

Predicción de la osteoporosis basada en el aprendizaje automático tras análisis de texturas de TC en pacientes con hiperparatiroidismo primario

Antonio Adarve Castro Castro¹, J.M. Castro-García²,
Virginia Soria Utrilla³, Pablo Salvatierra Sánchez¹

1. Hospital Universitario Virgen de la Victoria

2. Hospital General Universitario, Alicante,

3. Hospital Regional Universitario Carlos Haya, Málaga

1. Objetivo.

Evaluar el rendimiento de técnicas de aprendizaje automático supervisado para clasificar la densidad mineral ósea (DMO) normal versus anormal utilizando características clínicas y análisis de textura del tejido óseo espinal en pacientes con hiperparatiroidismo primario (HPP).

2. Material y Método.

Estudio retrospectivo con 219 pacientes diagnosticados de HPP, de los cuales 58 presentaron DEXA y TC que incluía la región lumbar por cualquier motivo entre los 6 meses previos hasta 1 mes después de DEXA. La DMO se midió en UH y para el análisis de textura se trazó un ROI en la primera vértebra lumbar de cada TC utilizando el software LifeX 7.3.0. Se utilizaron técnicas de aprendizaje automático (regresión logística, Naive Bayes, red neuronal y un método de conjunto que utiliza Naive Bayes y red neuronal), que fueron entrenadas con las variables estadísticamente más significativas obtenidas después del análisis univariante. Se calculó la exactitud, precisión, recuperación, puntaje F1 y AUC.

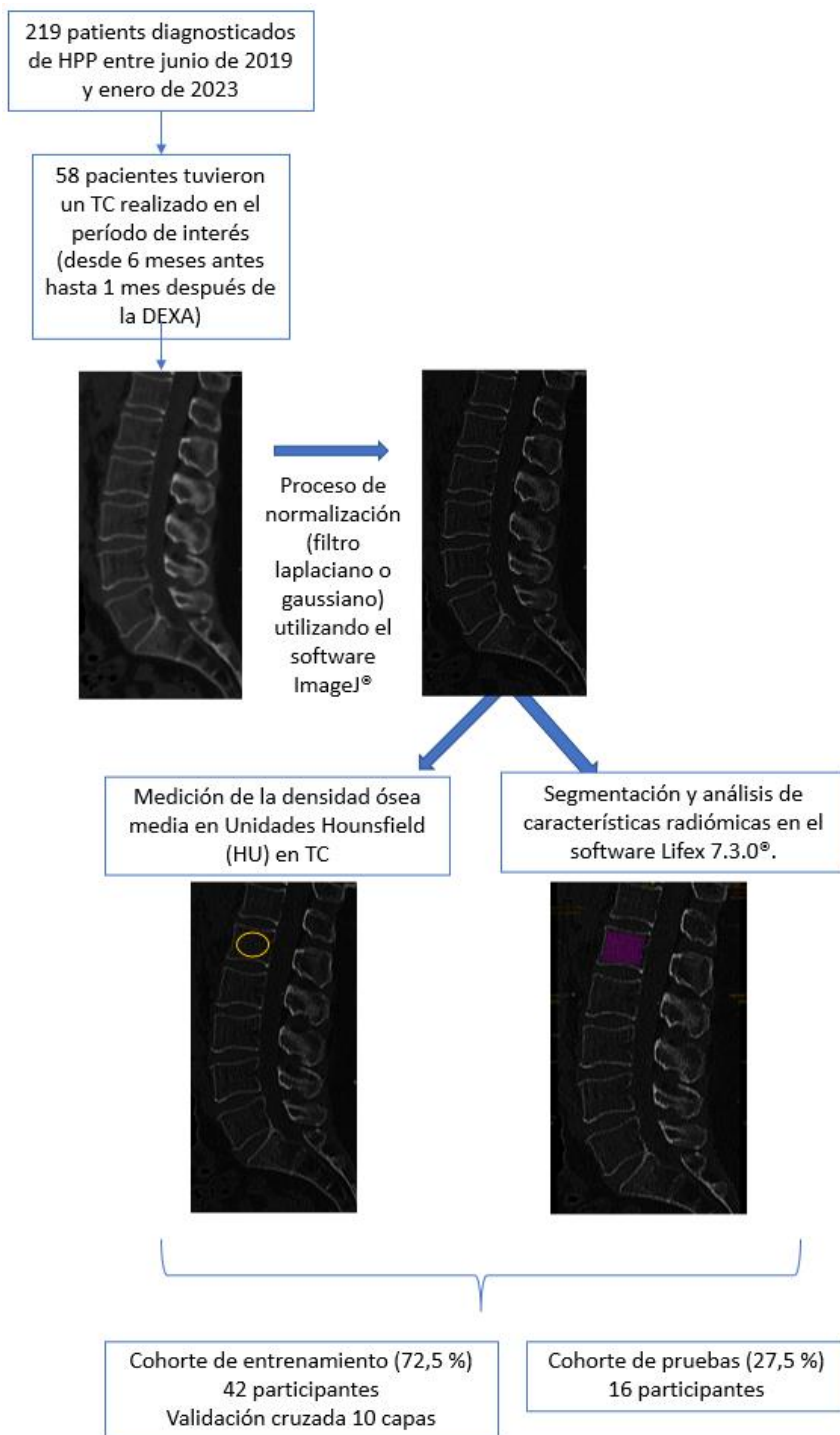


Fig 1. Diagrama de flujo del estudio

3. Resultados.

- Se reclutaron 218 pacientes diagnosticados de HPP, de los cuales a 58 se les había realizado una tomografía computarizada que incluía la región de la columna lumbar. Once de ellos (18,9%) eran hombres con una mediana de edad de 53 (49,5-58,2) años y 47 (81,1%) eran mujeres con una mediana de edad de 57 (52,3-64,8) años

