



VNIVERSITAT ID VALÈNCIA

Facultat de Medicina i Odontologia
Programa de Doctorado: 3139 MEDICINA
Línea de Investigación: APARATO LOCOMOTOR

Aplicación del análisis de textura a partir de imágenes de RM convencional en el estudio de pacientes con lumbalgia mecánica crónica

TESIS DOCTORAL

Presentada por:

Vicente José Climent Peris

DIRECTORES:

Dr. Luís Martí Bonmatí

Dr. Julio Doménech Fernández

Dr. Antonio Silvestre Muñoz

Valencia, marzo de 2023

INFORME DIRECTORES/AS Y TUTOR/A PARA DEPÓSITO DE TESIS

Director (es) / Codirector (es):

1.- Apellidos y nombre: Martí Bonmatí, Luis. N.I.F. 19838946C, Departamento/Instituto: Área clínica de imagen médica. Centro: Hospital Universitario y Politécnico La Fe. Valencia

2.- Apellidos y nombre: Silvestre Muñoz, Antonio N.I.F. 22686192G, Departamento/Instituto: Servicio de Cirugía Ortopédica y Traumatología. Hospital Clínico Universitario de Valencia. Departamento de Cirugía. Universitat de Valencia

3.- Apellidos y nombre: Doménech Fernández, Julio. N.I.F. 28989752T, Departamento/Instituto: Servicio de Cirugía Ortopédica. Centro: Clínica Universitaria de Pamplona. Pamplona

Tutor o tutora (si procede)

Apellidos y nombre: _____, N.I.F. _____, Departamento/Instituto: _____ Centro: _____

Directores/as y tutor/a, respectivamente, de la tesis doctoral: "Aplicación del análisis de textura a partir de imágenes de RM convencional en el estudio de pacientes con lumbalgia mecánica crónica"

de D/Dña. Vicente José Climent Peris,

estudiante del Programa de Doctorado **3139 Medicina** (RD99/2011) en Medicina de la Universitat de València, emiten informe favorable para la realización del depósito y la defensa de la tesis doctoral.

Fecha: 14 febrero de 2023

Luis Martí
Bonmatí
Firmado digitalmente por Luis Martí Bonmatí
Fecha: 2023.02.13 22:43:18 +01'00'

ANTONIO
SILVESTRE
MUÑOZ -
NIF:22686192G
Firmado digitalmente por ANTONIO SILVESTRE MUÑOZ -
NIF:22686192G
Fecha: 2023.02.16 18:27:57 +01'00'

JULIO|
DOMENECH
|
FERNANDEZ
Firmado digitalmente por JULIO|DOMENECH|
FERNANDEZ
Fecha: 2023.02.15 15:47:42 +01'00'

Fdo.: Dr Luís Martí Bonmatí

Fdo.: Dr Antonio Silvestre Muñoz

Fdo.: Dr Julio Doménech Fernández

Director 1

Director 2

Director 3

ESCUELA DOCTORAL
UNIVERSITAT DE VALÈNCIA

Facultat de Medicina i Odontologia, Av.Blasco.Ibañez nº 15, 46010-Valencia, Tel. 963864100 - Fax: 963864173
email: doctorado.medicina@uv.es
Web: www.uv.es/doctoratmedicina

A Dolo, mi mujer, a Silvia, Javier y Sergio, mis hijos

Agradecimientos

En primer lugar, quiero agradecer al Dr. Julio Doménech Fernández, en la actualidad director del departamento de Cirugía Ortopédica de la Clínica Universidad de Navarra, por ser el principal responsable de que me embarcara en este proyecto y que planteásemos un estudio relacionado con patología del raquis. Su experiencia tanto en lo relativo al tratamiento quirúrgico como en aspectos menos conocidos como la etiología de los distintos procesos y su fisiopatología, así como en la metodología de la investigación ha sido de inestimable ayuda durante todo el proceso de elaboración de esta Tesis Doctoral. En este largo viaje es imprescindible tener personas a tu lado que te ayuden con su experiencia y que te animen en los momentos de dificultad.

Esta investigación no hubiera sido posible sin la inestimable colaboración del Dr. Luís Martí Bonmatí, director del área clínica de imagen médica del Hospital Universitario y Politécnico La Fe de Valencia. He tenido la inmensa suerte de conocer a una gran persona que ha puesto a mi disposición su toda su experiencia en análisis de imagen y me ha permitido conocer un mundo hasta ahora desconocido para mí: la imagen médica más allá de lo que podemos ver con nuestros propios ojos. Siempre pendiente de la evolución del estudio, me ha transmitido la ilusión por seguir adelante.

Ha sido para mí un honor contar con la dirección en este proyecto del Dr. Antonio Silvestre Muñoz, jefe de servicio de COT del hospital Clínico Universitario de Valencia y profesor de la Universitat de Valencia. Agradezco profundamente poner a mi disposición su experiencia investigadora en COT y su ayuda especialmente en la fase de redacción de la memoria de esta Tesis Doctoral.

Mi más profundo agradecimiento por su colaboración a Ángel Alberich Bayarri y a Amadeo Ten Esteve, ingenieros biomédicos del grupo de investigación de imagen biomédica GIBI ²³⁰, por haberme instruido en este campo en el que ellos son expertos y por su trabajo en el análisis de texturas. También agradezco la inestimable ayuda de José Miguel Carot Sierra, catedrático en la escuela técnica superior de ingeniería industrial de la Universidad Politécnica de Valencia, responsable del complejo análisis estadístico de los datos y de Alejandro Rodríguez Ortega, investigador del departamento de ingeniería gráfica de la Universidad Politécnica de Valencia responsable del procesado de las variables de textura y del análisis estadístico que se ha llevado a cabo en este estudio.

Agradecer el apoyo incondicional de mi mujer e hijos, sin el cual no hubiera sido posible llegar al final de este difícil camino, así como los ánimos recibidos de familiares, amigos y compañeros de profesión.

Por último, quiero tener un recuerdo especial para mi padre, del que heredé la pasión por la medicina y que, allí donde esté, estará orgulloso de mi trayectoria en esta profesión.

Índice

ÍNDICE	1
INDICE DE TABLAS	5
ÍNDICE DE FIGURAS	7
ABREVIATURAS.....	9
1 INTRODUCCIÓN	11
1.1 CONCEPTOS DE DOLOR LUMBAR INESPECÍFICO Y LUMBALGIA MECÁNICA	13
1.2 EPIDEMIOLOGÍA DE LA LUMBALGIA MECÁNICA	15
1.3 APLICACIÓN DE LA RESONANCIA MAGNÉTICA (RM) EN EL ESTUDIO DEL DOLOR LUMBAR CRÓNICO	15
1.3.1 DEGENERACIÓN DEL DISCO INTERVERTEBRAL	16
1.3.2 PATOLOGÍA DE LOS PLATILLOS VERTEBRALES	19
1.3.3 PATOLOGÍA DE LOS ELEMENTOS POSTERIORES.....	20
1.3.4 ESTUDIO DE LA MUSCULATURA PARAVERTEBRAL.....	21
1.3.5 CORRELACIÓN CLÍNICO-RADIOLÓGICA.....	22
1.4 ANÁLISIS CUANTITATIVOS MEDIANTE RM.....	23
1.5 PRONÓSTICO DEL DOLOR LUMBAR INESPECÍFICO	24
1.6 RADIÓMICA: ANÁLISIS DE TEXTURA	25
1.7 PARÁMETROS DE TEXTURA (IMAGE FEATURES).....	25
1.8 PROCESO DEL ANÁLISIS RADIÓMICO.....	27
1.8.1 OBTENCIÓN DE LAS IMÁGENES	27
1.8.2 SEGMENTACIÓN DE LAS IMÁGENES.....	28
1.8.3 EXTRACCIÓN DE LOS PARÁMETROS DE TEXTURA	28
1.8.4 SELECCIÓN DE PARÁMETROS DE INTERÉS.....	29
1.8.5 MODELO PREDICTIVO.....	29
1.9 APLICACIONES CLÍNICAS DEL ANÁLISIS DE TEXTURA	30
1.10 APLICACIÓN DEL ANÁLISIS DE TEXTURA EN EL ESTUDIO DE LA COLUMNA VERTEBRAL.....	31

<u>2</u>	<u>HIPÓTESIS</u>	<u>33</u>
<u>3</u>	<u>OBJETIVOS</u>	<u>37</u>
3.1	OBJETIVO PRINCIPAL.....	39
3.2	OBJETIVOS SECUNDARIOS.....	39
<u>4</u>	<u>MATERIAL Y MÉTODOS</u>	<u>41</u>
4.1	DISEÑO DEL ESTUDIO	43
4.2	CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN	43
4.3	TAMAÑO MUESTRAL	44
4.4	SELECCIÓN DE LA MUESTRA	44
4.5	VARIABLES A ESTUDIO	44
4.5.1	VARIABLES CLÍNICAS.....	44
4.5.2	VARIABLES BASADAS EN LA RM LUMBAR.....	46
4.6	PROCEDIMIENTO	51
4.6.1	EVALUACIÓN CLÍNICA	51
4.6.2	ESTUDIO DE IMAGEN: RM LUMBAR	52
4.6.3	ANÁLISIS DE LA IMAGEN DE RM.....	52
4.7	ANÁLISIS DE DATOS	57
4.7.1	ESTADÍSTICOS DESCRIPTIVOS	57
4.7.2	ANÁLISIS DE LAS VARIABLES CUALITATIVAS	57
4.7.3	ANÁLISIS DE LAS VARIABLES DE TEXTURA.....	58
<u>5</u>	<u>ASPECTOS ÉTICOS.....</u>	<u>63</u>
<u>6</u>	<u>RESULTADOS</u>	<u>67</u>
6.1	CARACTERÍSTICAS DE LA MUESTRA.....	69
6.2	RESULTADOS DEL ANÁLISIS DE LAS VARIABLES CUALITATIVAS.....	71
6.2.1	GRADO DE DEGENERACIÓN DISCAL	71
6.2.2	CAMBIOS DE MODIC	76
6.2.3	ZONAS DE ALTA INTENSIDAD	77
6.2.4	DEGENERACIÓN DE LA MUSCULATURA RETROSOMÁTICA	77
6.2.5	ANÁLISIS DE REGRESIÓN MÚLTIPLE	78
6.3	RESULTADOS DEL ANÁLISIS DE TEXTURA	79
6.3.1	PARÁMETROS DE TEXTURA SELECCIONADOS	79

6.3.2	CORRELACIÓN DE LAS VARIABLES CLÍNICAS CON LOS PARÁMETROS DE TEXTURA	80
6.3.3	MODELOS DE REGRESIÓN.....	86
6.3.4	MODELOS DE CLASIFICACIÓN	88
7	<u>DISCUSIÓN.....</u>	<u>97</u>
7.1	HALLAZGOS PRINCIPALES	99
7.2	COMPARACIÓN CON OTROS ESTUDIOS	103
7.3	APLICACIONES CLÍNICAS DE LOS RESULTADOS.....	106
7.4	LIMITACIONES DEL ESTUDIO	107
7.5	LÍNEAS FUTURAS DE INVESTIGACIÓN.....	108
8	<u>CONCLUSIONES.....</u>	<u>109</u>
9	<u>BIBLIOGRAFÍA</u>	<u>113</u>
10	<u>ANEXOS</u>	<u>127</u>
	<u>ANEXO 1. HOJA DE INFORMACIÓN AL PACIENTE</u>	<u>129</u>
	<u>ANEXO 2. DOCUMENTO DE CONSENTIMIENTO INFORMADO ...</u>	<u>133</u>
	<u>ANEXO 3. CUESTIONARIO DE DISCAPACIDAD ROLLAND MORRIS</u>	<u>135</u>
	<u>ANEXO 4. INFORME CEI.....</u>	<u>137</u>
	<u>ANEXO 5. PUBLICACIÓN.....</u>	<u>139</u>

Índice de tablas

Tabla 1. Parámetros de textura básicos incluidos en la biblioteca PyRadiomics (53).....	48
Tabla 2. Características demográficas y clínicas de la muestra.	70
Tabla 3. Hallazgos estructurales de la muestra de 470 discos.	71
Tabla 4. Mayor grado de degeneración discal de cada paciente.	71
Tabla 5. Coeficiente de correlación de Spearmann entre el grado de degeneración discal en los distintos niveles y la puntuación END inicial.	72
Tabla 6. Coeficiente de correlación de Spearman entre el grado de degeneración discal y la puntuación END a los 6 meses.	73
Tabla 7. Coeficiente de correlación de Spearman entre el grado de degeneración discal y la puntuación RMQ inicial.	75
Tabla 8: Coeficiente de correlación de Spearman entre el grado de degeneración discal y la puntuación RMQ a los 6 meses.	75
Tabla 9. Cambios de Modic observados según tipo y localización.	76
Tabla 10: Zonas de alta intensidad (ZAI) y su localización.....	77
Tabla 11. Parámetros de textura seleccionados.....	79
Tabla 12. Resultados de la prueba de correlación de Spearman de los valores de los parámetros de textura con el valor END a los 6 meses.	80
Tabla 13. Resultados de la prueba U de Mann Whitney mediante la que se compara la distribución de los valores de las distintas variables de textura entre el grupo de pacientes con mejoría significativa del dolor y los que no mejoran.	82
Tabla 14. Resultados <i>Random Forest</i> de regresión en los distintos modelos analizados..	88

Tabla 15. Resultados obtenidos, aplicando RF de clasificación, con los distintos modelos entrenados en el 70 % de la muestra cuando los aplicamos al 30 % restante de la misma, para identificar pacientes con mejoría de END menor o igual al 30 %. 89

Tabla 16. Matriz de confusión sobre el 30 % de la muestra del modelo predictivo que incluye variables de textura seleccionadas y todas las variables sociodemográficas y clínicas, aplicando PCA, para predecir mejoría de END menor o igual al 30%. 90

Tabla 17. Resultados obtenidos, aplicando RF de clasificación, con los distintos modelos entrenados en el 70 % de la muestra cuando los aplicamos al 30 % restante de la misma, para identificar pacientes con discapacidad a los 6 m (RMQ > 4). 93

Índice de figuras

Figura 1. Tipos de Zonas de Alta Intensidad (ZAI). Forma redonda (A y D), cisura (B), vertical (C), oblicua (E) y alargada (F) (21).	18
Figura 2. Cambios Modic tipo I en imágenes de RM potenciadas en T1 (izquierda) y T2 (derecha).	19
Figura 3. Cambios Modic tipo II en imágenes de RM potenciadas en T1 (izquierda) y T2 (derecha).	20
Figura 4. Grados de degeneración de la musculatura retrosomática (26) en imágenes de RM transversales potenciadas en T1.	21
Figura 5. Proceso del análisis radiómico (48).	30
Figura 6. Proceso y análisis de las imágenes de RM realizado.	53
Figura 7. ROIs segmentadas en el corte sagital medio en la secuencia potenciada en T2.	54
Figura 8. ROIs segmentadas en el corte axial en la secuencia potenciada en T2.	55
Figura 9. Esquema del análisis estadístico realizado, incluyendo las variables clínicas y los parámetros de textura.	62
Figura 10. Diagrama de flujo del proceso de reclutamiento y seguimiento.	69
Figura 11. Grado de degeneración del disco L4-L5 en pacientes con mejoría de la puntuación END clínicamente relevante o sin ella ($p = 0,011$).	74
Figura 12. Valores del parámetro de textura GLSZM Zone Variance calculado en ROI de platillo vertebral ($p < 0,001$).	83
Figura 13. Valores del parámetro de textura Mean Absolute Deviation calculado en ROI disco intervertebral ($p = 0,001$).	83

Indice

Figura 14. Valores del parámetro de textura de primer orden Percentil 10 calculado en ROI disco intervertebral ($p = 0,001$).	84
Figura 15. Valores del parámetro de textura GLCM lmc 1 calculado en ROI disco intervertebral ($p = 0,001$).	84
Figura 16. Valores del parámetro de textura NGTDM Busyness calculado en ROI disco intervertebral ($p = 0,001$).....	85
Figura 17. Valores del parámetro de textura GLDM Large Dependence Low Gray Level Emphasis calculado en ROI Disco intervertebral ($p < 0,001$).	85
Figura 18. Distribución de los casos según el valor real y la predicción del valor END (escala numérica del dolor) a los 6 meses del modelo que incluye todas las variables.	86
Figura 19. Distribución de los casos según el valor real y la predicción del valor RMQ (escala Rolland Morris) a los 6 meses del modelo que incluye todas las variables.....	87
Figura 20. Curva ROC del modelo predictivo que incluye las variables de textura y todas las variables clínicas, aplicando PCA (análisis de componentes principales), para predecir mejoría de END menor o igual al 30 %. $AUC=0,71$	90
Figura 21. Nomograma de Nolan del modelo predictivo con mejor rendimiento para predecir mejoría de END (escala numérica del dolor) menor o igual al 30 %.	91
Figura 22. Nomograma de Nolan del modelo predictivo con mejor rendimiento para predecir puntuación en $RMQ > 4$ al final del seguimiento....	95

Abreviaturas

AUC: área bajo la curva ROC (*area under curve*)
CEST: *chemical exchange saturation transfer*
DE: desviación estándar
DICOM: *digital imaging and communication in medicine*
END: escala numérica del dolor
GLCM: *gray-level co-occurrence matrix*
GLDM: *gray-level dependence matrix*
GLRLM: *gray-level run-length matrix*
GLSZM: *gray-level size zone*
IBSI: *imaging biomarker standardization*
ICC: índice de correlación interclase
IMC: índice de masa corporal
MAD: desviación absoluta de la mediana
MAE: error absoluto medio
MRI: magnetic resonance image
NGTDM: *neighboring gray tone difference matrix*
ÖMPSQ: *Örebro Musculoskeletal Pain Screening Questionnaire*
OMS: organización mundial de la salud
PCA: análisis de componentes principales (*principal components analysis*)
RM: resonancia magnética
RMQ: *Rolland-Morris disability questionnaire*
RMSE: error cuadrático medio
ROC: característica operativa del receptor (*receiver operating characteristic*)
ROI: región de interés (*region of interest*)
RX: rayos X
SHAP: SHapley Additive exPlanations
T1: tiempo de relajación longitudinal
T2: tiempo de relajación transversal
TC: tomografía computadorizada
VPP: valor predictivo positivo
VPN: valor predictivo negativo
ZAI: zonas de alta intensidad

1 Introducción

Introducción

La lumbalgia mecánica inespecífica es en la actualidad uno de los principales problemas sanitarios de las sociedades desarrolladas, con una elevada prevalencia y un elevado porcentaje de cronificación. La OMS reconoce el dolor lumbar como una de las principales causas de discapacidad y de interferencia con la actividad laboral. Así, se ha publicado que el 84% de los individuos lo sufrirá a lo largo de su vida, que la prevalencia del dolor lumbar crónico es del 23% y que esta produce algún grado de discapacidad hasta en el 12% de la población (1). La mayoría de estos pacientes tienen estudios de imagen de la columna lumbar, principalmente con radiografía simple y resonancia magnética (RM). A pesar de ello debemos reconocer que no disponemos de herramientas diagnósticas que permitan establecer con precisión la relación causal entre los hallazgos observados (imagen, neurofisiológicos) y la clínica que presenta el paciente para así establecer un pronóstico de la evolución de los síntomas a medio plazo.

