Escritura - revisión y edición | 13. Validación |
| 7. Investigación | 14. Visualización |

Nota de aceptación:

Este artículo fue aprobado por la editora de la revista MSc. Dra. Natalia Tancredi.