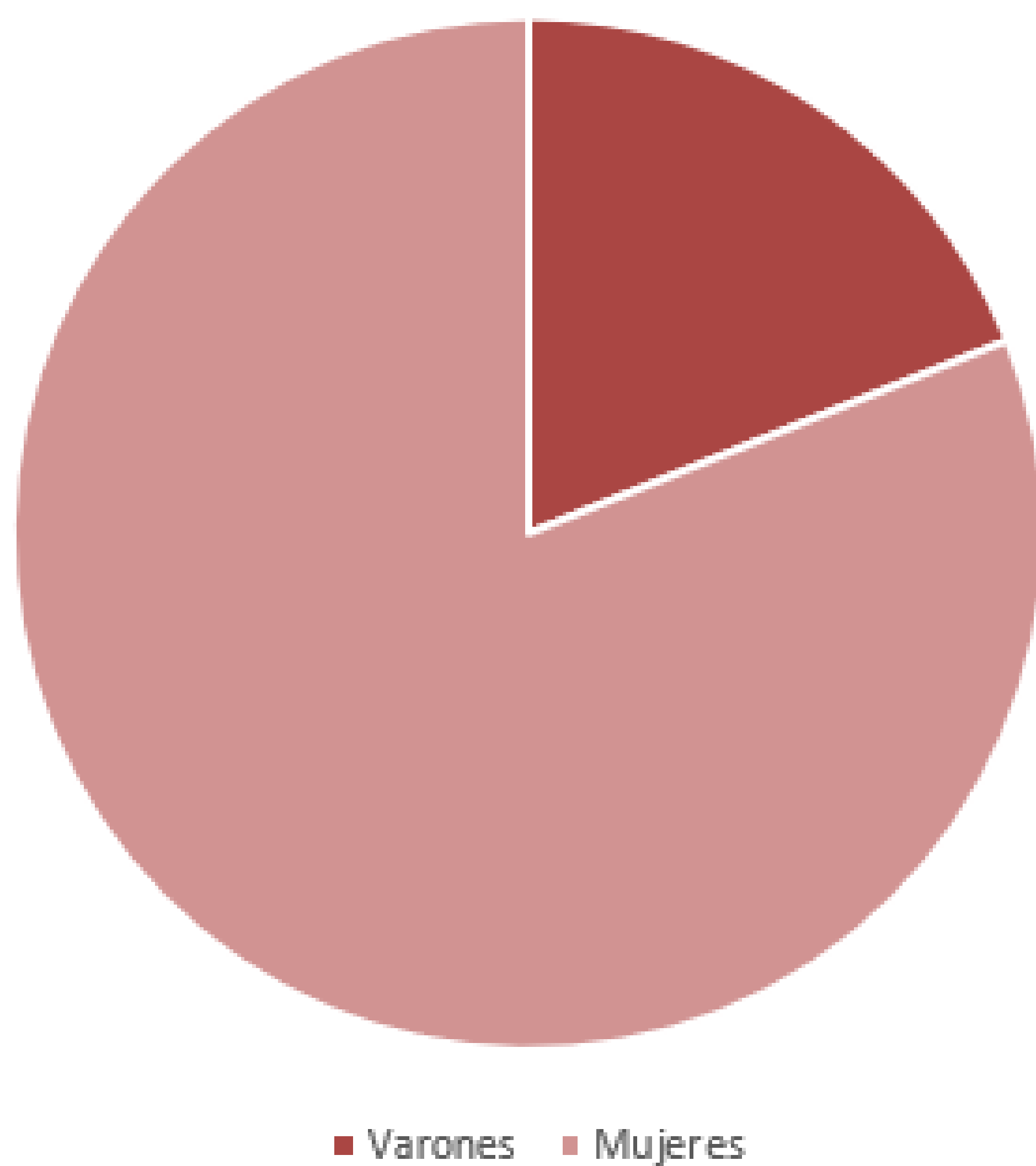


Fig 2. Distribución de pacientes con osteoporosis por género

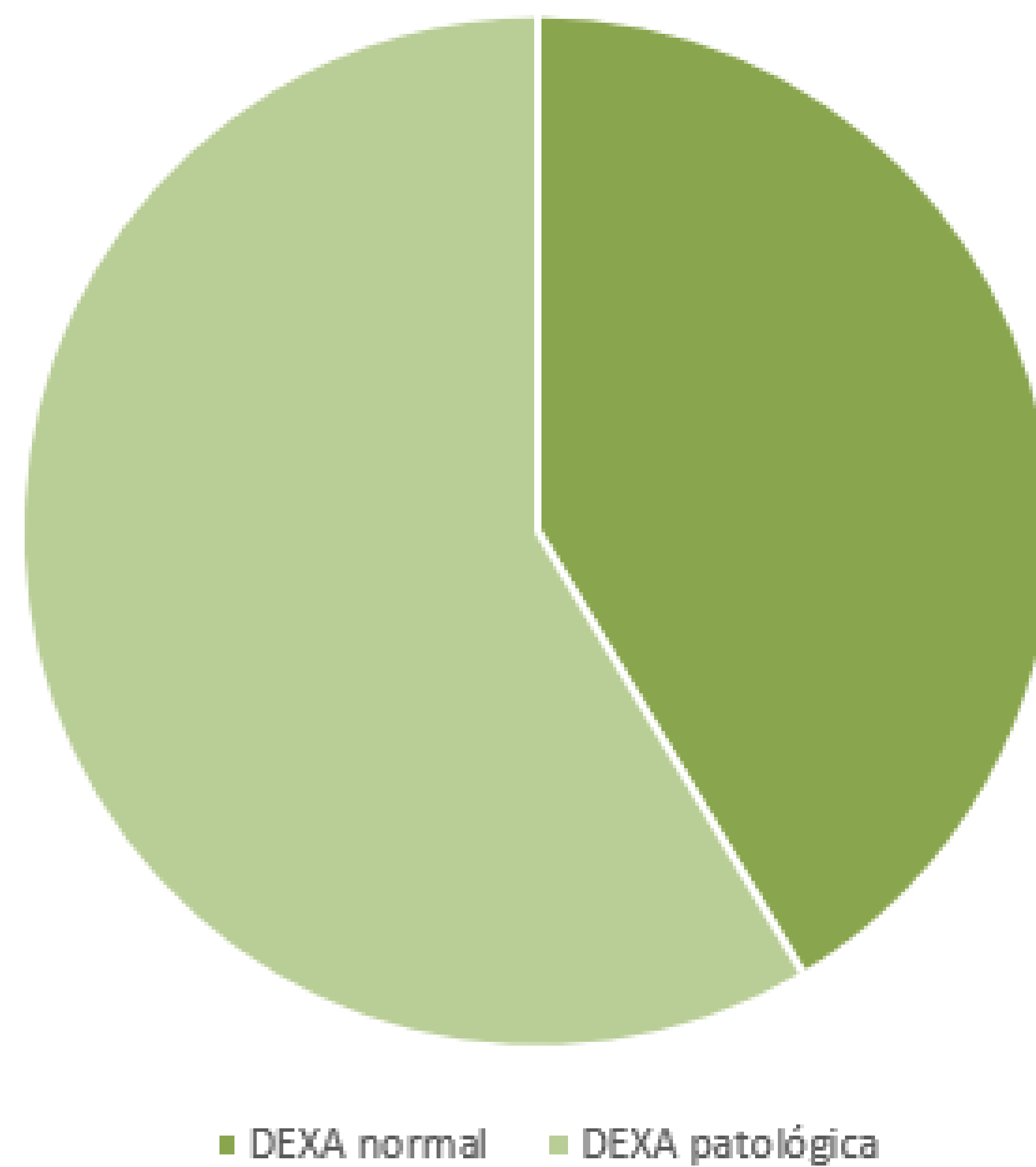


Fig 3. Distribución de DEXA normal y patológica

- En cuanto a los estudios de TC, 34 (58,6%) se realizaron con contraste intravenoso y 24 (41,4%) sin él, arrojando valores medios de UH de $134,5 \pm 36,5$ y $124,5 \pm 33,4$, respectivamente. No se observaron diferencias estadísticamente significativas al comparar los valores de UH entre los estudios con y sin contraste intravenoso ($p > 0,64$).
- Al evaluar UH en los ROI seleccionados de pacientes con DEXA normal vs DEXA anormal, se encontraron diferencias estadísticamente significativas ($p < 0,001$), siendo la UH media de $153 \pm 41,72$ en aquellos con DEXA normal y $107 \pm 33,04$ en aquellos con DEXA anormal (Tabla 4). Obtuvimos que un punto de corte de 106 UH mostró una especificidad del 90% y una sensibilidad del 61% para distinguir la presencia y ausencia de osteoporosis respecto al diagnóstico de DEXA, con un área bajo la curva de 0,84 ($p < 0,001$). Asimismo, un punto de corte de 160 UH proporcionó una especificidad del 89,9% y una sensibilidad del 60% para descartar la presencia de valores diagnósticos de osteoporosis en las exploraciones DEXA, con un área bajo la curva de 0,81 ($p < 0,001$).

ANÁLISIS DE TEXTURA