1.1 Conceptos de dolor lumbar inespecífico y lumbalgia mecánica

El dolor lumbar inespecífico se define como un dolor más o menos intenso, tensión o rigidez en el área comprendida entre el borde inferior de la 12ª costilla y el pliegue glúteo (zona lumbar), de etiología desconocida y en cuya etiopatogenia podrían estar implicados los discos intervertebrales, las articulaciones, y el tejido conectivo (2). En la gran mayoría de los casos no puede determinarse un diagnóstico anatomopatológico específico que explique los síntomas de dolor. De los pacientes que consultan en atención primaria por dolor lumbar menos del 10% tienen una causa evidenciada, incluyendo un 4% que se presentan con fractura por compresión, el 3% con estenosis de canal, un 2% con patología visceral (incluyendo prostatitis, endometriosis, enfermedad pélvica inflamatoria, nefrolitiasis, pielonefritis, absceso perirrenal, aneurisma de aorta, pancreatitis, colecistitis e incluso úlcera gástrica perforada), el 0,7% tumor primario o metástasis y un 0,01% con una infección (3). Una revisión sistemática Cochrane confirmó una incidencia entre el 3,6 y el 6,5% de las fracturas como causa de dolor lumbar (4).

Introducción

Wadell propuso en 1987 una clasificación simple y práctica del dolor lumbar, ampliamente aceptada, dividiendo a los pacientes en tres categorías: patología espinal específica, dolor radicular y dolor lumbar inespecífico (5). En la actualidad prácticamente todas las Guías de Práctica Clínica sobre el manejo de dolor lumbar establecen el concepto de lumbalgia inespecífica como entidad clínica.

En líneas generales, consideraremos como lumbalgia inespecífica al dolor lumbar no atribuible a patología visceral, traumatismo claro, tumor subyacente, infección, síndrome de cola de caballo, radiculopatía o estenosis de canal. Incluiremos dentro de la definición aquellos casos de dolor lumbar que presenten degeneración discal, roturas anulares, artrosis facetaria o hernia discal sin clínica de afectación radicular así como espondilolisis y espondilolistesis, que con frecuencia podemos encontrar en pacientes asintomáticos (6–8).

La Red Española de Investigadores en Dolencias de la Espalda (www.reide.org) define el “síndrome mecánico lumbar” como aquellas afecciones que cursan con dolor lumbar que se modifica en función de los esfuerzos, posturas y movimientos, y se asocia habitualmente a una limitación dolorosa de la movilidad. Puede acompañarse de dolor referido o irradiado y no se debe a fracturas ni a enfermedades oncológicas, metabólicas, endocrinas, infecciosas, digestivas, ginecológicas, ni vasculares. El diagnóstico sindrómico incluye los cuadros dolorosos debidos a una causa orgánica, como hernia discal o estenosis espinal, los asociados a alteraciones de la estática y la dinámica de la columna vertebral, como escoliosis o espondilolistesis, y los debidos a causas inespecíficas o desconocidas, que representan, como ya se ha comentado, la mayoría de los casos.

Según la duración de los síntomas, suele considerarse como dolor lumbar agudo cuando este dura menos de 6 semanas, subagudo entre 6 semanas y 3 meses y crónico cuando supera esta duración. La guía de práctica clínica COST B13 Working Group on Guidelines for Chronic Low Back Pain define el dolor lumbar crónico como aquel con una duración superior a las 12 semanas (7). Después de un episodio de dolor

lumbar inespecífico se producirá la recidiva de los síntomas en dos tercios de los pacientes (8). Entre el 28 y el 65% de los pacientes que presentan un episodio agudo no consiguen una recuperación completa después de un año del inicio de los síntomas (9).

1.2 Epidemiología de la lumbalgia mecánica

La Carga Global de enfermedad atribuible a la lumbalgia a nivel mundial se ha estimado en un 46%; es la primera causa de años vividos con discapacidad (*Years Lived with Disability*) en 45 de 50 países desarrollados (90%) y en 94 de 138 países en desarrollo (68%) (10). Según los resultados del estudio EPISER (11) la prevalencia de lumbalgia aguda estimada en la población española adulta (mayor de 20 años) es del 14,8% y la prevalencia estimada de personas con lumbalgia crónica entre los adultos españoles es del 7,7%; la prevalencia se incrementa con la edad hasta los 60 años y es más frecuente en mujeres (17,8%) que en varones (11,3%). La Encuesta de Salud de la Comunidad Valenciana de 2010 determinó que el dolor lumbar es la tercera enfermedad crónica diagnosticada en la población adulta (15% en mujeres y 9% en hombres) (12). En España la lumbalgia constituye la causa más frecuente de baja laboral y la primera causa de discapacidad en menores de 45 años. El coste económico atribuible al dolor lumbar es de 8.946 millones de euros por año, representando el 0,68% del Producto Interior Bruto español (13). En nuestro país, el dolor lumbar es la primera causa de discapacidad ajustada a años de vida, por encima de la cardiopatía isquémica (14).

1.3 Aplicación de la Resonancia Magnética (RM) en el estudio del dolor lumbar crónico

La RM se ha convertido en la actualidad en la técnica de imagen más utilizada en el diagnóstico del dolor lumbar y que en muchas ocasiones sustituye a la radiografía simple como primera prueba de imagen a realizar. Sin embargo su uso generalizado no se ha visto reflejado en los resultados de salud en los pacientes con lumbalgia inespecífica (15). La imagen de RM convencional permite estudiar los distintos elementos que forman parte de la unidad funcional vertebral: el disco intervertebral,

Introducción

los cuerpos vertebrales, las articulaciones intervertebrales, así como la musculatura paravertebral y las estructuras neurales.

1.3.1 Degeneración del disco intervertebral

El núcleo pulposos ocupa sobre el 50% del volumen central del disco intervertebral. Está compuesto por condrocitos y una red de fibras de colágeno embebidos en una matriz compuesta por proteoglicanos, responsable de la hidratación del disco. Esta composición determina la alta intensidad que muestra en la RM en las secuencias potenciadas en T2. El anillo fibroso rodea toda la circunferencia del núcleo pulposos, formado por una densa red de fibras de colágeno organizadas de forma concéntrica; se muestra con baja intensidad en la RM. Las fibras más periféricas se denominan fibras de Sharpey y se anclan en los platillos vertebrales. El disco intervertebral normal esta inervado únicamente en las capas más periféricas del anillo fibroso, siendo aneurales el núcleo pulposos y las capas más internas del anillo fibroso.

La degeneración discal se caracteriza por una reducción de la producción de proteoglicanos con la consecuente deshidratación del disco, con un incremento de las fibras de colágeno. Esto se traduce en una disminución de la intensidad de señal en la imagen de RM en T2 y se acompaña de un colapso progresivo del disco (16). En el proceso de degeneración discal las terminaciones nerviosas crecen en número y penetran en el interior del disco intervertebral (17).

1.3.1.1 Gradación de la degeneración discal

Pfirrmann en 2001 (18) estableció un sistema de gradación de la degeneración del disco intervertebral cuando se analizan las imágenes de RM potenciadas en T2. Este sistema tiene en cuenta datos como la intensidad de señal, la estructura del disco, las diferencias entre el núcleo y el anillo fibroso y la altura discal. Considera cinco grados ordenados de menor a mayor gravedad:

- Grado I: Disco sano. Aparece homogéneo e hiperintenso en T2, de color blanco brillante. Altura normal del disco.

- Grado II: Disco hiperintenso, pero no homogéneo. Sigue manteniendo una altura normal y puede diferenciarse el núcleo del anillo fibroso.

- Grado III: Aparece una señal de intensidad intermedia, en un color grisáceo. No puede distinguirse correctamente el núcleo del anillo fibroso, y la altura del disco puede estar conservada o aparecer discretamente disminuida.

- Grado IV: El disco aparece hipointenso, de color gris oscuro y no homogéneo. Se ha perdido la diferencia entre núcleo y anillo fibroso, y la altura del disco puede ser normal o estar moderadamente disminuida.

- Grado V: La estructura del disco sigue siendo heterogénea, con una señal hipointensa que aparece negra en T2. No hay distinción entre núcleo y anillo fibroso, y el espacio del disco aparece colapsado.

Esta clasificación ha demostrado tener una buena reproducibilidad, tanto intra como inter-observador (19).

1.3.1.2 Zonas de Alta Intensidad

Otro hallazgo característico del proceso de degeneración discal son las denominadas zonas de alta intensidad (ZAI). Estas regiones hiperintensas en las imágenes potenciadas en T2 se localizan en el anillo fibroso, predominantemente en la zona posterior del disco intervertebral. Histológicamente se corresponden con un tejido de granulación hipervascularizado e inflamatorio, situado alrededor de una fisura central que permite su penetración por material nuclear (20). Se han establecido varios tipos de ZAI dependiendo de su localización (ZAI anterior o ZAI posterior), del tipo de señal que presentan las imágenes de RMN obtenidas con secuencias potenciadas en T1 (T1-isointenso, T1-hiperintenso y T1-hipointenso) y de su forma (redonda, cisura, vertical, oblicua y alargada) (21) Figura 1.

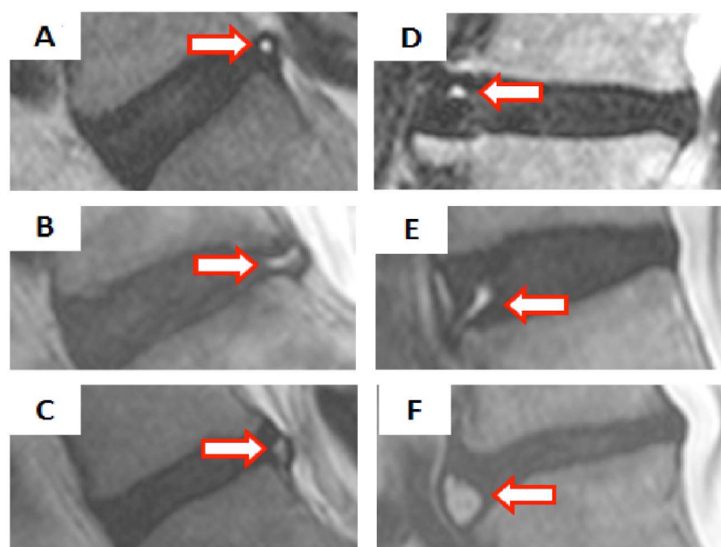


Figura 1. Tipos de Zonas de Alta Intensidad (ZAI). Forma redonda (A y D), cisura (B), vertical (C), oblicua (E) y alargada (F) (21).

1.3.1.3 Hernia discal

La degeneración del anillo fibroso, su pérdida de capacidad de contención y su fisuración condicionan con el tiempo la salida del núcleo pulposo y la formación de las hernias discales. Por consenso el abombamiento del disco que afecta a más del 25% de su circunferencia no se considera hernia discal, tratándose de abombamientos asimétricos adaptativos a las deformidades adyacentes (22). La hernia discal se define como la salida del material discal más allá del espacio intervertebral, focal, localizada en un punto de rotura del anillo fibroso y afecta a menos del 25% de su circunferencia en el plano axial. En general se consideran dos tipos de hernia: la hernia protrusa, con base amplia y con prominencia del núcleo pulposo sobre el canal, pero sin rotura del anillo fibroso, y la hernia extrusa, con una base estrecha con respecto al cuerpo más ancho, en la que el anillo fibroso interrumpido permite la salida del núcleo pulposo al canal. Las hernias extrusas puede evolucionar a secuestro discal cuando pierden el contacto con el resto del disco en el espacio intervertebral, y a hernia

migra cuando el material discal se desplaza craneal o caudalmente respecto al punto de extrusión (22).

1.3.2 Patología de los platillos vertebrales

Modic estableció en 1988 (23) una clasificación de las alteraciones óseas en los platillos vertebrales observadas en las imágenes de RM, que reflejan los cambios histológicos subyacentes. Se describen tres tipos de cambios en los platillos:

- Tipo I: edema óseo relacionado con la inflamación de la vértebra, observándose en las imágenes T1 una señal hipointensa y en las T2 una señal hiperintensa (Figura 2).
- Tipo II: reflejan la degeneración grasa, objetivándose en las imágenes en T1 y T2 como señal hiperintensa, que se anula con la supresión grasa (Figura 3).
- Tipo III: está relacionado con la esclerosis ósea trabecular, observándose de manera hipointensa en las imágenes potenciadas en T1 y T2.



Figura 2. Cambios Modic tipo I en imágenes de RM potenciadas en T1 (izquierda) y T2 (derecha).

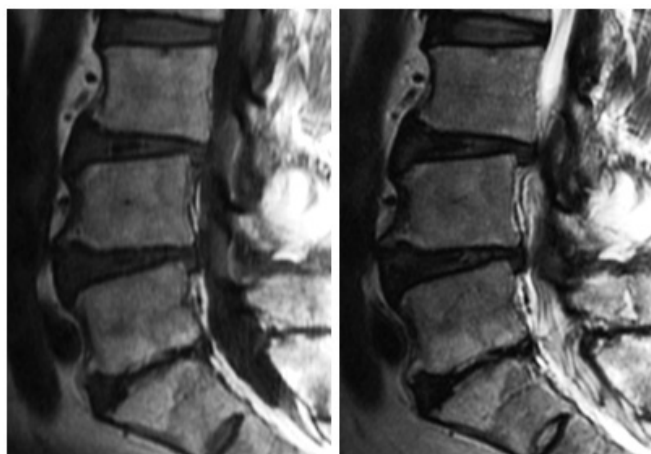


Figura 3. Cambios Modic tipo II en imágenes de RM potenciadas en T1 (izquierda) y T2 (derecha).

Se considera, como hemos visto, que los cambios de Modic tipo I representan cambios inflamatorios, los cambios tipo II manifiestan la conversión de la medula ósea hematopoyética normal a grasa medular ósea como consecuencia de isquemia y los cambios tipo III representan cambios de esclerosis ósea subcondral (24). La etiología y fisiopatología de estos cambios no ha sido completamente aclarada; podría tratarse de la asociación de factores mecánicos, microtraumatismos y diversos factores bioquímicos e isquémicos secundarios a un incremento de las tensiones soportadas por el hueso subcondral que originarían microfracturas.

Los cambios tipo I y II suelen localizarse en el tercio anterior del platillo vertebral y en los dos últimos espacios lumbares. Además, pueden coexistir los cambios I y II, así como los II y III. Los cambios tipo I pueden permanecer estables, resolverse o evolucionar a tipo II. El tipo II durante mucho tiempo se consideró estable, podría evolucionar a tipo I. El tipo III se trata de un estadio evolutivo terminal (25).

1.3.3 Patología de los elementos posteriores

Las imágenes de RM ponen de manifiesto algunos hallazgos localizados en los elementos vertebrales posteriores que podemos considerar patológicos en pacientes con dolor lumbar. Así, podemos

encontrar con frecuencia degeneración de las articulaciones facetarias con derrame articular, edema óseo, engrosamiento sinovial o quistes facetarios, espondilolistesis degenerativa y espondilólisis.

1.3.4 Estudio de la musculatura paravertebral

La degeneración de la musculatura retrosomática y su infiltración grasa secundaria también puede valorarse mediante RM y se ha intentado relacionar con el dolor lumbar. El método de valoración de la degeneración grasa de la musculatura paravertebral más utilizado es el visual subjetivo, evaluando los cortes axiales en los niveles estudiados, y clasificando la infiltración grasa como grado 0 (0-10% de infiltración), grado 1 (10-50% de infiltración), y grado 2 (>50% de infiltración) (26) (Figura 4). Diversos estudios han puesto de manifiesto la infiltración grasa de la musculatura paravertebral en pacientes con dolor lumbar agudo y crónico; sin embargo, no se ha podido establecer una clara relación causal, pudiendo la atrofia muscular deberse al desuso y por tanto ser una consecuencia del dolor lumbar (26).

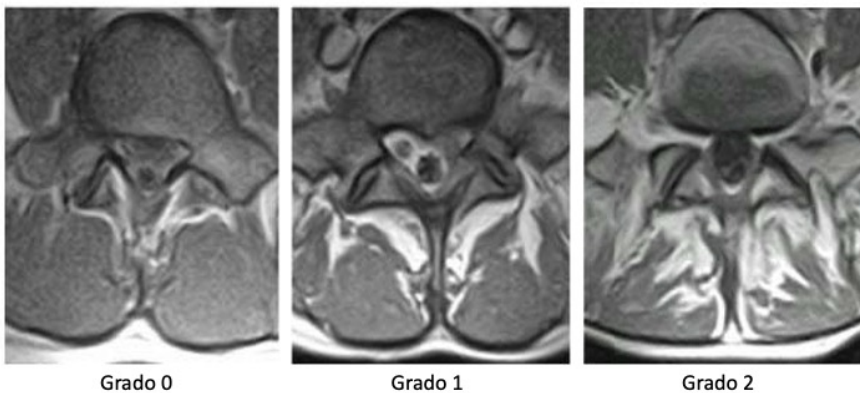


Figura 4. Grados de degeneración de la musculatura retrosomática (26) en imágenes de RM transversales potenciadas en T1.

1.3.5 Correlación clínico-radiológica

El desarrollo de las técnicas de imagen, en especial la RM como se ha comentado en los puntos anteriores, ha permitido obtener información detallada de las anomalías estructurales de la columna. Sin embargo, existe una gran dificultad para establecer una relación de causalidad entre el dolor lumbar y hallazgos como la degeneración discal, la artrosis facetaria, la protrusión/hernia discal o las alteraciones en la alineación del raquis. Estudios clásicos demuestran la existencia de abombamientos o protrusiones discales en más de 60% de individuos asintomáticos entre 25 y 50 años (27), con roturas del anillo fibroso en más del 50%, de los cuales más del 90% presentaban ZAI (28). Tonosu y cols (29), en un estudio longitudinal de 10 años de duración, observaron que no existía una asociación clara entre la degeneración discal severa (Pfirmann 4 o 5), abombamiento discal, ZAI, espondilolistesis o cambios de Modic y la aparición de dolor lumbar. En otro estudio longitudinal durante 3 años Suri y cols (30) observaron que la aparición de dolor lumbar no se relacionaba con la aparición de nuevos cambios en la RM respecto a los existentes al inicio del estudio. Sin embargo, se apreciaba una mayor incidencia de cambios en aquellos pacientes que se presentaban con síntomas radiculares. Un metaanálisis sobre estudios que comparaban sujetos sanos con pacientes con edad menor o igual a 50 años concluyó que diversos hallazgos en RM, tales como abombamiento discal, degeneración, extrusión, protrusión, cambios Modic tipo I o la espondilólisis, eran más frecuentes en individuos con dolor lumbar que en asintomáticos, aunque sin embargo no encontró relación del dolor lumbar con cualquier otro cambio de Modic, ZAI, estenosis central de canal, fisuras anulares o espondilolistesis (31). Un estudio reciente sobre más de 700 pacientes menores de 30 años con dolor lumbar mostró una prevalencia del 59% de la existencia en la RM de degeneración discal grado de Pfirmann 3 o superior (32). En un estudio transversal en 52 pacientes con dolor lumbar se observó que la presencia de cambios degenerativos, que se observaron con mayor frecuencia en L4-L5, no se relacionaron con la intensidad del dolor ni demostraron valor predictivo para la presencia de dolor neuropático, sin embargo el sexo femenino y la intensidad del dolor si

demonstraron valor predictivo para la presencia de dolor neuropático (33). En otro estudio prospectivo con un seguimiento de 30 años se observó que los cambios degenerativos a los 20 años de edad predisponen a un deterioro progresivo de los discos intervertebrales, pero no se relacionan con el dolor o la discapacidad (34). Rahyussalim y cols en una revisión sistemática reciente observan que hasta el 60% de la población general sin dolor lumbar puede presentar signos degenerativos discales y este porcentaje puede superar el 80 - 90% en mayores de 55 años asintomáticos (35). Dado que la prevalencia de hallazgos en la RM de pacientes con dolor lumbar es elevada y la correlación con los hallazgos clínicos muy controvertida es necesario seguir investigando en este campo.

1.4 Análisis cuantitativos mediante RM

Los estudios de imagen por RM convencional permiten obtener resultados cualitativos o semicuantitativos permitiendo las clasificaciones o gradaciones que se han comentado en los apartados anteriores. Todos estos parámetros y las distintas escalas de valoración basadas en la RM (como por ejemplo la escala de degeneración discal de Pfirrmann) tienen un valor limitado como biomarcadores para la detección de la enfermedad en estadios iniciales, así como escaso valor pronóstico. Además, estas valoraciones cualitativas siempre son subjetivas y están expuestas a una mayor variabilidad inter-observador.

En la actualidad un método de cuantificación es el cálculo de la intensidad de señal a nivel del disco en el corte sagital medio en las secuencias potenciadas en T2 (36, 37). Esta señal refleja las propiedades bioquímicas y se relaciona con el contenido en proteoglicanos del disco. La reducción de la señal del disco se relaciona con el grado de degeneración del mismo. Técnicas de RM más avanzadas (tales como T2*, T1rho, transferencia de magnetización, *chemical exchange saturation transfer* CEST mapping, MRI sintética MAGiC) permiten un estudio tanto desde el del punto de vista bioquímico como estructural (38-41). Sin embargo, estos procedimientos requieren unos protocolos de adquisición de imágenes específicos, incrementan la duración de las exploraciones, generan una información sin una validación amplia, y aumenta los costes

respecto a la RM convencional, dificultando su aplicación en la práctica clínica habitual. Por lo que sabemos en la actualidad no hay estudios que relacionen estos hallazgos con la clínica del paciente y su pronóstico.

1.5 *Pronóstico del dolor lumbar inespecífico*

Se han descrito diversas herramientas que integran variables clínicas, con la finalidad de identificar aquellos pacientes con riesgo de mala evolución cuando padecen un proceso de dolor lumbar. Karan *y cols* (42) evaluaron en un estudio prospectivo la capacidad predictiva de tres cuestionarios denominados *Örebro Musculoskeletal Pain Screening Questionnaire* (ÖMPSQ), *Predicting the Inception of Chronic Pain* (PICKUP) y STarT-Back. Los tres cuestionarios consiguieron un AUC cercana al 0.70, aunque todas ellas infravaloraron el riesgo de mal pronóstico (42). Los mismos autores realizaron una revisión sistemática concluyendo que la capacidad de estos métodos de cribado para clasificar correctamente a un paciente en alto o bajo riesgo de sufrir dolor crónico estaría entre el 60 y el 70%, y de sufrir discapacidad entre el 70 y el 80% (43). Curiosamente, un estudio prospectivo posterior demostró que el ÖMPSQ y el STarT-Back no coincidían a la hora de determinar el riesgo de cronicidad en los mismos individuos (44). Estos cuestionarios recogen síntomas clínicos y aspectos psicosociales para establecer el cálculo predictivo sobre la cronificación del dolor lumbar, sin considerar los hallazgos de imagen o las alteraciones anatomopatológicas.

En la actualidad existe una gran incertidumbre sobre la capacidad de las alteraciones estructurales observadas en las pruebas de imagen para identificar aquellos pacientes que tienden a cronificarse. La búsqueda de marcadores pronósticos de resolución, permanencia o agravamiento del dolor lumbar se ha constituido en una de las prioridades de la investigación en la patología de la columna vertebral.

La RM, a pesar de ser claramente la técnica de imagen de elección en la patología de la columna vertebral, presenta una correlación baja con la clínica y, por lo tanto, aporta un valor pronóstico limitado (45). De Shepper *y cols* apreciaron un aumento pequeño (el AUC de 77 pasó 78) en

la capacidad de predecir la mala evolución del dolor lumbar al añadir los hallazgos de RM a la valoración clínica del paciente (46).

1.6 Radiómica: análisis de textura

La radiómica es el proceso de extracción de una gran cantidad de datos cuantificables, objetivos y no visibles al ojo humano que están contenidos en las imágenes obtenidas por las técnicas de imagen, incluyendo la TAC y la RM. Mediante este proceso la imagen radiológica se transforma en un amplio conjunto de parámetros que, mediante el uso de sofisticadas herramientas computacionales, permiten desarrollar modelos que mejoraran la capacidad diagnóstica, valor pronóstico y predictivo de la imagen en determinadas patologías. El análisis de estos parámetros junto con los hallazgos clínicos, resultados de laboratorio, análisis genéticos, y otras pruebas diagnósticas aportarían información de gran valor a la hora de establecer un plan terapéutico en un paciente determinado. La radiómica constituye una fuente potencial de biomarcadores que ya han demostrado su utilidad en el diagnóstico y fenotipado del cáncer, así como en su pronóstico, predicción de respuesta al tratamiento y monitorización de la evolución de la enfermedad en un paciente determinado (47, 48).

1.7 Parámetros de textura (image features)

A partir de la imagen radiológica pueden obtenerse distintos tipos de información clínicamente relevantes. Por una parte, tenemos los parámetros subjetivos cualitativos, utilizados habitualmente por los radiólogos para describir las lesiones. Por otra parte, los parámetros objetivos cuantitativos que se obtienen a partir de la imagen mediante la aplicación de algoritmos matemáticos computacionales. Estos parámetros cuantitativos se clasifican en distintas categorías (49, 50):

- Parámetros de forma: se trata de la descripción del área de interés (ROI, por *Region of Interest* en inglés), sus propiedades geométricas como el volumen, diámetros en diferentes direcciones, máxima superficie o esfericidad.

Introducción

- Parámetros estadísticos de primer orden: se basan en el estudio de la intensidad de señal individual de los vóxeles sin tener en cuenta su distribución espacial. Estos parámetros se basan en el estudio del histograma y con él se obtienen distintos valores una ROI como la media, mediana, máximo, grado de uniformidad de la intensidad de la imagen (*entropía*), así como la asimetría (*skewness*) o la morfología más o menos plana del histograma (*curtosis*).
- Parámetros estadísticos de segundo orden: también conocidos como parámetros de textura propiamente dichos, se obtienen mediante cálculos estadísticos de las relaciones de los vóxeles de la imagen entre sí. Estos valores nos proporcionan información de la distribución espacial de los vóxeles de distintas intensidades de gris y el grado de heterogeneidad espacial de la zona estudiada. Estos parámetros pueden obtenerse a partir de la “*grey-level co-occurrence matrix*” (GLCM), que cuantifica la incidencia de vóxeles con la misma intensidad a una distancia dada a lo largo de una dirección fija. Algunos de los estadísticos que podemos obtener son autocorrelación, contraste, correlación, *cluster prominence*, *cluster shade*, *cluster tendency*, disimilitud, energía, entropía, homogeneidad, probabilidad máxima, suma de cuadrados, suma de promedios, suma de varianzas, y suma de entropía, entre otros.
- Parámetros estadísticos de segundo orden: obtenidos a partir de la “*grey-level run-length matrix*” (GLRLM), cuantifica los vóxeles consecutivos con la misma intensidad a lo largo de una dirección fija. Los parámetros que se pueden obtener son: *short run emphasis*, *long run emphasis*, *gray-level nonuniformity*, *run-length nonuniformity*, *run percentage*, *low gray-level run emphasis* y *high gray-level run emphasis*.
- Otras matrices propuestas para el análisis de textura son la “*Gray Level Size Zone*” (GLSZM), la “*Neighboring Gray Tone Difference Matrix*” (NGTDM) y la “*Gray Level Dependence Matrix*” (GLDM).

Puede obtenerse una gran variedad de valores de nivel superior aplicando diversos métodos matemáticos; muchos de estos valores resultaran ser redundantes por lo que posteriormente deberán procesarse los resultados para elegir aquellos parámetros que sean más adecuados.

1.8 Proceso del análisis radiómico.

El análisis radiómico se realiza siguiendo una serie de etapas metodológicamente bien definidas, cada una de cuales presenta unas características y dificultades específicas, y que se desarrollan a continuación (49, 51).

1.8.1 Obtención de las imágenes

La primera fase del proceso consiste en la adquisición de las imágenes de la columna vertebral del paciente, normalmente mediante una exploración de TC o RM. Un aspecto importante a tener en cuenta en esta primera fase es la variabilidad existente en los parámetros de adquisición de estas imágenes. Esta variabilidad sería de menor importancia si se va a hacer un análisis sólo cualitativo de las alteraciones, pero a la hora de transformar la imagen en datos condicionará los resultados extraídos del análisis de textura. Una forma de evitar esta limitación sería excluir aquellos valores que se ven más influenciados por la adquisición de las imágenes. Por ejemplo, dado que la intensidad de señal en las imágenes de RM está condicionada por diversos factores como son las características del campo magnético del equipo, los parámetros de adquisición y el procesado de la imagen, estos deben presentar valores similares. Sin embargo, la intensidad de señal puede cambiar repitiendo la exploración exactamente igual y con el mismo equipo al mismo individuo en dos momentos distintos. Para corregir esta variabilidad puede cuantificarse la relación entre intensidades de vóxeles entre sí, que no depende de la intensidad individual; otra opción es realizar una normalización de la señal antes de proceder con el análisis.

1.8.2 Segmentación de las imágenes

La segmentación de las imágenes consiste en delimitar la región de interés (ROI) como máscara que posteriormente vamos a analizar. Se trata de una fase fundamental en el procesado de imágenes ya que todos los datos posteriores se van a generar a partir del volumen concreto que seleccionemos en el estudio analizado. Podríamos decir que es el equivalente a la toma de una biopsia para el estudio anatomopatológico de una lesión, donde la correcta localización de la obtención de la muestra influirá en el resultado obtenido. La segmentación puede realizarse de forma manual o automática. El hecho de que las lesiones suelen mostrar formas irregulares genera un problema o sesgo de reproducibilidad en la segmentación manual. Además, los métodos manuales suponen un considerable esfuerzo de tiempo porque las imágenes RM y TC contienen un elevado número de imágenes por estudio. Probablemente, la mejor opción en este proceso es la segmentación automática y posteriormente corregida manualmente, aunque todavía diversos autores consideran que la segmentación manual es la más veraz a pesar de la alta variabilidad inter-observador que presenta (50). Nuestra postura y experiencia favorece la segmentación automática con edición y validación posterior por una persona experimentada.

1.8.3 Extracción de los parámetros de textura

Una vez realizada la segmentación, se aplican herramientas informáticas desarrolladas para el cálculo de los parámetros de textura a los que se ha hecho referencia en el apartado 1.7. Algunos de los paquetes informáticos más utilizados en el análisis de textura son:

- maZda. Programa informático que permite realizar todo el proceso de análisis de textura a partir de imágenes en los formatos más habituales como DICOM, permite el cálculo de parámetros basados en el histograma, gradiente, GLCM y GLRLM (52).

- PyRadiomics. Se trata de un paquete de código abierto para la extracción de características radiómicas de las imágenes médicas. Este paquete tiene como objetivos establecer un estándar de referencia para el análisis radiómico y proporcionar una plataforma de código abierto probada y mantenida para una extracción fácil y reproducible de las características radiómicas. Permite el cálculo de los parámetros referenciados en el apartado 1.7, pudiendo obtener más de 1.000 valores distintos, la mayor parte de ellos según las definiciones establecidas por la *Imaging Biomarker Standardization Initiative* (IBSI) (53).

1.8.4 Selección de parámetros de interés

Muchos de los parámetros obtenidos en el análisis radiómico son redundantes, por lo que debemos realizar una selección con cribado tras identificar aquellos que potencialmente pueden ser útiles para obtener un modelo que pueda tener una aplicación clínica en el proceso que estamos analizando. Dos de los procedimientos más utilizados a este fin son el análisis de agrupaciones (*clusters*) y el análisis de componentes principales (*PCA*). El análisis de agrupaciones consiste en crear grupos de parámetros de características similares, con una alta redundancia intra-agrupación y baja correlación inter-agrupación, para seleccionar un parámetro representativo de la agrupación que será el empleado para el análisis posterior. El método PCA consiste en crear un grupo pequeño de variables no correlacionadas a partir de un gran conjunto de variables correlacionadas de forma que representemos la variedad del conjunto de datos con el menor número posible de variables.

A partir de estas variables consideradas reproducibles, informativas y no redundantes, podemos llevar a cabo el análisis de asociación (50).

1.8.5 Modelo predictivo

Mediante un análisis estadístico de las variables obtenidas puede elaborarse un modelo predictor que estime las probabilidades de que suceda un evento dado. Para establecer la correlación de los datos

Introducción

computacionales con datos clínicos (diagnóstico de fenotipo, efecto del tratamiento o resultado clínico final) pueden emplearse tanto descriptores estadísticos clásicos, como análisis multivariante y agrupaciones nosológicas. También pueden relacionarse mediante modelos de agrupamiento supervisado empleando Inteligencia Artificial, siendo el *Random Forest* uno de los métodos con mejor rendimiento desde el punto de vista pronóstico (48, 54).

La figura 5 resume de forma esquemática todo el proceso del análisis radiómico (48).

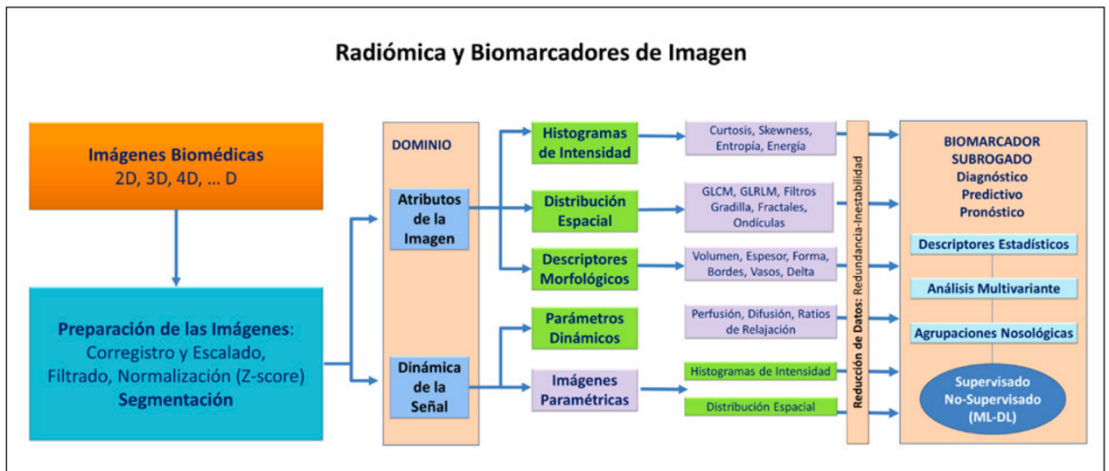


Figura 5. Proceso del análisis radiómico (48).