- Para todos los estudios de TC se obtuvo un conjunto completo de 38 variables de primer orden y 18 variables de segundo orden. Se observaron diferencias estadísticamente significativas en diversas variables de primer y segundo orden al comparar pacientes con DEXA normal y anormal. Entre las variables de primer orden, la intensidad mediana, los percentiles basados en la intensidad y varias medidas estadísticas basadas en la intensidad mostraron diferencias significativas. Para las variables de segundo orden, la falta de uniformidad del nivel de Gray exhibió significación estadística (Tabla 1). Al analizar los resultados del análisis de textura por separado para los estudios de TC con y sin contraste, no se encontraron diferencias estadísticamente significativas entre pacientes con DEXA normal y anormal ($p > 0,05$).

RENDIMIENTO DEL MODELO DE APRENDIZAJE MÁQUINA

- Las variables que se incluyeron para entrenar y probar las cuatro técnicas de aprendizaje automático fueron: UH, Kurtosis IPH6, nivel máximo de gris 84IY y GLRLM nivel de gris no uniforme R5YN. Los resultados de exactitud, precisión, recuperación, F1 y área bajo la curva (curva ROC) para el grupo de entrenamiento después de un análisis basado en validación cruzada de 10 veces y de probarlos en el grupo de prueba se resumen en la Tabla 5. Los valores de área bajo la curva para el grupo de prueba fueron 0,716, 0,783, 0,833 y 0,916 respectivamente para la red neuronal, la regresión logística, Naive Bayes y el modelo conjunto conjunto de red neuronal y Naive Bayes (Tabla 2).

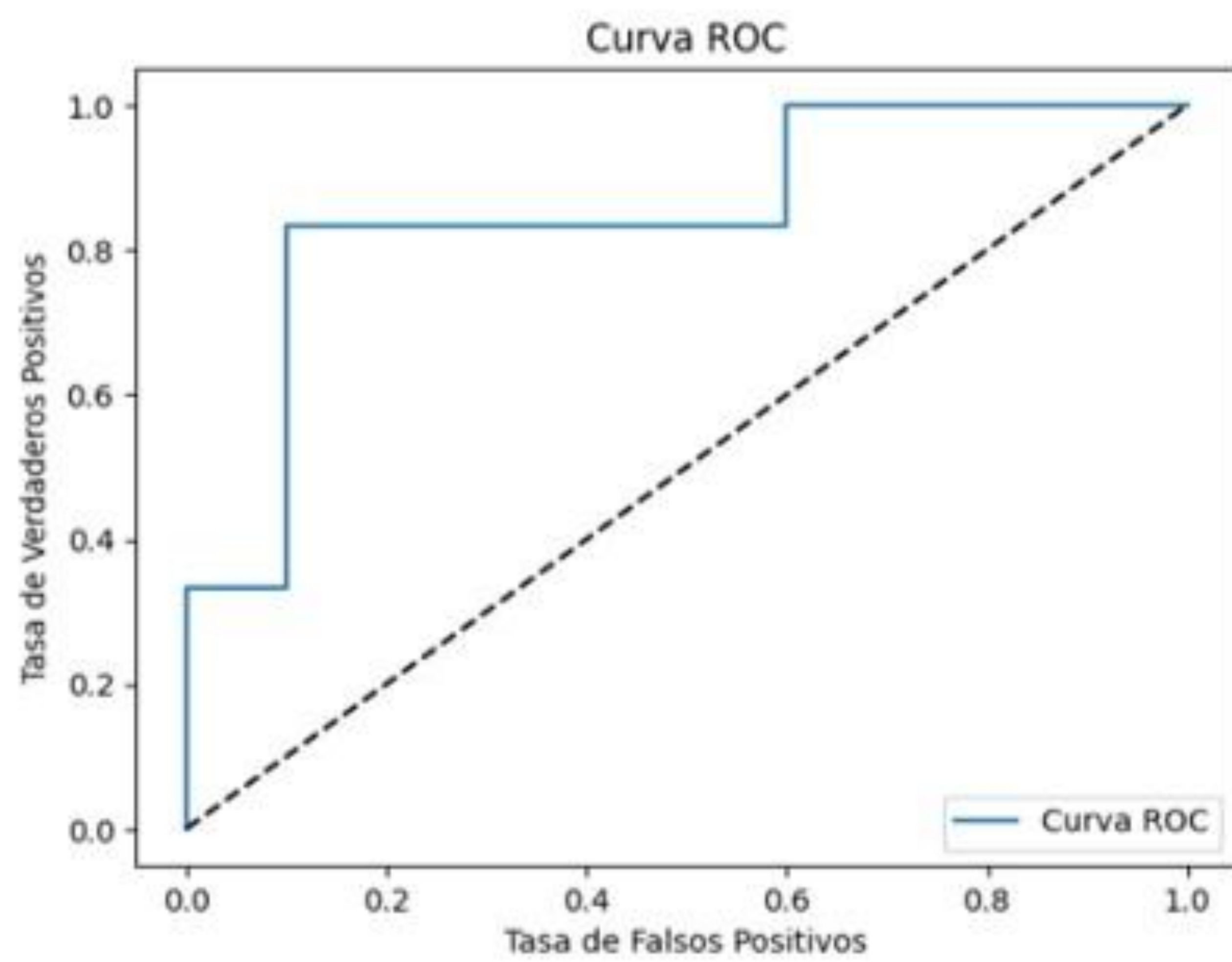
Tabla 1. Variables en pacientes con DEXA normal y anormal.

Variables		Pacientes con DEXA normal (n = 24)	Pacientes con DEXA patológica (n = 34)	p
Edad (años)		55 (51-60.5)	56 (51.75-64.25)	>0.05
Sexo	Hombre	7 (29%)	4 (12%)	>0.05
	Mujer	17 (71%)	30 (88%)	
Tipo de TC	Sin contraste iv	10 (42%)	14 (41%)	>0.05
	Con contraste iv	14 (58%)	20 (59%)	
Valores de UH		153 ± 41.72	107 ± 33.04	<0.001
Análisis de texturas	Intensidad media	19 (18-23.75)	17 (15-19)	<0.001
	Percentil 10 de intensidad de base	5 (4-7.75)	3.5 (2.75-5.25)	<0.05
	Percentil 25 de intensidad de base	12 (10-14.75)	9.5 (8-12)	<0.01
	Percentil 50 de intensidad de base	28 (25.25-32)	24.5 (23-27)	<0.001
	Percentil 90 de intensidad de base	35 (32-40)	32 (29-34.25)	<0.001
	_Kurtosis IBSI:IPH6	0.6 (0.46-0,82)	0.53 (0.37-0.75)	<0.01
	Nivel máximo de gris IBSI:84IY	93 (85-113.75)	81.5 (72-92)	<0.01
	Rango intercuartílico IBSI:SALO	15 (15-18.5)	14 (14-16.26)	<0.05
	Range de intensidad IBSI:20JQ	93 (85-113.75)	81.5 (72-92)	<0.01
	Media absoluta basada en intensidad desviación IBSI:4FUA	9.0051E15 (8.548E15-1.1146E16)	8.12135E+15 (7.26E+15-8.742E+15)	<0.01
	Desviación media absoluta basada en intensidad IBSI:N72L	8.9618E15 (8.3721E15-9.9117E15)	8.138E15 (5.4973E15-9.5189E15)	<0.05
	Intensidad cuadrática media IBSI: 5ZWQ	2.178E16 (7.4589E15-2.4135E16)	1.8066E16 (2.096E15-2.1828E16)	<0.01
	Falta de uniformidad del nivel de gris GLRLM R5YN	222.75 (192.75-252)	235.88 (214.63-248)	<0.01