1.9 Aplicaciones clínicas del análisis de textura

En los últimos años se han publicado diversos trabajos, especialmente en imagen de tumores, imagen del sistema nervioso e imagen cardiaca, en los que se evalúa el papel del análisis de textura y el valor de determinados parámetros como biomarcadores con valor diagnóstico, predictivo y pronóstico. Así, algunas aplicaciones clínicas del análisis radiómico pueden ser:

- Estudio de tumores para mejorar el diagnóstico, estadiaje, respuesta al tratamiento y pronóstico: cáncer de mama (55-57), glioma (58-60), hepatocarcinoma (61-63), sarcomas de partes blandas (64, 65) y cáncer de páncreas (66, 67), entre otros.
- Estudio de patologías neurodegenerativas como la enfermedad de Alzheimer (68).
- Estudio de disfunción orgánica como en imagen cardiaca (69).
- Estudio de patología del aparato locomotor mediante estudios cuantitativos de la estructura ósea en pacientes con osteoporosis (70) o estudios del hueso subcondral en el extremo proximal de la tibia en pacientes con artrosis (71). En tumores del aparato locomotor la radiómica ha demostrado capacidad de diferenciar entre condrosarcoma de bajo grado y encondroma (72).

1.10 Aplicación del análisis de textura en el estudio de la columna vertebral

Algunos estudios han analizado la posibilidad de aplicar el análisis de textura a la patología del disco intervertebral sobre las imágenes médicas convencionales de RM. El análisis de textura ha resultado ser más sensible que la imagen en secuencias potenciadas en T2 para detectar un disco patológico (73). La pérdida de la homogeneidad del núcleo pulposo muestra una gran correlación con la degeneración discal, en especial en las fases más iniciales de este proceso. Además el análisis de la textura se postula como una herramienta para monitorizar la evolución de dicho proceso involutivo o la respuesta al tratamiento y que, por su sencillez, podría incorporarse a la práctica clínica habitual (74). Huber y cols demostraron mayor reproductibilidad del análisis de textura frente a los métodos cualitativos para establecer el grado de afectación por RM en pacientes con estenosis de canal lumbar (75). Abdollah y cols encontraron diferencias significativas entre pacientes con dolor lumbar inespecífico y controles en parámetros como contraste, homogeneidad y energía a nivel del platillo inferior de L5 así como el superior de S1; también observaron diferencias significativas en contraste y homogeneidad a nivel del disco

Introducción

L5-S1 (76). Ketola y cols demostraron que el análisis de textura de los discos intervertebrales mejoraba la capacidad de diagnosticar por RM a pacientes con dolor lumbar inespecífico frente a la clasificación de Pfirmann o los cambios de Modic, en especial cuando se aplicaba a los espacios L4-L5 y L5-S1 (77). También se ha aplicado el análisis de textura para el estudio de la musculatura paraespinal, permitiendo cuantificar su nivel de infiltración grasa (78) e incluso relacionar la textura con la fuerza muscular (79).

Aunque los estudios analizados ponen de manifiesto la capacidad del análisis de textura para identificar un disco patológico frente a uno sano y una mayor capacidad de identificar a individuos con dolor lumbar frente al análisis convencional de la RM, no se han encontrado estudios que analicen su capacidad para correlacionarse con el grado de dolor, la discapacidad y el pronóstico de los pacientes a medio plazo.

2 Hipótesis

El análisis cuantitativo de la textura del disco intervertebral, de los platillos vertebrales adyacentes y de la musculatura paravertebral lumbar en las imágenes de RM de la columna vertebral lumbar está directamente relacionado con los valores proporcionados por las escalas de dolor y discapacidad en pacientes con lumbalgia mecánica crónica, permitiendo determinar el pronóstico a los 6 meses de estos pacientes.

3 Objetivos

El diseño de esta Tesis Doctoral se establece para dar respuesta al desarrollo principal y a los desglosados como secundarios. En concreto, se establecieron los siguientes objetivos:

3.1 Objetivo principal

Evaluar la utilidad del análisis radiómico sobre las imágenes de RM de los componentes principales de columna lumbar en la evaluación clínica de los pacientes con dolor lumbar inespecífico.

3.2 Objetivos secundarios

- Establecer la correlación entre los distintos parámetros de textura obtenidos a partir de imágenes de RM con los valores de las escalas de dolor y de discapacidad en pacientes con dolor lumbar crónico.
- Determinar si estos parámetros de textura tienen un valor pronóstico y si permiten prever cual será la respuesta de los pacientes tras 6 meses de tratamiento conservador.
- Evaluar la correlación existente entre los criterios subjetivos grado de degeneración discal, cambios de Modic, presencia de ZAI y degeneración de la musculatura paravertebral, con la intensidad del dolor y la discapacidad en pacientes con dolor lumbar crónico.
- Investigar si estos parámetros de grado de degeneración discal, cambios de Modic, presencia de ZAI y degeneración de la musculatura paravertebral tienen un valor pronóstico y si permiten prever cual será la respuesta de los pacientes tras 6 meses de tratamiento conservador.
- Comparar las diferencias entre el análisis radiómico y la lectura radiológica subjetiva con respecto a las correlaciones con las escalas de dolor, escalas de discapacidad y el pronóstico a los 6 meses tras tratamiento conservador.
- Establecer una metodología de evaluación de pacientes con dolor lumbar crónico con impacto en la práctica clínica habitual.

4 Material y métodos

Material y métodos

4.1 *Diseño del estudio*

Se diseñó un estudio prospectivo, observacional no intervencionista en el que se invitó a participar a pacientes consecutivos, diagnosticados de lumbalgia mecánica crónica, y reclutados en las consultas de los servicios de Rehabilitación y de Cirugía Ortopédica y Traumatología de un área sanitaria de Valencia que atiende a una población de 320.000 habitantes. En el diseño y redacción de esta Tesis Doctoral se siguieron las recomendaciones y lista de verificación de la declaración TRIPOD (80).

4.2 *Criterios de inclusión y exclusión*

Seguendo los criterios diagnósticos de la Guía de Práctica Clínica COST-B-13 (7) se establecieron los siguientes **criterios de inclusión**:

- 1) Edad entre 18 y 65 años.
- 2) Lumbalgia mecánica inespecífica de 6 meses como mínimo de duración.
- 3) Indicación de una exploración con RM de la columna lumbar.
- 4) Recibir tratamiento rehabilitador durante 6 meses y acudir a revisión clínica posterior.
- 5) Firmar el consentimiento informado

Como ***criterios de exclusión*** se consideraron los siguientes:

- 1) Dolor de origen específico etiquetado como fractura vertebral, neoplasia en columna o estructuras nerviosas, espondiloartropatía, o espondilodiscitis.
- 2) Haber recibido tratamiento rehabilitador previamente a la visita inicial.
- 3) Cirugía previa de raquis.
- 4) Retraso mental, enfermedad mental grave, abuso o dependencia de sustancias y analfabetismo que impidan la cumplimentación de cuestionarios de discapacidad.

4.3 *Tamaño muestral*

Para el cálculo del tamaño muestral se utilizó el software G-Power versión 3.1. Considerando como clínicamente relevante un valor de R² de 0,15 y para un estudio de correlación, el cálculo a priori para obtener un tamaño del efecto de 0,387 con un error alfa de 0,05 y una potencia (1-error beta) de 0,9 el tamaño muestral necesario es de 62 pacientes.

Se estableció un tamaño muestral de 80 pacientes para compensar eventuales pérdidas de seguimiento.

4.4 *Selección de la muestra*

Se invitó a participar de forma consecutiva a todos los pacientes que acudieron a las consultas externas de Rehabilitación y de Cirugía Ortopédica y Traumatología del mismo Departamento de Salud de la Comunitat Valenciana por dolor lumbar, y que cumplieran los criterios de inclusión, sin ninguno de los criterios de exclusión, dentro de la actividad normal de estas unidades clínicas. Tras la adecuada información sobre el estudio, aquellos individuos que aceptaron participar firmaron el documento de consentimiento informado (Anexo 1).

4.5 *Variables a estudio*

Se consideraron las siguientes variables clínicas y de imagen:

4.5.1 *Variables clínicas*

4.5.1.1 *Datos epidemiológicos*

Se recogieron las variables demográficas, edad y sexo, y las antropométricas básicas peso y altura, y se calculó el Índice de Masa Corporal (IMC) en Kg/cm².

4.5.1.2 *Intensidad del dolor*

Para cuantificar la intensidad del dolor que presentaba el paciente se utilizó la Escala Numérica del Dolor (END).

La END, introducida por Downie en 1978 (81), es una de las escalas más empleadas, ofrece alta sensibilidad y genera datos que pueden ser estadísticamente analizados (82). El paciente debe asignar al dolor que sufre en ese momento un valor numérico entre dos puntos extremos (0 a 10). Aunque al sujeto se le pide que utilice valores numéricos para indicar el nivel de su dolor, la utilización de palabras claves, así como unas instrucciones previas (0 es no tener dolor y 10 es el máximo dolor imaginable), favorecen que el paciente conceptualice su dolor en términos numéricos. Con este tipo de escala el dolor se mide sólo según su intensidad y es una escala discreta, no continua. Es útil como instrumento de medida para valorar la respuesta a un tratamiento seleccionado (83). Podemos considerar dolor leve si el paciente puntúa el dolor como menor de 3, dolor moderado si la valoración se sitúa entre 4 y 5, o dolor severo si la valoración es igual o superior a 6 (84). Se considera que el mínimo cambio clínicamente relevante sería del 30 % en esta puntuación (85).

Se evaluó el dolor que presentaban los pacientes en el momento del contacto inicial y en una visita de seguimiento a los 6 meses.

4.5.1.3 Grado de discapacidad

Para cuantificar la discapacidad que el dolor lumbar le genera al paciente se utilizó la versión validada al español del Rolland-Morris Disability Questionnaire (RMQ) (86) (Anexo 2).

La escala de RMQ valora la discapacidad funcional que presenta el paciente con dolor lumbar. Se publicó por primera vez en el año 1983 por M. Roland y R. Morris (87), realizándose una revisión en el año 2000 (88). Consiste en un cuestionario de 24 preguntas relacionadas con la vida diaria, valorando como afecta el dolor lumbar a su actividad normal, sumando 1 punto por cada respuesta afirmativa del paciente y 0 por cada respuesta negativa. Los resultados van desde 0 a 24, siendo 0 nula discapacidad y 24 la máxima discapacidad posible. Un resultado de 4 o menor se considera discapacidad leve sin repercusión sobre la calidad de vida del paciente (89).

Material y métodos

Se evaluó la discapacidad que presentaban los pacientes en el contacto inicial y en una visita de seguimiento a los 6 meses.

4.5.2 Variables basadas en la RM lumbar

Se analizaron múltiples variables basadas en la RM convencional para su posterior correlación con las variables clínicas.

4.5.2.1 Variables cualitativas

4.5.2.1.1 Degeneración discal

- Se identificó el grado de degeneración de los discos intervertebrales según la escala de Pfirrmann (18) para cada uno de los discos de manera independiente desde L1-L2 hasta L5-S1. Para esta tarea clasificadora se usaron los tres cortes sagitales centrales en la secuencia T2.
- En cada uno de los discos lumbares se identificó la presencia de Zonas de Alta Intensidad (ZAI) (21) en la parte posterior del anillo fibroso analizando todos los cortes sagitales T2 disponibles en cada estudio.

4.5.2.1.2 Patología de los platillos vertebrales

- Se investigó la eventual presencia de cambios de Modic y tipo (23) en cada uno de los niveles discales de la columna lumbosacra. La presencia y clasificación se obtuvo mediante el análisis de los cortes sagitales potenciados tanto en T1 como en T2.

4.5.2.1.3 Degeneración muscular

- Se determinó el grado de degeneración grasa, en una escala de 0 a 2, según los criterios establecidos por Hildebrandt M y cols (26). Para esta tarea se analizó la musculatura paravertebral posterior en los cortes transversales potenciados en T2 a nivel del disco L4-L5.

4.5.2.2 Variables de textura

Las variables de textura se calcularon a partir de la segmentación manual de diversas ROI en las imágenes de RM convencional: disco intervertebral y platillos vertebrales superior e inferior en los niveles L3-L4 a L5-S1 y musculatura retrosomática a nivel L4-L5. Para cada ROI se calcularon los parámetros básicos incluidos en la biblioteca de la aplicación PyRadiomics (53).

Estos parámetros de textura de las áreas segmentadas se dividen en varios grupos que se describen a continuación y se enumeran en la tabla 1:

4.5.2.2.1 Parámetros de primer orden

Los estadísticos de primer orden, basados en el histograma de la escala de grises, describen la distribución de las intensidades de gris de los vóxeles dentro de las distintas ROI. En este grupo se incluyen 18 parámetros.

4.5.2.2.2 Parámetros de forma

En este grupo se incluyen 14 parámetros que describen la forma y tamaño tridimensional de cada ROI y son independientes del nivel de intensidad en la escala de grises.

4.5.2.2.3 Gray Level Co-occurrence Matrix (GLCM)

En este grupo se incluyen 24 parámetros que describen la frecuencia con la que un nivel de gris aparece en una relación espacial específica con otro valor de gris, dentro de una ROI.

4.5.2.2.4 Gray Level Size Zone Matrix (GLSZM)

En este grupo se incluyen 16 parámetros que cuantifican las zonas con determinada intensidad de gris; estas zonas se definen por el número de vóxeles conectados que presentan la misma intensidad en la escala de grises.

4.5.2.2.5 Gray Level Run Length Matrix (GLRLM)

En este grupo se incluyen 16 parámetros que cuantifican los recorridos del nivel de gris en una imagen. Un recorrido de nivel de gris se define como la longitud, en número de píxeles, de píxeles consecutivos que tiene el mismo valor de nivel de gris.

4.5.2.2.6 Neighboring Gray Tone Difference Matrix (NGTDM)

En este grupo se incluyen 5 parámetros y cuantifican la diferencia entre la intensidad de gris de un vóxel y el promedio de los vóxeles de su alrededor.

4.5.2.2.7 Gray Level Dependence Matrix (GLDM)

En este grupo se incluyen 14 parámetros que cuantifican las dependencias de los niveles de gris en la imagen. La dependencia se define por el número de vóxeles conectados que son dependientes de un vóxel central.

Tabla 1. Parámetros de textura básicos incluidos en la biblioteca PyRadiomics (53).

Parámetros de primer orden	
1	Energía
2	Energía total
3	Entropía
4	Mínimo
5	Percentil 10
6	Percentil 90
7	Máximo
8	Media
9	Mediana
10	Rango intercuartil
11	Rango
12	Desviación media absoluta de la media
13	Desviación media absoluta de la media robusta
14	Raíz cuadrada de la media
15	Skewness
16	Kurtosis

17	Varianza
18	Uniformidad
Parámetros de forma 3D	
1	Mesh volume
2	Voxel volume
3	Surface area
4	Surface área to volumen ratio
5	Esfericidad
6	Diámetro 3D máximo
7	Diámetro 2D máximo (corte)
8	Diámetro 2D máximo (columna)
9	Diámetro 2D máximo (fila)
10	Longitud eje mayor
11	Longitud eje menor
12	Longitud eje mínimo
13	Elongación
14	Flatness
Gray Level Co-occurrence Matrix (GLCM)	
1	Autocorrelation
2	Join average
3	Cluster prominence
4	Cluster shade
5	Cluster tendency
6	Contrast
7	Correlation
8	Difference Average
9	Difference entropy
10	Difference variance
11	Joint Energy
12	Joint Entropy
13	Informational Measure of correlation 1
14	Informational Measure of correlation 2
15	Inverse difference moment
16	Maximal correlation coefficient
17	Inverse difference moment normalized
18	Inverse difference
19	Inverse difference normalized
20	Inverse variance
21	Maximum probability

22	Sum average
23	Sum entropy
24	Sum of squares
Gray Level Size Zone Matrix (GLSZM)	
1	Small area emphasis
2	Large area emphasis
3	Gray level non-uniformity
4	Gray level non-uniformity normalized
5	Site-zone non-uniformity
6	Site-zone non-uniformity normalized
7	Zone percentage
8	Gray level variance
9	Zone variance
10	Zone entropy
11	Low gray level zone emphasis
12	High gray level zone emphasis
13	Small area low gray level zone emphasis
14	Small area high gray level zone emphasis
15	Large area low gray level zone emphasis
16	Large area high gray level zone emphasis
Gray Level Run Length Matrix (GLRLM)	
1	Short run emphasis
2	Long run emphasis
3	Gray level non-uniformity
4	Gray level non-uniformity normalized
5	Run length non-uniformity
6	Run length non-uniformity normalized
7	Run percentage
8	Gray level variance
9	Run variance
10	Run entropy
11	Low gray level run emphasis
12	High gray level run emphasis
13	Short run low gray level emphasis
14	Short run high gray level emphasis
15	Long run low gray level emphasis
16	High run low gray level emphasis
Neighboring Gray Tone Difference Matrix (NGTDM)	
1	Coarseness

2	Contrast
3	Busyness
4	Complexity
5	Strength
Gray Level Dependence Matrix (GLDM)	
1	Small dependence matrix
2	Large dependence matrix
3	Gray level non-uniformity
4	Dependence non-uniformity
5	Dependence non-uniformity normalized
6	Gray level variance
7	Dependence variance
8	Dependence entropy
9	Low gray level emphasis
10	High gray level emphasis
11	Small dependence low gray level emphasis
12	Small dependence high gray level emphasis
13	Large dependence low gray level emphasis
14	Large dependence high gray level emphasis

4.6 *Procedimiento*

El protocolo de estudio que se llevó a cabo en todos los pacientes fue el siguiente:

4.6.1 *Evaluación clínica*

En la primera visita en consultas externas se recogieron los datos epidemiológicos del paciente, se estableció el grado de dolor lumbar según la END y se entregó al paciente el cuestionario Rolland-Morris para cumplimentarlo en ese momento.

En todos los pacientes se indicó seguir con un tratamiento rehabilitador, según el protocolo de tratamiento habitual del centro donde se desarrolló el estudio.

Se realizó una visita de seguimiento al paciente a los 6 meses de la primera; en esta visita se determinó de nuevo el grado de dolor según la

Material y métodos

END y se entregó el cuestionario Rolland-Morris para su cumplimentación por el paciente en ese momento.