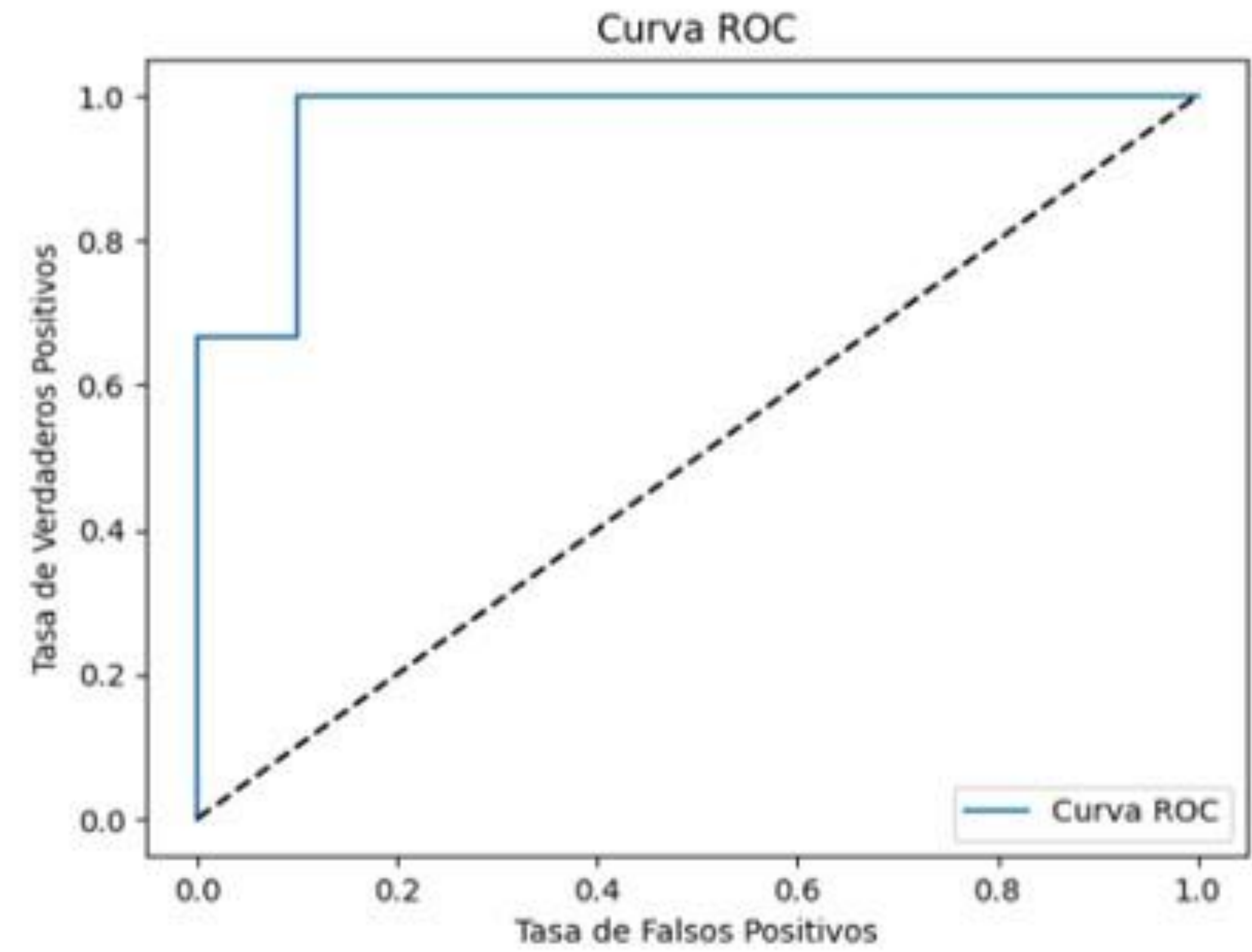
Tabla 2. Resultados de exactitud, precisión, recuperación F1 y curva ROC para el grupo de entrenamiento y prueba.

		Red neuronal	Regresión logística	Naive Bayes	Modelo conjunto de red neuronal y Naive Bayes
Cohorte de entrenamiento	Exactitud	0.645	0.868	0.847	0.881
	Precisión	0.816	0.871	0.854	0.884
	Recuperación	0.70	0.863	0.846	0.880
	F1	0.72	0.868	0.846	0.880
Cohorte de prueba	Exactitud	0.687	0.812	0.875	0.937
	Precisión	0.555	0.800	1.000	1.000
	Recuperación	0.833	0.666	0.666	0.833
	F1	0.666	0.727	0.800	0.909
	Área bajo la curva	0.716	0.783	0.833	0.916