4.6.2 Estudio de imagen: RM lumbar

En la primera visita se solicitó una RM lumbar, siguiendo el protocolo habitual del centro en el estudio del dolor lumbar crónico. Esta exploración se llevó a cabo con diversos equipos: en 70 casos la RM se realizó con un equipo Siemens Symphony 1,5T, en 2 casos Siemens Avanto 1,5T, en 8 casos Philips Achieva 3T, en 11 Philips Panorama HFO 1T y en 3 GE Signa Excite 1,5T. El estudio siempre incluyó, entre otras imágenes, las secuencias axiales potenciadas en T2 y las secuencias sagitales potenciadas en T1 y T2.

4.6.3 Análisis de la imagen de RM

El análisis realizado sobre las imágenes de RM convencional se resume en la Figura 6. Las imágenes de RM de los pacientes fueron importadas del PACS en formato DICOM y anonimizadas para su estudio analítico.

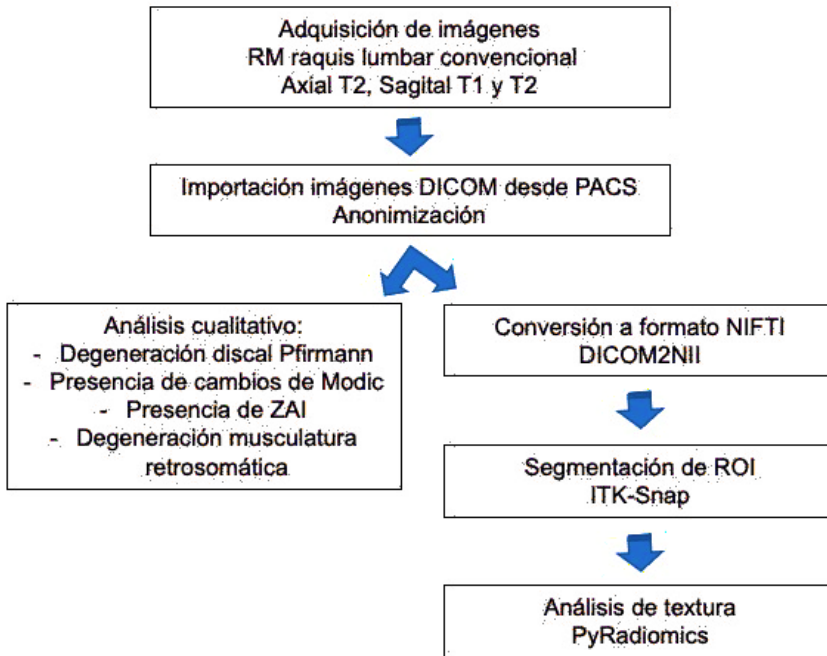


Figura 6. Proceso y análisis de las imágenes de RM realizado.

4.6.3.1 Parámetros cualitativos

A partir de las imágenes de RM se obtuvieron los parámetros cualitativos enunciados en el apartado 4.5.2.1. sobre degeneración discal (escala de Pfirmann, ZAI), platillos vertebrales (cambios de Modic y tipo) y degeneración muscular.

4.6.3.2 Análisis de textura

El análisis de textura se realizó siguiendo el esquema de trabajo descrito en el apartado 1.9 y que de forma detallada se desarrolla a continuación.

Partimos de las imágenes de RM de la columna lumbar obtenidas en el estudio convencional realizado al paciente, seleccionando para su análisis la secuencia sagital potenciada en T2 y la axial potenciada en T2.

Material y métodos

Todas las imágenes de los pacientes estaban almacenadas en formato DICOM. En primer lugar, las imágenes en formato DICOM fueron transformadas a formato NIFTI mediante el software dicom2nii.

Utilizando el software ITK-SNAP 3.6.0 se realizó la segmentación manual de las ROI siguientes:

- En el corte sagital medio de la secuencia potenciada en T2 se segmentaron los discos intervertebrales L3-L4, L4-L5, L4-S1, los platillos vertebrales superiores e inferiores de cada uno de estos discos, y una ROI localizada en el líquido cefalorraquídeo para la normalización de la intensidad de imagen (Figura 7).

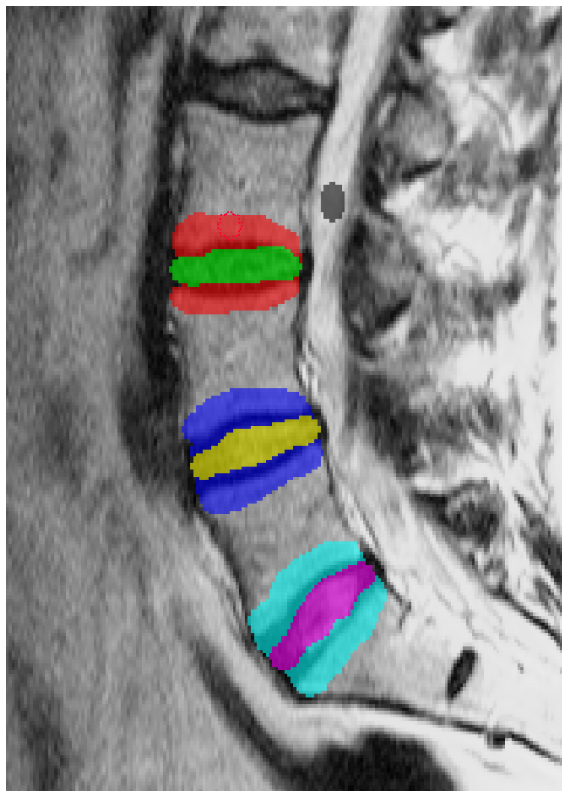


Figura 7. ROI segmentadas en el corte sagital medio en la secuencia potenciada en T2.

- En el corte transversal potenciado en T2 sobre el disco intervertebral L4-L5 se segmentaron dos ROIs correspondientes a la musculatura paravertebral lumbar derecha e izquierda. También se segmentaron dos ROIs situadas simétricamente a la línea media en la grasa subcutánea que se utilizaron para la normalización de la intensidad de la imagen (Figura 8).

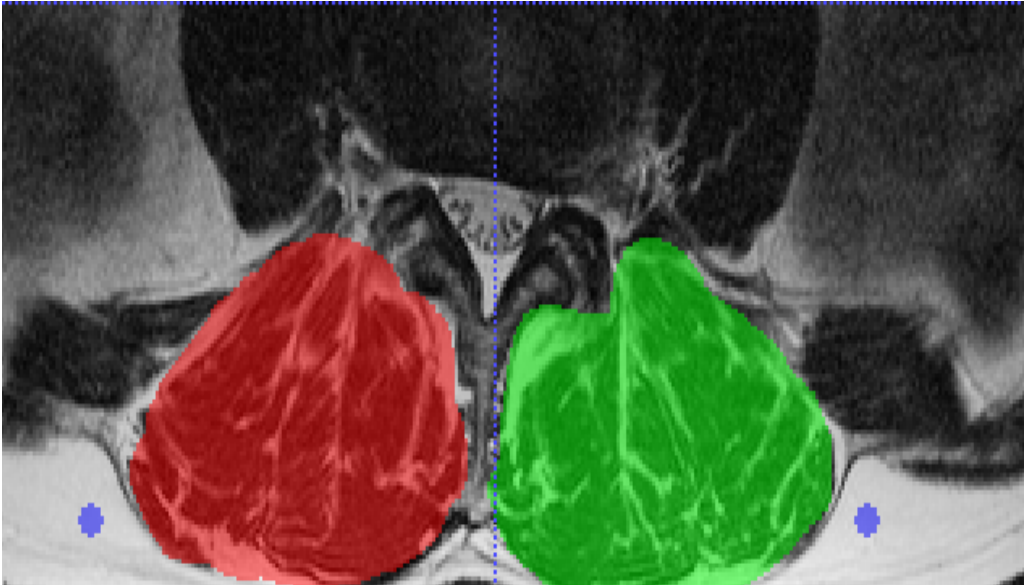


Figura 8. ROIs segmentadas en el corte axial en la secuencia potenciada en T2.

La segmentación se realizó dos veces en cada uno de los pacientes, en distinto momento temporal, con el fin de establecer la reproductibilidad del procedimiento y eliminar el efecto negativo que pudiera generarse al realizar la segmentación manual, promediándose posteriormente los valores obtenidos para cada ROI.

A partir de las ROIs seleccionados se procedió al cálculo de las variables de textura. Para ello se utilizó el software PyRadiomics (53) y se obtuvieron los 107 parámetros enumerados anteriormente.

Las variables de textura obtenidas de las ROI de disco de los tres niveles analizados se promediaron obteniendo un valor único por paciente

Material y métodos

para cada una de ellas; el mismo procedimiento se realizó con las variables obtenidas a partir de las 6 ROI de los platillos vertebrales de cada paciente.

Previo al análisis estadístico se realizó una selección de las variables de textura de interés, aquellas características que mayor información pudieran aportar en nuestro estudio, eliminando aquellas que su aportación a un modelo estadístico fuera inapreciable o aquellas que fueran redundantes, pudiendo dar lugar a problemas de multicolinealidad. Para hacer frente a este problema se planteó una estrategia de selección de características que se basaba en 4 pasos (90):

- Evaluación de la robustez de la variable. Dado que cada ROI se segmentó dos veces en dos momentos temporales distintos para cada paciente, se analizó la variabilidad de cada uno de los parámetros mediante la correlación intraclase (ICC). Esta técnica permite evaluar la fiabilidad de cada una de las características calculadas y ofrece una sola calificación con la que poder filtrar los parámetros. En este estudio se estableció que todas aquellas características cuyo ICC fuera inferior a 0.85 no se utilizarán en el análisis estadístico posterior.
- Evaluación de la información. Se calculó la desviación absoluta de la mediana (MAD) de cada característica. MAD es una medida de la dispersión similar a la desviación estándar pero más robusta a los valores atípicos. El criterio que se siguió fue incluir aquellas características cuyo MAD fuera superior a 0.
- Evaluación del rendimiento del pronóstico. Se calculó el índice de concordancia (C-Index) que mide la media de la frecuencia con la que un modelo predictor dice que X es mayor que Y cuando, en los datos observados, X es en realidad mayor que Y. Por lo tanto, utilizando los resultados obtenidos para las segmentaciones realizadas en dos momentos temporales, podemos ver la concordancia que habrá para cada uno de los parámetros. El criterio exclusión fue eliminar aquellas características cuyo C-Index fuera inferior a 0.58.

- Reducción de la redundancia. Reducir la alta dimensionalidad de los parámetros reduciendo aquellas variables cuya correlación fuera alta (>0.9), manteniendo en el estudio aquellas características cuyo C-Index fuera mayor.

4.7 *Análisis de datos*

4.7.1 *Estadísticos descriptivos*

Se obtuvieron los estadísticos descriptivos expresando las variables continuas en media (con desviación estándar) o mediana (con rango intercuartil) según su normalidad o no y las variables categóricas en número (con porcentaje).

La distribución normal se comprobó con el test de Kolmogorof-Smirnov para determinar el empleo de pruebas paramétricas o no paramétricas.

4.7.2 *Análisis de las variables cualitativas*

Para analizar la correlación entre dolor y discapacidad con los hallazgos estructurales en imagen se utilizó el coeficiente de correlación de Spearman. Se estudió la correlación entre el grado de Pfirrmann de degeneración discal de cada espacio y la puntuación END y RMQ que presentaban los pacientes en la visita inicial, así como la correlación con dichos valores a los 6 meses, con el fin de determinar su posible valor pronóstico en la evolución de la enfermedad. También se realizó este análisis considerando solamente el disco con mayor degeneración en cada paciente y con el promedio del grado de degeneración de todos los discos analizados. También se estudió si existían diferencias entre el grado de degeneración discal de los distintos espacios, el espacio con más degenerado y el valor promedio entre los pacientes que presentaban mejoría del END clínicamente relevante (mayor o igual al 30% de la puntuación inicial) y los que no, así como entre pacientes con discapacidad ($RMQ > 4$) o sin ella al final del seguimiento; estos análisis se realizaron utilizando la prueba U, no paramétrica, de Mann-Whitney de muestras independientes.

Material y métodos

La relación entre la presencia de cambios de Modic, ZAI y grado de degeneración de la musculatura retromiocrural con el dolor y la discapacidad al final del seguimiento se analizó mediante la prueba U, no paramétrica, de Mann-Whitney de muestras independientes, comparando la distribución de los valores de END y RMQ entre los pacientes con estos hallazgos radiológicos o sin ellos.

Se consideró significativo un valor $p < 0,05$.

Estos análisis fueron realizados utilizando el software IBM SPSS Statistics (versión 25).

4.7.3 Análisis de las variables de textura

4.7.3.1 Correlación de las variables clínicas con los parámetros de textura

Se estudió la correlación entre el valor de dolor END a los 6 meses y los distintos parámetros de textura seleccionados calculando el coeficiente de correlación de Spearman.

Se analizó la distribución de los valores de las variables de textura seleccionadas en el grupo formado por los pacientes que habían presentado una mejoría del dolor END de al menos un 30 % y en el grupo formado por aquellos que no. Se compararon estas distribuciones mediante la prueba U, no paramétrica, de Mann-Whitney de muestras independientes.

Los mismos análisis se realizaron para determinar la correlación entre el valor de discapacidad RMQ a los 6 meses con las variables de textura y la distribución de los valores de los mismos en dos grupos formados por los pacientes con un RMQ igual o menor a 4 y mayor a 4 al final del seguimiento.

Estos análisis fueron realizados utilizando el software IBM SPSS Statistics (versión 25).

4.7.3.2 Modelos Predictivos

Con el fin de elaborar modelos predictivos que permitieran determinar la situación clínica del paciente, grado de dolor y discapacidad al final del seguimiento se utilizaron algoritmos de “*machine learning supervisado*”.

El aprendizaje supervisado trabaja con datos etiquetados, es decir, datos para los que ya conoce la respuesta de destino. En base a un histórico de datos, trata de buscar patrones relacionándolos con un campo especial, llamado objetivo, dadas unas variables de entrada. A través de dicho histórico de datos, el algoritmo puede aprender a asignar una etiqueta de salida o función que le permita predecir el atributo objetivo para una nueva acción. El aprendizaje supervisado se utiliza en dos tipos de problemas: regresión y clasificación.

El análisis de regresión se aplica cuando queremos predecir el valor de una variable continua. Para evaluar el rendimiento de estos modelos se calcula la diferencia entre la predicción y el valor real de la variable. La métrica más comúnmente utilizada para las tareas de regresión es el error cuadrático medio (RMSE) y representa a la raíz cuadrada de la distancia cuadrada promedio entre el valor real y el valor pronosticado. Los valores más bajos de RMSE indican un mejor ajuste. RMSE es una buena medida de la precisión con que el modelo predice la respuesta, y es el criterio más importante para ajustar si el propósito principal del modelo es la predicción. Otra métrica aplicada en los problemas de regresión es el error absoluto medio (MAE), promedio de la diferencia absoluta entre el valor observado y los valores predichos.

El análisis de clasificación se aplica en el estudio de variables categóricas; a partir del análisis de determinadas variables conocidas, predeciremos el valor de la variable que estamos estudiando, seleccionando entre las distintas opciones posibles, por ejemplo, si un paciente mejorará o no con un determinado tratamiento. Para evaluar el rendimiento de estos modelos se elabora una matriz de confusión a partir de la cual se calculan diversos parámetros como exactitud (*accuracy*),

Material y métodos

sensibilidad (*recall*), especificidad, valor predictivo positivo (*precision*), valor predictivo negativo y área bajo la curva ROC.

4.7.3.2.1 Análisis de regresión

A partir de las variables sociodemográficas, variables clínicas y parámetros de textura obtenidos se aplicó un modelo Random Forest de regresión haciendo un k-fold con $k = 3$ (método de validación cruzada en el que la muestra se divide en 3 grupos, el modelo se entrena en dos de ellos y se aplica en el tercero, realizando dicho proceso tres veces, cambiando los grupos, para obtener el resultado). Se realizaron varios modelos, en cada caso incorporando todas las variables o las 20 primeras variables más importantes según su valor SHAP (*Shapley Additive exPlanations*, mide la contribución de cada variable a la construcción del modelo y se utiliza como método de selección de variables); variables centradas y escaladas con media 0 y varianza unitaria. Las predicciones que se ensayaron fueron las siguientes:

- Predicción del dolor (END) a los 6 meses
- Predicción de la variación absoluta del dolor (END) a los 6 meses
- Predicción del % de variación del dolor (END) a los 6 meses
- Predicción de la discapacidad (RMQ) a los 6 meses
- Predicción de la variación absoluta de la discapacidad (RMQ) a los 6 meses
- Predicción del % de variación de la discapacidad (RMQ) a los 6 meses

Para evaluar el rendimiento de estos modelos, la diferencia entre la predicción y el valor real, se calculó el error cuadrático medio (RMSE) y el error absoluto medio (MAE).

4.7.3.2.2 Análisis de clasificación

A partir de las variables de textura seleccionadas se realizó un análisis de clasificación con el fin de obtener un modelo predictivo que

permitiera encuadrar a los pacientes en diversas categorías, según su respuesta al tratamiento:

- Pacientes con mejoría del dolor (END) inferior a 30% (mínimo cambio clínicamente relevante) a los 6 meses.
- Pacientes con puntuación del cuestionario de discapacidad RMQ mayor a 4 a los 6 meses; los individuos con puntuación igual o menor a 4 se considera que no tienen discapacidad significativa.

Para realizar este análisis se utilizó el algoritmo Random Forest de clasificación y se formularon varios modelos predictivos con distintas combinaciones de variables:

- Parámetros de textura calculados para todas las ROI.
- Parámetros de textura y valores iniciales de dolor (END) y discapacidad (RMQ).
- Parámetros de textura, edad, sexo, IMC, valores iniciales END y RMQ.

En todos los modelos el análisis se realizó incluyendo los parámetros de textura seleccionados mediante el método explicado anteriormente y se repitieron aplicando el análisis de componentes principales (PCA).