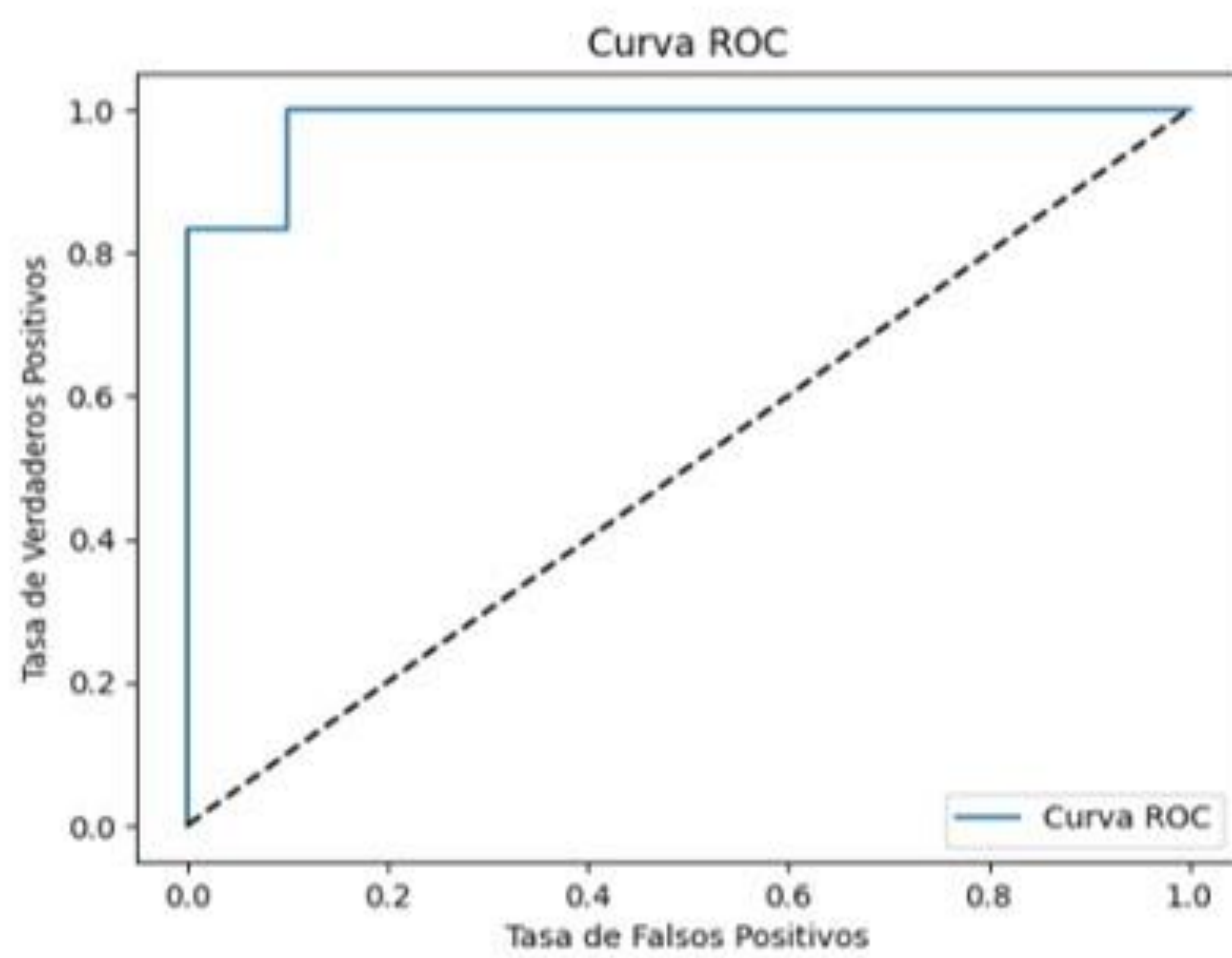
a) Neural network



b) Logistic regression



c) Naive Bayes



d) Mixed model neural network and naive bayes

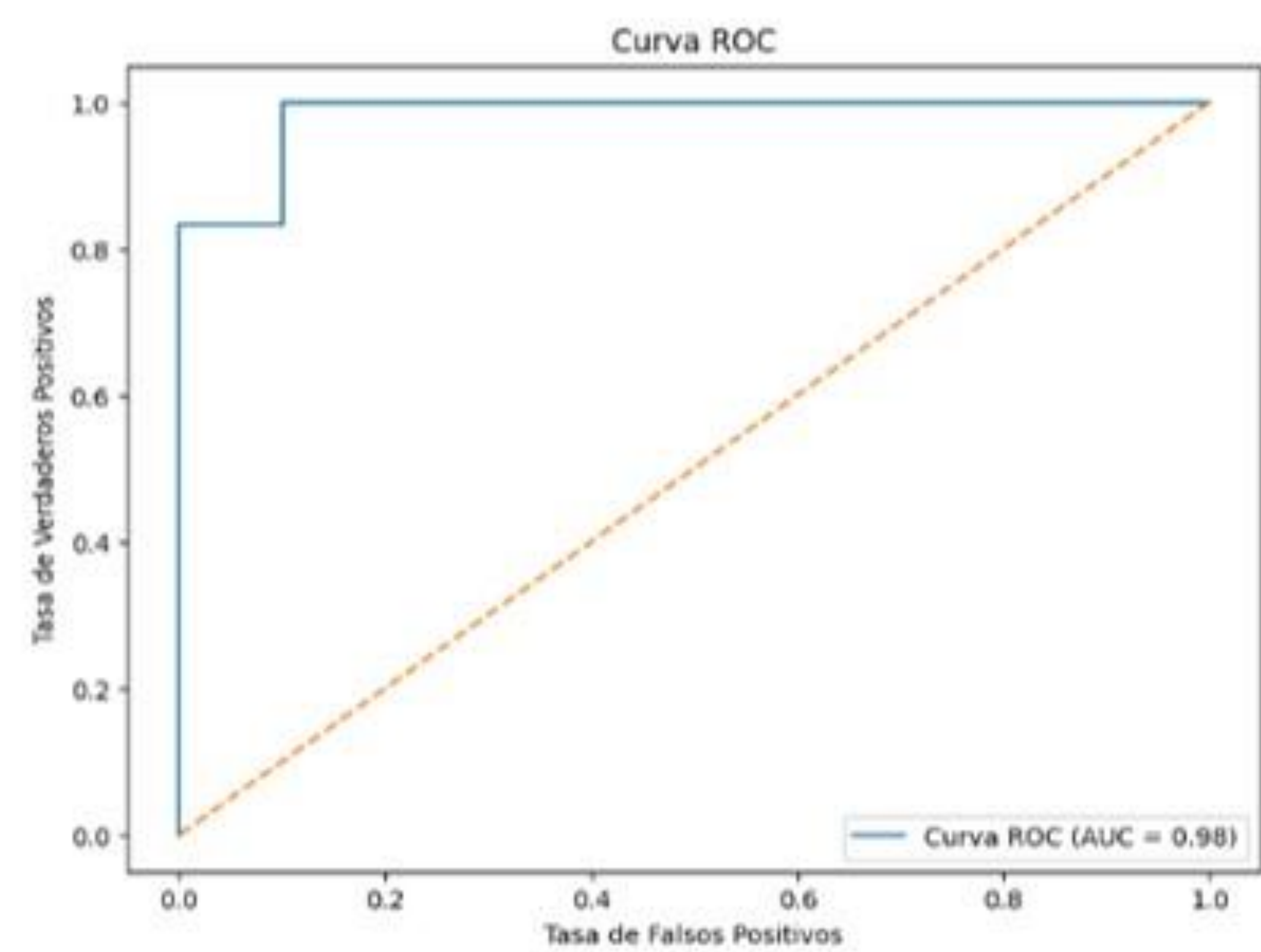


Fig 4. Resultados de la curva ROC de los diferentes modelos de algoritmos de aprendizaje automático (red neuronal, regresión logística, Naive Bayes y modelo de conjunto) para el grupo de prueba.

4. Discusión

- En este estudio, desarrollamos un modelo predictivo para la densidad mineral ósea (DMO) utilizando cuatro técnicas distintas de aprendizaje automático, incorporando variables derivadas del análisis de textura a partir de imágenes de TC del primer cuerpo vertebral lumbar. El modelo de conjunto, basado en una red neuronal de modelo mixto con Naive Bayes, exhibió un alto poder discriminatorio con un área bajo la curva (AUC) de 0,916 para clasificar la DMO normal frente a anormal. Esto sugiere que las características radiómicas extraídas de imágenes de TC de la columna lumbar pueden servir como un biomarcador prometedor para predecir la DMO anormal.
- Nuestros hallazgos revelaron que los valores medios de la Unidad Hounsfield (HU) en pacientes con osteoporosis eran significativamente más bajos que aquellos con densidad ósea normal. En particular, los puntos de corte de HU específicos, como 106 HU con 90 % de especificidad y 61 % de sensibilidad, y 160 HU con 89,9 % de especificidad y 60 % de sensibilidad, demostraron un fuerte potencial diagnóstico para distinguir la presencia y ausencia de osteoporosis en comparación con el diagnóstico DEXA. Estos resultados se alinean con estudios previos de Zou et al. [2] y Pickhardt et al. [6], enfatizando la confiabilidad de los valores de HU en el diagnóstico de osteoporosis.