Cada modelo se entrenó con el 70% de la muestra y se aplicó en el 30% restante para su validación interna. Con las predicciones obtenidas en el 30% de la muestra se elaboró en cada caso una matriz de confusión y se calcularon los siguientes parámetros: exactitud (*accuracy*), sensibilidad (*recall*), especificidad, valor predictivo positivo (*precision*), valor predictivo negativo y área bajo la curva ROC. En el modelo con mejor rendimiento se calcularon las razones de verosimilitud positiva y negativa, el nomograma de Nolan, el Odds ratio diagnóstico y la F1 score.

El proceso de análisis estadístico de las variables de textura se resume en la Figura 9.

Material y métodos

Todos los modelos y análisis estadísticos se realizaron usando Python 3.6.10 (Python Software Foundation, <https://www.python.org>). Los modelos Random Forest fueron construidos usando los paquetes Sklearn (91) y Pandas (92).

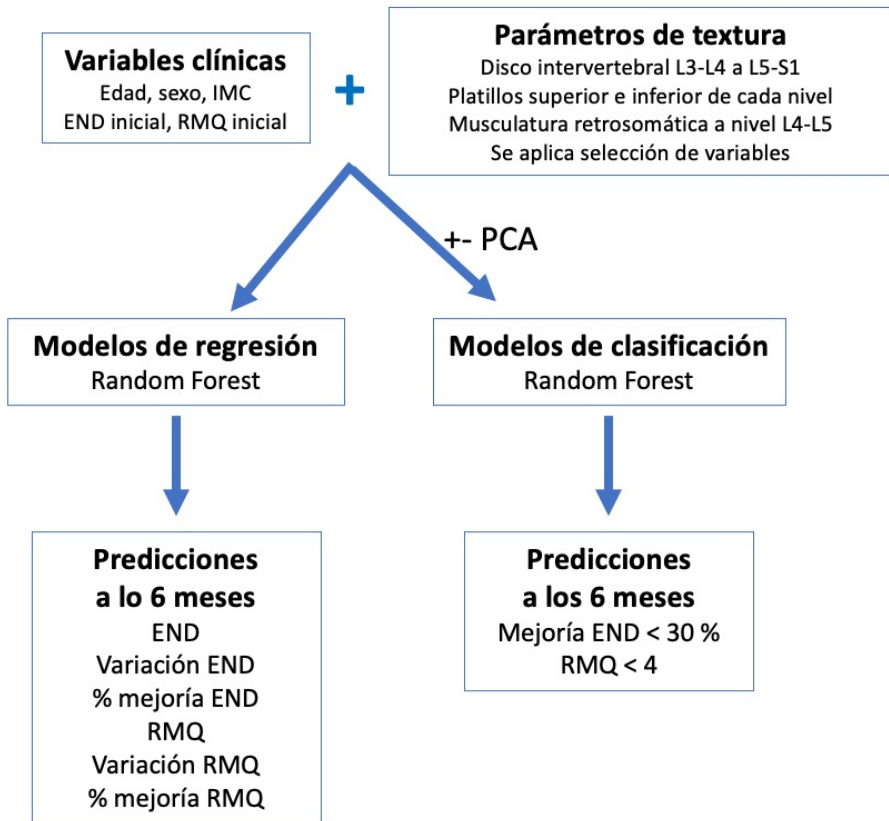


Figura 9. Esquema del análisis estadístico realizado, incluyendo las variables clínicas y los parámetros de textura.

5 Aspectos éticos

Este es un estudio descriptivo sin intervención. El proyecto se ajustó a las directrices existentes en España y UE para la protección de los pacientes en los estudios clínicos respecto a la recogida, almacenamiento y custodia de datos personales.

Se utilizaron contraseñas personales y encriptación de datos, salvaguardando la identidad de los participantes y respetando su derecho de confidencialidad. Además, los investigadores se adhirieron a la Declaración de Helsinki sobre investigación médica (93).

A todos los pacientes elegibles se les dio información oral y escrita (consentimiento informado) sobre el estudio. El presente proyecto fue aprobado por el Comité Ético de Investigación del Hospital Arnau de Vilanova de Valencia (HAV CIE 10/2017) (Anexo 3).

6 Resultados

Resultados

6.1 Características de la muestra

Se incluyeron en el estudio un total de 100 pacientes, de los cuales 94 completaron todo el seguimiento. El proceso de reclutamiento y seguimiento de los pacientes se recoge en la Figura 10.

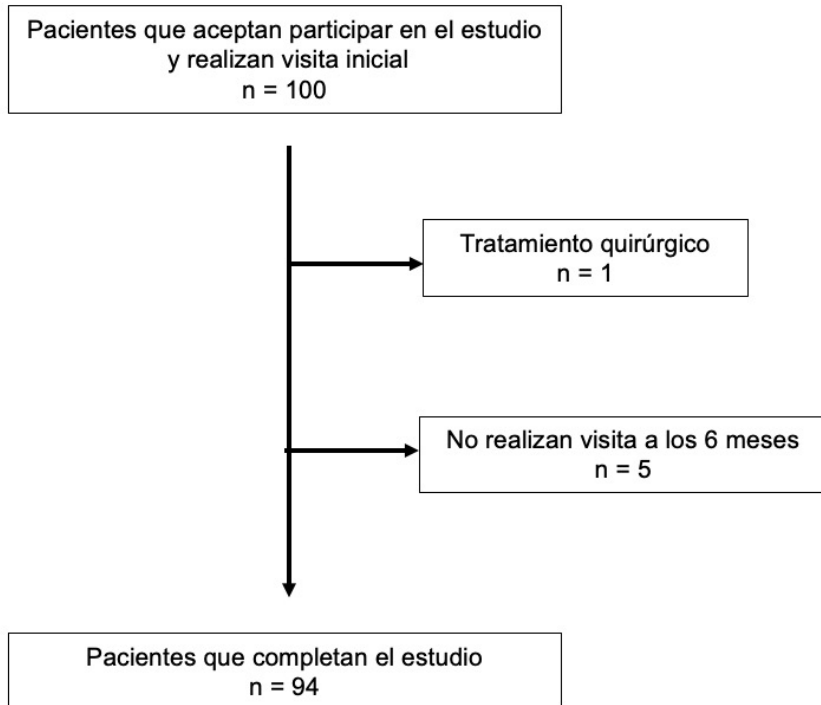


Figura 10. Diagrama de flujo del proceso de reclutamiento y seguimiento.

Las características demográficas y clínicas de la muestra se recogen en la Tabla 2.

Globalmente se trató de una muestra de pacientes de ambos sexos con discreto predominio de mujeres (61%), con una media de edad de 49 años, por lo que podemos considerar que fueron pacientes jóvenes en su mayoría. El IMC se situó en torno a 26 Kg/m², que clasificamos como sobrepeso (IMC entre 25-26,9 Kg/m²).

El valor de la mediana para la intensidad del dolor al inicio del estudio según la END fue de 7, lo que consideramos dolor severo. Al

Resultados

inicio, el valor para la discapacidad se situó en 10 según la escala RMQ, que se clasificó como moderada.

Tabla 2. Características demográficas y clínicas de la muestra. Variables categóricas expresadas en media (+-DE) edad y talla, el resto expresadas en mediana (rango intercuartil).

Edad (años)		49,57 (+/-8,92)
Sexo (%)	Varón	37 (39,4)
	Mujer	57 (60,6)
Peso (Kg)		76,4 (20,5)
Talla (m)		1,68 (+/-0,084)
IMC (Kg/m ²)		25,86 (5,17)
END Inicial (0-10)		7 (3)
END Final (0-10)		6 (4)
RM Inicial (0-24)		10 (7)
RM final (0-24)		5 (7)

Al final del seguimiento, en la visita a los 6 meses, la mediana del dolor según la END se situó en 6, lo que se considera dolor moderado y la discapacidad se situó en 5, que se clasificó como moderada. Si analizamos la evolución individual de los pacientes, 21 (20,3%) empeoraron su puntuación en la END, 19 (22,3%) no cambiaron y 54 (57,4%) refirieron mejoría; 31 (33%) pacientes mejoraron al menos un 30 % de su dolor, lo que se considera una mejoría clínicamente relevante.

Si analizamos la evolución de la discapacidad, 18 pacientes (19,15%) empeoraron su puntuación RMQ, en 5 (5,32%) no cambió y 71 (75,53%) mejoraron su grado de discapacidad; 40 (42,4%) pacientes presentaron un RMQ menor o igual a 4, lo que indica que no presentaban discapacidad al final del seguimiento.

6.2 Resultados del análisis de las variables cualitativas

6.2.1 Grado de degeneración discal

6.2.1.1 Análisis descriptivo

En la Tabla 3 se muestra la gradación de Pfirmann de degeneración discal obtenida. Los discos superiores de la columna lumbar muestran un menor grado de degeneración discal (Pfirmann 2), mientras que los segmentos distales muestran un mayor grado de degeneración (Pfirmann 3 y 4).

Tabla 3. Hallazgos estructurales de la muestra de 470 discos en 94 pacientes.

	Pfirmann 1	Pfirmann 2	Pfirmann 3	Pfirmann 4	Pfirmann 5
L1-L2	4	62	20	7	1
L2-L3	3	56	22	11	2
L3-L4	3	36	35	18	2
L4-L5	1	14	41	33	5
L5-S1	2	18	23	38	13

En la Tabla 4 se clasifican a los pacientes según el disco con mayor grado de degeneración; más de la mitad presentaban un grado de degeneración Pfirmann 4 (54,25%) en alguno de sus discos lumbares.

Tabla 4. Mayor grado de degeneración discal de cada paciente.

Pfirmann 1	Pfirmann 2	Pfirmann 3	Pfirmann 4	Pfirmann 5
0	3 (3,19%)	20 (21,28%)	51 (54,25%)	20 (21,28%)

6.2.1.2 Correlación con la variable dolor

Analizando todos los discos no se encontró correlación entre el grado de degeneración discal y el grado de dolor según la Escala Numérica del Dolor en la visita inicial (Tabla 5). Tampoco se apreció correlación eligiendo solamente el disco más degenerado en cada caso (coeficiente de correlación de Spearman $R = 0,174$, $p = 0,144$) o utilizando para el cálculo el promedio de la puntuación Pfirmann de los discos estudiados (coeficiente de correlación de Spearman $R = 0,010$, $p = 0,923$).

Tabla 5. Coeficiente de correlación de Spearmann entre el grado de degeneración discal en los distintos niveles y la puntuación END inicial.

Nivel del Disco	R	p-valor
L1-L2	-0,157	0,130
L2-L3	-0,094	0,369
L3-L4	-0,061	0,557
L4-L5	0,056	0,591
L5-S1	0,134	0,197
Disco peor	0,193	0,062
Promedio Pfirmann	0,010	0,923

La correlación entre la degeneración discal y el dolor a los 6 meses mostró resultados similares a los obtenidos con los momentos iniciales, observando solamente una correlación débil a nivel de L4-L5 ($R = 0,244$; $p = 0,039$). Los resultados de este análisis se recogen en la Tabla 6.

Tampoco se apreció correlación eligiendo solamente el disco más degenerado en cada caso (coeficiente de correlación de Spearman $R = 0,220$; $p = 0,062$) o utilizando para el cálculo el promedio de la puntuación Pfirmann de los discos estudiados (coeficiente de correlación de Spearman $R = -0,54$; $p = 0,605$).

El grado de degeneración discal del nivel L4-L5 mostró una diferencia estadísticamente significativa entre el grupo de pacientes con mejoría clínicamente relevante a los 6 meses (mediana = 3, rango = 4, n = 31) o sin ella (mediana = 3, rango = 3, n = 63); $U = 680,500$, $p = 0,011$ (Figura 11). Sin embargo, no se encontraron diferencias estadísticamente significativas en el resto de los cálculos realizados en el resto de niveles.

Tabla 6. Coeficiente de correlación de Spearman entre el grado de degeneración discal en los distintos niveles y la puntuación END a los 6 meses (*diferencia estadísticamente significativa).

Nivel del Disco	R	p-valor
L1-L2	-0,198	0,960
L2-L3	-0,079	0,511
L3-L4	0,115	0,337
L4-L5	0,244	0,039*
L5-S1	0,107	0,369
Disco peor	0,040	0,702
Promedio Pfirmann	-0,54	0,605

Resultados

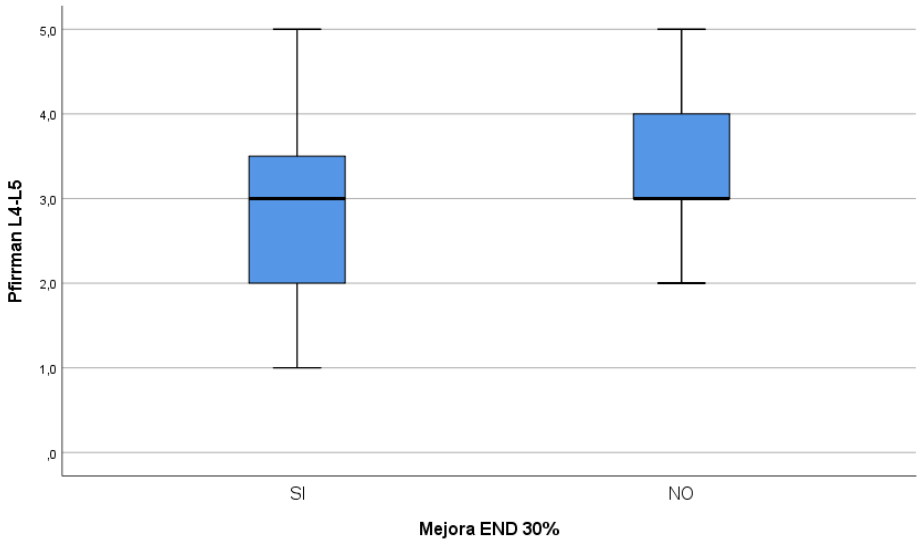


Figura 11. Grado de degeneración del disco L4-L5 en pacientes con mejoría de la puntuación END clínicamente relevante o sin ella ($p = 0,011$).

6.2.1.3 Correlación con la variable discapacidad

Analizando todos los discos no se encontró correlación entre el grado de degeneración discal y el grado de discapacidad según la escala RMQ en la visita inicial; tampoco se apreció correlación eligiendo solamente el disco más degenerado en cada caso ni considerando el promedio del grado de Pfirman de los discos estudiados para el cálculo. Los resultados del análisis se recogen en la Tabla 7.

Analizando todos los discos de la muestra, no se encontró asociación estadísticamente significativa entre la degeneración discal y la discapacidad a los 6 meses. Considerando el disco con mayor degeneración de cada paciente o el promedio del grado de Pfirman de los discos estudiados tampoco se encontró que existiera correlación estadísticamente significativa. Los resultados de este análisis se recogen en la Tabla 8.

Tabla 7. Coeficiente de correlación de Spearman entre el grado de degeneración discal y la puntuación RMQ inicial.

Nivel del Disco	R	p-valor
L1-L2	-0,057	0,634
L2-L3	0,007	0,950
L3-L4	-0,227	0,054
L4-L5	0,016	0,894
L5-S1	0,049	0,683
Disco peor	0,020	0,865
Promedio Pfirmann	0,020	0,845

Tabla 8. Coeficiente de correlación de Spearman entre el grado de degeneración discal y la puntuación RMQ a los 6 meses.

Nivel del Disco	R	p-valor
L1-L2	-0,082	0,431
L2-L3	0,097	0,350
L3-L4	-0,104	0,317
L4-L5	-0,006	0,957
L5-S1	-0,105	0,313
Disco peor	0,17	0,884
Promedio Pfirmann	-0,024	0,822

Resultados

No se encontraron diferencias estadísticamente significativas cuándo se comparó el grado de degeneración discal (Pfirmann de cada espacio, Pfirmann peor y promedio) entre los grupos con discapacidad (RMQ>4) o sin ella al final del seguimiento.

6.2.2 Cambios de Modic

6.2.2.1 Análisis descriptivo

El 77,02% de los pacientes no presentaron cambios de Modic. Cinco pacientes (5,31%) presentaron cambios Modic I y 19 (20,21%) presentaron cambios Modic II, coexistiendo ambos en dos niveles distintos en un paciente. No se hallaron en ninguno de los casos cambios Modic III. El 96,8% de estos cambios se localizaron en los espacios L4-L5 y L5-S1. En la Tabla 9 se recogen la localización y tipo de estos cambios de Modic.

Tabla 9. Cambios de Modic observados según tipo y localización.

	Modic I	Modic II	Modic III
L1-L2	0	0	0
L2-L3	0	0	0
L3-L4	2	1	0
L4-L5	1	6	0
L5-S1	2	12	0

6.2.2.2 Correlación con las variables clínicas

No se encontró diferencia estadísticamente significativa en la puntuación en la escala END en la visita a los 6 meses entre los pacientes que presentaban cambios de Modic tipo I y/o II (mediana 6, n = 23) y los que no los presentaban (mediana 6, n = 71) U = 644,5; p = 0,13.

Tampoco se encontró diferencia en la puntuación RMQ a los 6 meses entre estos dos grupos (mediana 5 / 5) U = 758; p = 0,58. No se analizó la relación con los cambios de Modic tipo III debido a que no se encontraron en la muestra de pacientes.

6.2.3 Zonas de Alta Intensidad

6.2.3.1 Análisis descriptivo

Se encontraron ZAI en 45 pacientes que afectaban a 65 de los 470 discos analizados (13,8%) y se localizaron con mayor frecuencia en los discos L4-L5 y L5-S1. Los resultados se recogen en la tabla 10.

Tabla 10: Zonas de alta intensidad (ZAI) y su localización.

Localización	L1-L2	L2-L3	L3-L4	L4-L5	L5-S1
Nº ZAI	2	5	10	22	26

6.2.3.2 Correlación con las variables clínicas

No se encontró diferencia estadísticamente significativa en la puntuación en la escala END a los 6 meses entre los pacientes que presentaban ZAI a nivel de los discos intervertebrales (mediana = 6; n = 45) y los que no las presentaban (mediana = 6; n = 49) $U = 1084,5$; $p = 0,891$.

No se encontraron diferencias en la puntuación RMQ a los 6 meses entre estos dos grupos (medianas = 5/5) $U = 1077,5$; $p = 0,849$.

6.2.4 Degeneración de la musculatura retrosomática

6.2.4.1 Análisis descriptivo

Cuarenta y un pacientes (43,62%) presentaron un grado de degeneración grasa de la musculatura moderado y 27 (28,72%) presentaron degeneración severa; en 26 (27,66%) de los casos analizados no se apreció degeneración grasa de la musculatura retrosomática.

6.2.4.2 Correlación con las variables clínicas

No se encontró diferencia estadísticamente significativa en la puntuación en la escala END en la visita a los 6 meses entre los pacientes

Resultados

que presentaban degeneración grasa de la musculatura retrosomática (Mediana = 6, n = 68) y los que no la presentaban (Mediana = 6, n = 26) $U = 797,5$; $p = 0,461$; tampoco se encontró diferencias en la presencia de degeneración grasa entre los pacientes que mejoraban un 30% su END y los que no ($U = 902,4$, $p = 0,442$).

No se encontró diferencia significativa en la puntuación RMQ a los 6 meses entre estos dos grupos, mediana 5 en ambos, $U = 982$, $p = 0,406$.

6.2.5 Análisis de regresión múltiple

Dado que no se observaron correlaciones de las distintas variables independientes (grado de Pfirmann, sexo, edad e IMC) con las variables dependientes (END y RMQ) no se pudieron construir ecuaciones predictivas válidas que explicaran el dolor y la discapacidad iniciales o al final del seguimiento a los 6 meses.

6.3 Resultados del análisis de textura

6.3.1 Parámetros de textura seleccionados

Aplicando el método de reducción de dimensionalidad se seleccionaron un total de 45 parámetros de textura del total de 107 parámetros básicos de la biblioteca PyRadiomics calculados inicialmente. Fueron seleccionadas 14 variables calculadas en ROI de disco intervertebral, 23 calculadas en ROI de musculatura retrosomática y 8 en platillo vertebral. Estas variables seleccionadas se detallan en la tabla 11. En adelante mantenemos su denominación en inglés para evitar ambigüedad y favorecer el intercambio de conocimiento.

Tabla 11. Parámetros de textura seleccionados.

ROI Platos vertebral	ROI musculatura retrosomática
Shape Surface Area	Shape Major Axis Length
Shape Major Axis Length	Shape Minor Axis Length
Shape Maximum 2D Diameter Column	Shape Voxel Volume
First order Percentile 90	Shape Sphericity
First order Total Energy	Shape Surface Volume Ratio
GLCM Imc2	Shape Maximum 2D Diameter Column
GLZCM Gray Level Variance	Shape Surface Volume Ratio
GLSZM Zone Variance	First order Energy
ROI Disco intervertebral	First order Kurtosis
Shape Maximum 2D Diameter Column	First order 10 Percentile
Shape Surface Volume Ratio	First order Median
Shape Voxel Volume	First order Maximum
First order Root Mean Squared	GLCM Idn
First order Mean Absolute Deviation	GLCM Idm
First order Percentil 10	GLSZM Gray Level Non Uniformity
First order Total Energy	GLSZM Low Gray Level Zone Emphasis
First order Skewness	GLSZM Large Area High Gray Level Emphasis
First order Kurtosis	GLSZM Zone Variance
GLCM Imc1	NGTDM Strength
GLCM Cluster Shade	NGTDM Coarseness
GLSZM Large Area High Gray Level Emphasis	NGTDM Busyness
NGTDM Busyness	GLDM Dependence Non Uniformity
GLDM Large Dependence Low Gray Level Emphasis	GLDM Large Dependence Low Gray Level Emphasis

Resultados

6.3.2 Correlación de las variables clínicas con los parámetros de textura

De las 45 de las variables de textura analizadas, 14 mostraron tener una correlación con el valor END a los 6 meses. En la Tabla 12 se recogen los resultados del análisis de correlación entre el valor END a los 6 meses y las variables de textura seleccionadas.

Tabla 12. Resultados de la prueba de correlación de Spearman de los valores de los parámetros de textura con el valor END (escala numérica del dolor) a los 6 meses. En negrita las p estadísticamente significativas.

Parámetros de textura seleccionados	R	p-valor
ROI Plátanos vertebral		
Shape Surface Area	0,102	0,328
Shape Major Axis Length	0,009	0,930
Shape Maximum 2D Diameter Column	0,168	0,105
First order Percentile 90	-0,194	0,06
First order Total Energy	0,043	0,683
GLCM Imc2	-0,293	0,004
GLZCM Gray Level Variance	-0,288	0,05
GLSZM Zone Variance	0,368	<0,001
ROI musculatura retrosomática		
Shape Major Axis Length	0,033	0,755
Shape Minor Axis Length	-0,028	0,788
Shape Voxel Volume	-0,095	0,363
Shape Sphericity	-0,169	0,104
Shape Maximum 2D Diameter Column	0,007	0,950
Shape Surface Volume Ratio	0,237	0,022
First order Energy	0,015	0,890
First order Kurtosis	-0,116	0,268
First order 10 Percentile	0,162	0,122
First order Median	0,129	0,217
First order Maximum	-0,093	0,376
GLCM Idn	-0,003	0,978
GLCM Idm	0,122	0,242
GLSZM Gray Level Non Uniformity	-0,012	0,913
GLSZM Low Gray Level Zone Emphasis	0,224	0,031
GLSZM Large Area High Gray Level Emphasis	-0,075	0,477

GLSZM Zone Variance	0,368	<0,001
NGTDM Strength	-0,135	0,195
NGTDM Coarseness	0,083	0,427
NGTDM Busyness	0,204	0,050
GLDM Dependence Non Uniformity	-0,129	0,219
GLDM Large Dependence Low Gray Level Emphasis	0,228	0,028
ROI Disco intervertebral		
Shape Maximum 2D Diameter Column	0,076	0,469
Shape Surface Volume Ratio	0,013	0,898
Shape Voxel Volume	0,022	0,830
First order Root Mean Squared	0,052	0,620
First order Mean Absolute Deviation	-0,190	0,066
First order Percentil 10	0,356	<0,001
First order Total Energy	0,046	0,657
First order Skewness	0,360	<0,001
First order Kurtosis	0,167	0,108
GLCM Imcl	0,174	0,093
GLCM Cluster Shade	0,277	0,007
GLSZM Large Area High Gray Level Emphasis	-0,222	0,032
NGTDM Busyness	0,312	0,002
GLDM Large Dependence Low Gray Level Emphasis	0,298	0,004

En el análisis de la distribución del valor de los parámetros de textura entre los grupos con y sin mejoría significativa del dolor, de las 45 variables estudiadas, encontramos que 19 de ellas mostraban diferencia significativa entre ambos grupos. Los resultados del análisis de estas variables se muestran en la Tabla 13. La distribución de los valores de los parámetros de textura con p-valor igual o menor a 0,001 se representan gráficamente en las figuras 12 a 17.

Resultados

Tabla 13. Resultados de la prueba U de Mann Whitney mediante la que se compara la distribución de los valores de las distintas variables de textura entre el grupo de pacientes con mejoría significativa del dolor (grupo 1) y los que no mejoran (grupo 2). Variables con diferencia significativa entre grupos ($p < 0,05$).

Parámetro de textura	Grupo 1 n= 31 Mediana	Grupo 2 n= 63 Mediana	U	p-valor
ROI Plátano vertebral				
Shape Maximum 2D Diameter Column	25	27	732	0,049
GLCM Imc2	1	0,99	643	0,007
GLSZM Gray Level Variance	63,56	38,04	654	0,009
GLSZM_Zone Variance	0,29	0,67	524	<0,001
GLCM Cluster Shade	474,33	194,33	696	0,024
ROI musculatura				
Shape Sphericity	0,398	0,390	691	0,028
GLSZM Low Gray Level Zone Emphasis	0,015	0,020	675	0,020
GLSZM Zone Variance	1,25	3,01	717	0,047
GLDM Large Dependence Low Gray Level Emphasis	0,78	0,13	636	0,008
NGTDM Strength	16,44	14,20	718	0,048
NGTDM Busyness	0,19	0,24	640	0,009
ROI Disco Intervertebral				
First order Mean Absolute Deviation	20,48	12,39	566	0,001
First order 10 Percentile	-53,03	-48,02	598	0,002
First order Skewness	0,35	0,50	623	0,004
GLCM Imc1	-0,32	-0,41	560	0,001
GLCM Cluster Shade			642	0,007
GLSZM Large Area High Gray Level Emphasis	190,47	266,20	672	0,014
NGTDM Busyness	0,19	0,41	556	0,001
GLDM Large Dependence Low Gray Level Emphasis	0,61	1,35	647	0,008

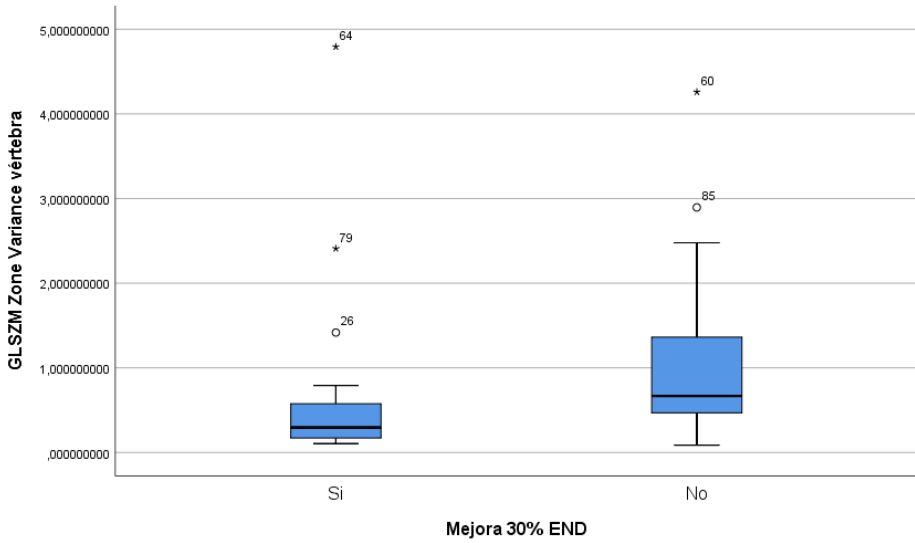


Figura 12. Valores del parámetro de textura GLSZM Zone Variance calculado en ROI de platillo vertebral ($p < 0,001$).

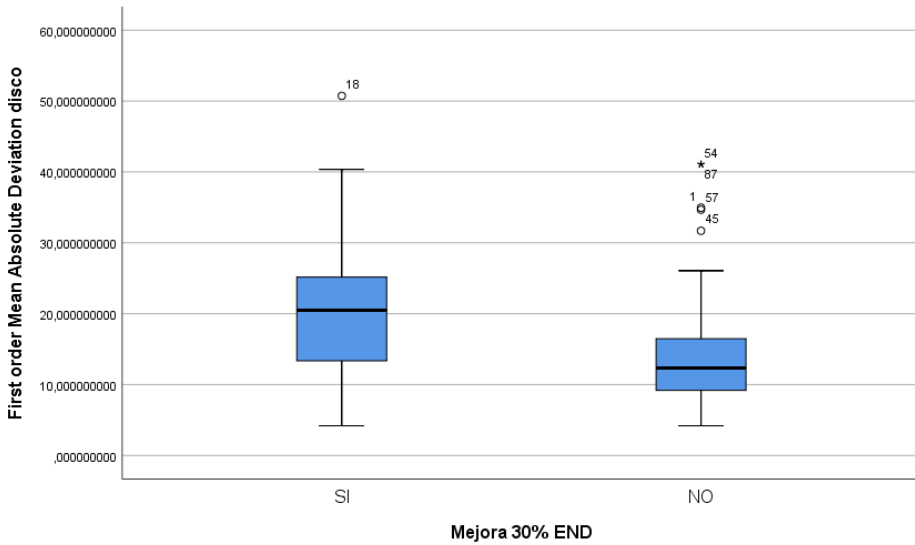


Figura 13. Valores del parámetro de textura Mean Absolute Deviation calculado en ROI disco intervertebral ($p = 0,001$).

Resultados

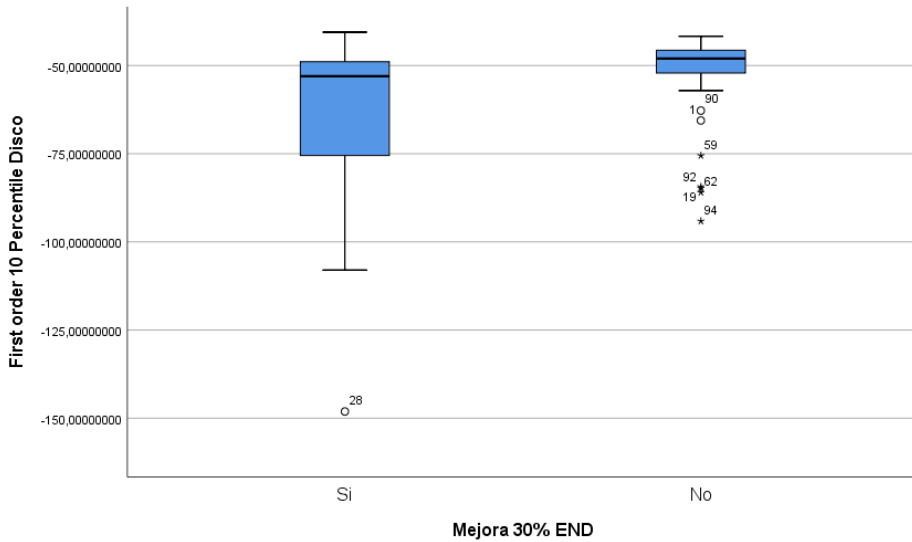


Figura 14. Valores del parámetro de textura de primer orden Percentil 10 calculado en ROI disco intervertebral ($p = 0,001$).

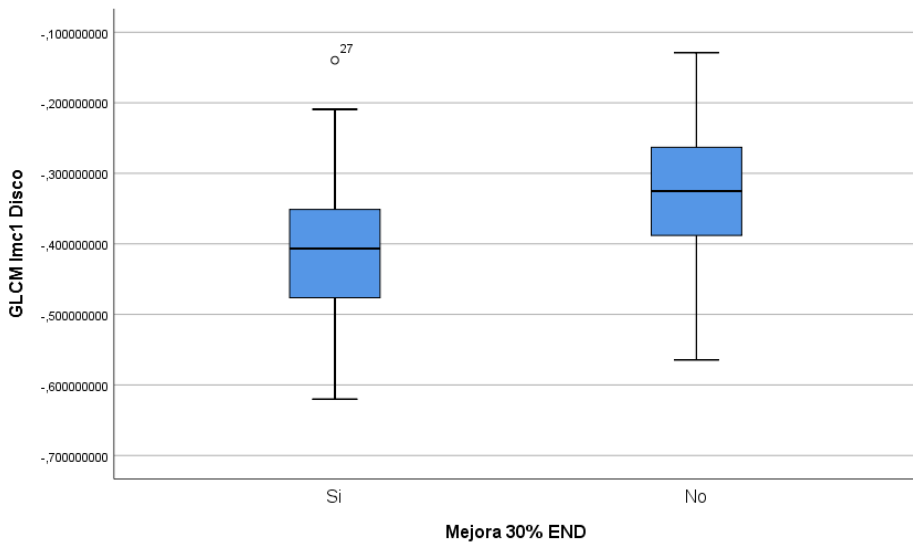


Figura 15. Valores del parámetro de textura GLCM lmc 1 calculado en ROI disco intervertebral ($p = 0,001$).

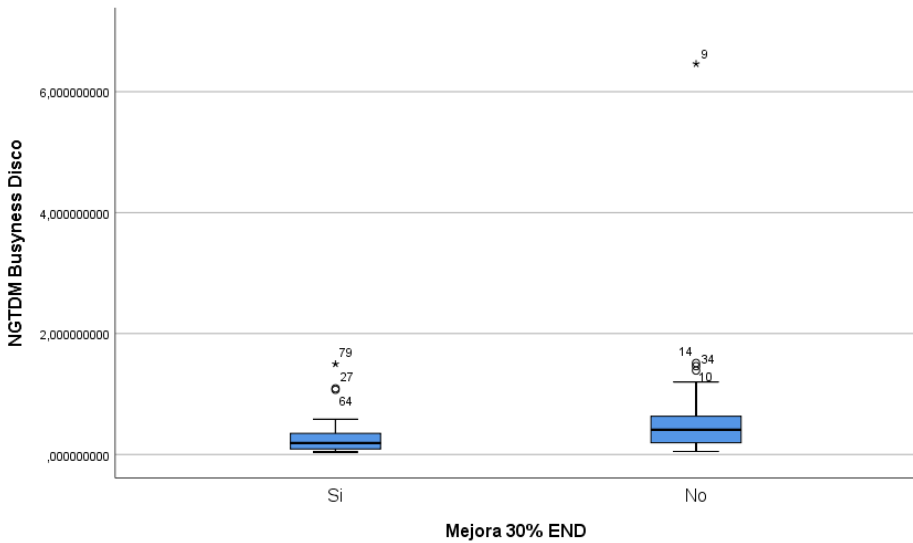


Figura 16. Valores del parámetro de textura NGTDM Busyness calculado en ROI disco intervertebral ($p = 0,001$).