- El análisis de textura, un enfoque novedoso en nuestro estudio, mostró resultados comparables a la literatura existente para distinguir pacientes con y sin osteoporosis. Si bien estudios anteriores exploraron el análisis de textura en diferentes regiones anatómicas, nuestro enfoque en las imágenes de TC de la columna lumbar, típicamente el sitio para DEXA, agrega especificidad a la predicción de la osteoporosis. [8,13, 14,15].

- Las características radiómicas como la curtosis y la falta de uniformidad del nivel de gris, utilizadas por otros investigadores [8,9], fueron parte integral de nuestros modelos predictivos. La curtosis refleja el pico de distribución de intensidad [10,16], y la falta de uniformidad del nivel de gris mide la variabilidad, ambas importantes para evaluar las alteraciones de la microestructura ósea debido a la pérdida de DMO [16, 9].
- A diferencia de estudios anteriores sobre patología de la columna, nuestra investigación se concentró exclusivamente en pacientes con cambios en la densidad ósea atribuidos al hiperparatiroidismo primario (PHP) [2,7,10]. La creciente disponibilidad de estudios de TC mejora la aplicabilidad de nuestro modelo radiómico, lo que podría impulsar más pruebas DEXA confirmatorias en caso de resultados patológicos y minimizar los costos innecesarios de pruebas para resultados normales [6,8].
- Sin embargo, nuestro estudio tiene limitaciones, incluido un diseño de centro único con un tamaño de muestra relativamente pequeño. Las técnicas de segmentación manual y normalización de imágenes pueden introducir sesgos de medición y limitar la significación estadística, respectivamente. No obstante, la estrategia de reclutamiento consecutivo de pacientes mitiga el sesgo de selección. Es importante destacar que nuestro enfoque, a pesar de las limitaciones, presenta un método de detección simple y rentable para predecir la DMO anormal en pacientes con PHP, abriendo vías para estudios prospectivos que validen su utilidad en la práctica clínica.

- El estudio también destaca la importancia de los valores de la Unidad Hounsfield (HU) para discernir la osteoporosis, mostrando puntos de corte específicos con una especificidad y sensibilidad impresionantes en comparación con el estándar de oro, DEXA. Estos hallazgos se alinean con la literatura existente y validan la utilidad de los valores de HU como una herramienta de diagnóstico confiable para la osteoporosis en el contexto de PHP.
- Además, nuestra exploración del análisis de textura, centrada específicamente en imágenes de TC de la columna lumbar, contribuye al creciente conjunto de pruebas que respaldan su aplicabilidad en la predicción de la osteoporosis. Las características radiómicas como la curtosis y la falta de uniformidad del nivel de gris surgieron como componentes cruciales en nuestros modelos predictivos, enfatizando su potencial como indicadores significativos de las alteraciones de la microestructura ósea relacionadas con la pérdida de DMO.
- A pesar de las limitaciones del estudio, incluido un diseño de centro único y un tamaño de muestra relativamente modesto, nuestro enfoque ofrece un método de detección valioso y rentable para predecir la DMO anormal en pacientes con PHP. La simplicidad y accesibilidad de nuestro modelo en la práctica clínica lo convierten en una herramienta prometedora para identificar personas en riesgo de osteoporosis, guiar futuras decisiones de diagnóstico y reducir potencialmente los costos de pruebas innecesarias.

5. Conclusiones.

Las técnicas de aprendizaje automático supervisadas permiten una clasificación precisa de la DMO normal versus anormal aprovechando las características clínicas y el análisis de la textura del tejido óseo espinal en pacientes con HPP sin gastos adicionales.

6. Referencias.

1. Kerschán-Schindl K. (2016) Prevention and rehabilitation of osteoporosis. *Wien Med Wochenschr.* DOI: 10.1007/s10354-015-0417-y
2. Zou D, Li W, Deng C, Du G, Xu N (2019) The use of CT Hounsfield unit values to identify the undiagnosed spinal osteoporosis in patients with lumbar degenerative diseases. *Eur Spine J.* DOI: 10.1007/s00586-018-5776-9
3. Vitetta GM, Ravera A, Mensa et al (2019) Actual role of color-doppler high-resolution neck ultrasonography in primary hyperparathyroidism: a clinical review and an observational study with a comparison of 99mTc-sestamibi parathyroid scintigraphy. *J Ultrasound.* DOI: 10.1007/s40477-018-0332-3
4. Barsony J. (2014) Diagnostic and Laboratory Evaluation of Hyperparathyroidism. In: Terris DJ, Duke WS, Pasiaka JL, editors. *Parathyroid Surgery. Fundamental and Advanced Concepts.* San Diego: Plural Publishing; p. 23–36.
5. Kanis JA, Cooper C, Rizzoli R, Reginster JY. (2019) European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporos Int.* DOI: 10.1007/s00198-018-4704-5
6. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N. (2013) Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. *Ann Intern Med.* DOI: 10.7326/0003-4819-158-8-201304160-00003
7. Choi MK, Kim SM, Lim JK. (2016) Diagnostic efficacy of Hounsfield units in spine CT for the assessment of real bone mineral density of degenerative spine: correlation study between T-scores determined by DEXA scan and Hounsfield units from CT. *Acta Neurochir (Wien).* DOI: 10.1007/s00701-016-2821-5.
8. Jiang YW, Xu XJ, Wang R, Chen CM. (2022) Radiomics analysis based on lumbar spine CT to detect osteoporosis. *Eur Radiol.* DOI: 10.1007/s00330-022-08805-4.
9. Biamonte E, Levi R, Carrone F et al. (2022) Artificial intelligence-based radiomics on computed tomography of lumbar spine in subjects with fragility vertebral fractures. *J Endocrinol Invest.* DOI: 10.1007/s40618-022-01837-z.
10. Xue Z, Huo J, Sun X et al. (2022) Using radiomic features of lumbar spine CT images to differentiate osteoporosis from normal bone density. *BMC Musculoskelet Disord.* DOI: 10.1186/s12891-022-05309-6.
11. Sebro R, De la Garza-Ramos C. (2023) Opportunistic screening for osteoporosis and osteopenia from CT scans of the abdomen and pelvis using machine learning. *Eur Radiol.* DOI: 10.1007/s00330-022-09136-0.
12. Yasaka K, Akai H, Mackin D et al. (2017) Precision of quantitative computed tomography texture analysis using image filtering. *Medicine (Baltimore).* DOI: 10.1097/MD.00000000000006993.
13. Kawashima Y, Fujita A, Buch K et al. (2019) Using texture analysis of head CT images to differentiate osteoporosis from normal bone density. *Eur J Radiol.* DOI: 10.1016/j.ejrad.2019.05.009.
14. Valentinič A, Trebeschi S, Kaesmacher J et al. (2019) Opportunistic osteoporosis screening in multi-detector CT images via local classification of textures. *Osteoporos Int.* DOI: 10.1007/s00198-019-04910-1.
15. Liyu Liu, Meng Si, Hecheng Ma et al. (2022) A hierarchical opportunistic screening model for osteoporosis using machine learning applied to clinical data and CT images. *BMC Bioinformatics.* DOI: 10.1186/s12859-022-04596-z
16. Radiomic Features, Pyradiomics (2023) Available via <https://pyradiomics.readthedocs.io/en/latest/features.html>. Accessed 20 Jul 2023