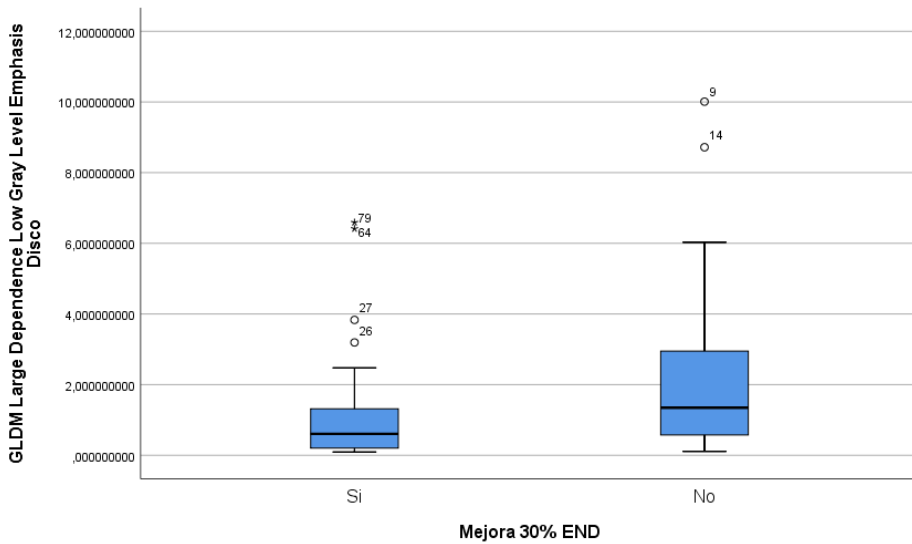


Figura 17. Valores del parámetro de textura GLDM Large Dependence Low Gray Level Emphasis calculado en ROI Disco intervertebral ($p < 0,001$).

Resultados

Los resultados del análisis de correlación entre el valor RMQ a los 6 meses y las variables de textura seleccionadas solamente mostraron un resultado positivo para el parámetro GLCM Cluster Shade calculado en ROI disco intervertebral ($R = 0,235$; $p = 0,022$).

De las 45 variables analizadas, solamente se encontraron diferencias significativas en la distribución de los valores de la variable de textura GLCM Cluster Shade calculada en ROI disco intervertebral entre el grupo de pacientes con discapacidad al final del seguimiento, $RMQ > 4$ (mediana 12,33; $n = 54$) o sin ella (mediana -3.12, $n = 40$) $U = 810$; $p = 0,039$.

6.3.3 Modelos de regresión

6.3.3.1 Modelos predictivos para dolor

El modelo de regresión para la predicción del valor de END a los 6 meses ofrece un $RMSE = 2.71$ y un $MAE = 2.25$. Si seleccionamos las 20 variables más importantes según su valor SHAP los resultados son $RMSE = 2.58$ y $MAE = 2.07$. Los resultados se representan gráficamente en la Figura 18.

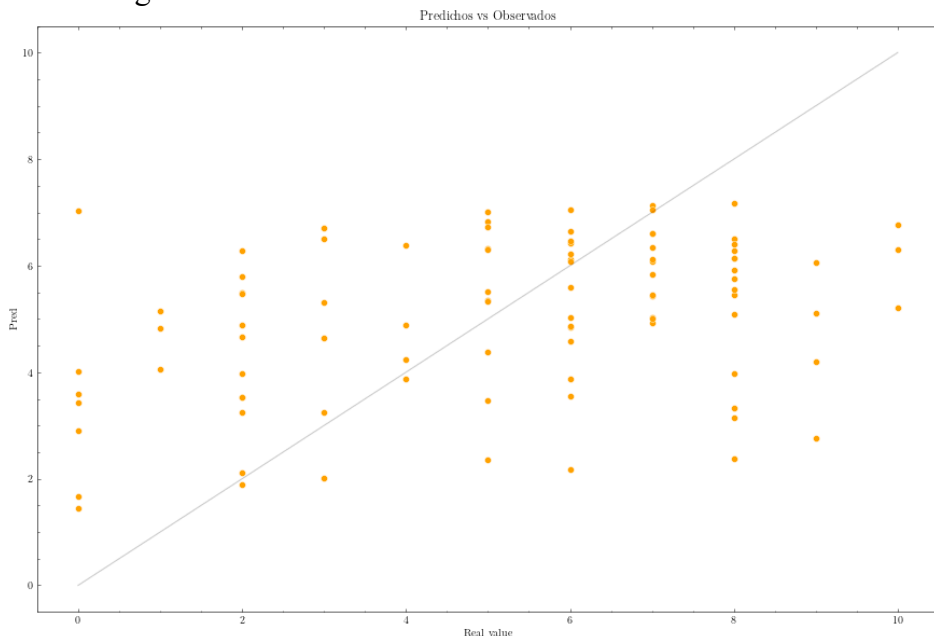


Figura 18. Distribución de los casos según el valor real y la predicción del valor END (escala numérica del dolor) a los 6 meses del modelo que incluye todas las variables.

6.3.3.2 Modelos predictivos para discapacidad

El modelo para la predicción del valor RMQ a los 6 meses ofrece un RMSE = 4.97 y un MAE = 4.06. Si seleccionamos las 20 variables más importantes según su valor SHAP los resultados son RMSE = 4.69 y MAE = 3.73. Los resultados se representan gráficamente en la Figura 19.

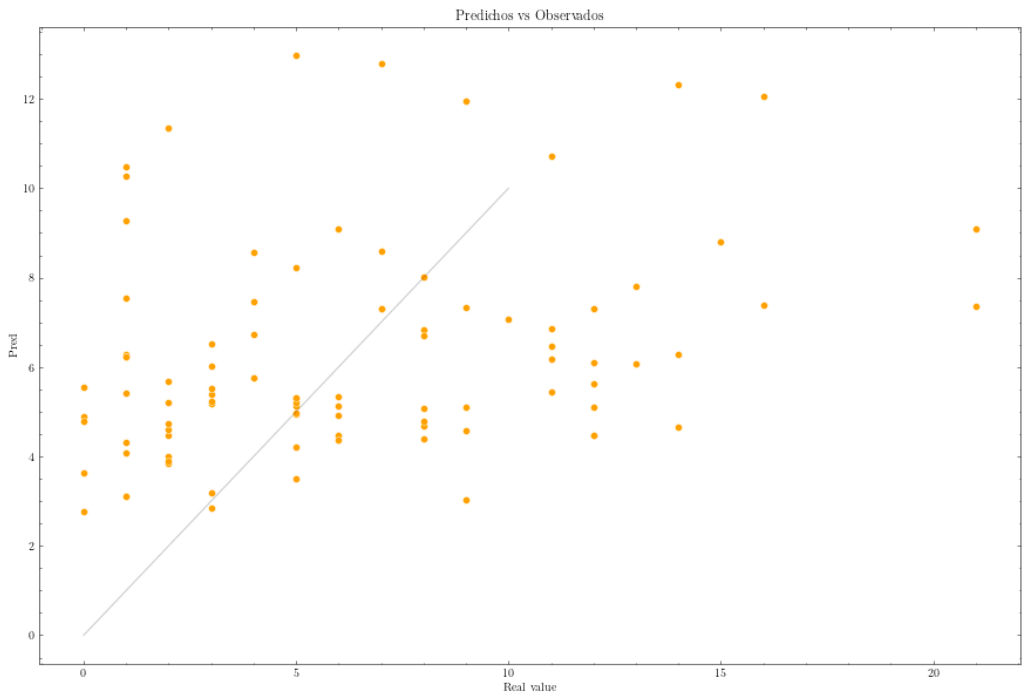


Figura 19. Distribución de los casos según el valor real y la predicción del valor RMQ (escala Rolland Morris) a los 6 meses del modelo que incluye todas las variables.

Los resultados de los distintos modelos de regresión realizados se resumen en la Tabla 14.

Resultados

Tabla 14. Resultados *Random Forest* de regresión en los distintos modelos analizados. Los valores altos del error cuadrático medio (RMSE) y del error absoluto medio (MAE) nos indican que las predicciones del modelo se alejan de los valores reales de cada una de las variables.

Predicción	RMSE	MAE
END 6 meses	2,71	2,25
Variación END	2,65	2,03
% mejoría END	46,5	34,2
RMQ 6 meses	4,69	3,73
Variación RMQ	5,08	4,09
% mejoría RMQ	53,8	43,1

6.3.4 Modelos de clasificación

6.3.4.1 Modelos predictivos para dolor

La Tabla 15 recoge los resultados obtenidos para los distintos modelos predictivos entrenados en el 70 % de la muestra para clasificar a los pacientes que no van a presentar una mejoría clínicamente significativa (reducción del 30 % o más en END) cuando se aplican sobre el 30 % de la muestra restante. Todos los modelos entrenados para identificar correctamente a dichos pacientes ofrecen sensibilidad alta, con baja especificidad.

El modelo con mejor capacidad predictiva es el que incluye en el análisis los parámetros de textura seleccionados y todas las variables clínicas incluidas en el estudio: edad, sexo, IMC, END y RMQ, aplicando PCA. La matriz de confusión para este modelo se muestra en la tabla 16. Los valores calculados para el mismo fueron: exactitud (*accuracy*) 0,79; sensibilidad (*recall*) 0,86; especificidad 0,57; valor predictivo positivo (*precision*) 0,86; valor predictivo negativo 0,40 y el área bajo la curva ROC 0,71 (Figura 20). En la tabla 16 se recogen otros indicadores calculados y en la figura 21 se representa el nomograma de Nolan para este modelo.

Tabla 15. Resultados obtenidos, aplicando Random Forest de clasificación, con los distintos modelos entrenados en el 70 % de la muestra cuando los aplicamos al 30 % restante de la misma, para identificar pacientes con mejoría de END menor o igual al 30 %. En negrita el modelo con el mejor resultado.

Variables incluidas	PCA	Exactitud (Accuracy)	Sensibilidad (Recall)	Especificidad	VPP (Precision)	VPN	AUC ROC
Selec V Textura (sVT)	NO	0,71	0,81	0,43	0,81	0,43	0,62
Selec V Textura (sVT)	SI	0,71	0,81	0,43	0,81	0,43	0,62
sVT + Edad + Sexo + I MC	NO	0,75	0,86	0,43	0,82	0,50	0,64
sVT + Edad + Sexo + IMC	SI	0,75	0,86	0,43	0,82	0,50	0,64
sVT + END + RMQ	NO	0,64	0,71	0,43	0,79	0,33	0,57
sVT + END + RMQ	SI	0,71	0,81	0,43	0,79	0,43	0,62
sVT +Todas Variables	NO	0,71	0,86	0,29	0,78	0,28	0,57
sVT +Todas Variables	SI	0,79	0,86	0,57	0,86	0,57	0,71
Sólo variables clínicas	NO	0,68	0,86	0,14	0,75	0,25	0,5
Sólo variables clínicas	Si	0,46	0,57	0,14	0,66	0,10	0,36

Resultados

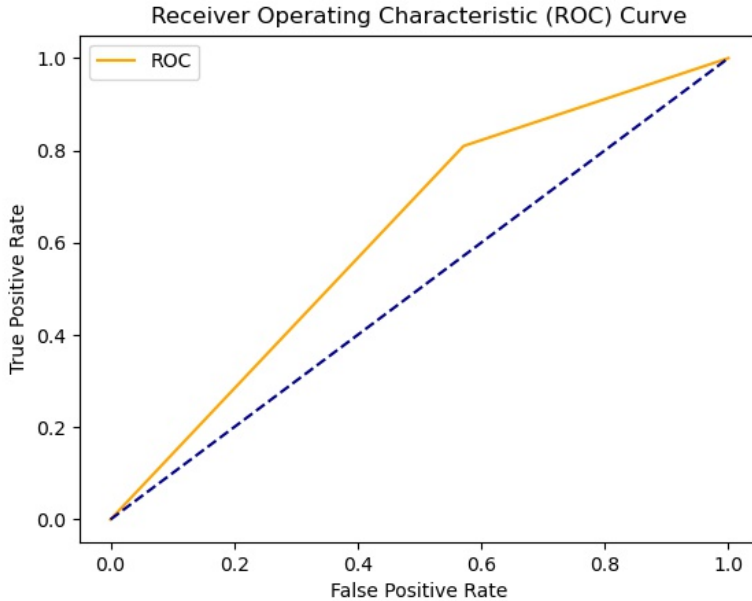


Figura 20. Curva ROC del modelo predictivo que incluye las variables de textura y todas las variables clínicas, aplicando PCA (análisis de componentes principales), para predecir mejoría de END menor o igual al 30 %. AUC=0,71.

Tabla 16. Matriz de confusión sobre el 30 % de la muestra del modelo predictivo que incluye variables de textura seleccionadas y todas las variables sociodemográficas y clínicas, aplicando PCA, para predecir mejoría de END menor o igual al 30%.

		Predicción	
		Positivo	Negativo
Observación	Positivo	18	3
	Negativo	3	4
Sensibilidad (<i>Recall</i>): 0,86		AUC ROC: 0,71	
Especificidad: 0,57		Razón de verosimilitud +: 2	
Exactitud (<i>Accuracy</i>): 0,79		Razón de verosimilitud -: 0,25	
VPP (<i>Precision</i>): 0,86		Odds Ratio diagnóstico: 8	
VPN: 0,57		F1 score: 0,86	

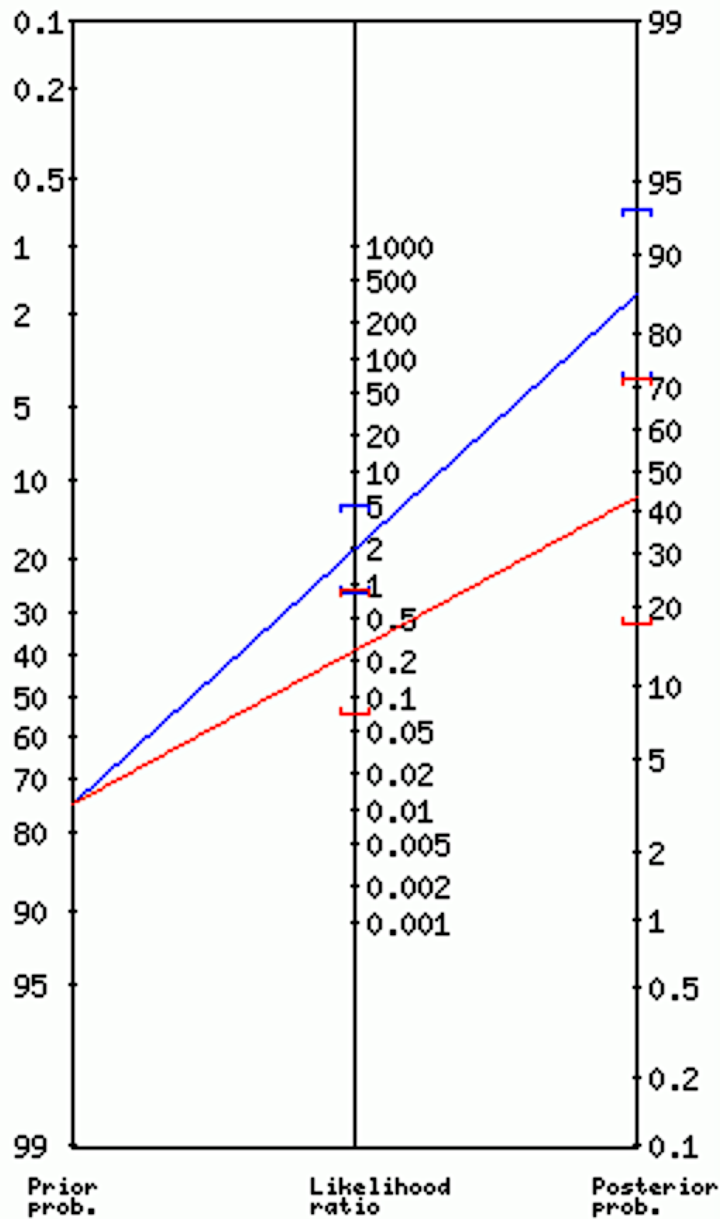


Figura 21. Nomograma de Nolan del modelo predictivo con mejor rendimiento para predecir mejoría de END (escala numérica del dolor) menor o igual al 30 %.

6.3.4.2 Modelos predictivos para discapacidad

La Tabla 17 recoge los resultados obtenidos para los distintos modelos predictivos entrenados en el 70 % de la muestra para clasificar a los pacientes que vayan a presentar algún grado de discapacidad al final del seguimiento ($RMQ > 4$) cuando se aplican sobre el 30 % de la muestra restante.

Todos los modelos presentaron mayor sensibilidad que especificidad, pero su capacidad predictiva fue inferior para discapacidad que la obtenida para dolor. El mejor resultado se obtuvo con el modelo que incluyó en el análisis los parámetros de textura seleccionados y las variables edad, sexo, IMC, sin incluir END y RMQ iniciales, sobre los que se aplicó PCA. La matriz de confusión para este modelo y los distintos indicadores calculados se muestran en la tabla 18. Para este modelo los valores calculados fueron: exactitud (*accuracy*) 0,61; sensibilidad (*recall*) 0,83; especificidad 0,20; valor predictivo positivo (*precision*) 0,65; valor predictivo negativo 0.4 y el área bajo la curva ROC 0,52.

La incorporación de las variables de textura al modelo no mejora la capacidad predictiva del mismo, obteniendo resultados similares al modelo que incluye solamente las variables clínicas y sociodemográficas.

Tabla 17. Resultados obtenidos, aplicando Random Forest de clasificación, con los distintos modelos entrenados en el 70 % de la muestra cuando los aplicamos al 30 % restante de la misma, para identificar pacientes con discapacidad a los 6 m ($RMQ > 4$). En negrita los modelos con mejor resultado.

Variables incluidas	PCA	Exactitud (Accuracy)	Sensibilidad (Recall)	Especificidad	VPP (Precision)	VPN	AUC ROC
Selección de Textura (sVT)	NO	0,43	0,5	0,3	0,56	0,25	0,4
Selección de Textura (sVT)	SI	0,54	0,72	0,2	0,62	0,29	0,46
sVT + Edad + Sexo + IMC	NO	0,39	0,5	0,2	0,53	0,18	0,35
sVT + Edad + Sexo + IMC	SI	0,61	0,83	0,20	0,65	0,4	0,52
sVT + END + RMQ	NO	0,36	0,44	0,20	0,50	0,17	0,32
sVT + END + RMQ	SI	0,50	0,61	0,30	0,61	0,30	0,46
sVT + Todas Variables	NO	0,36	0,5	0,1	0,5	0,1	0,3
sVT + Todas Variables	SI	0,50	0,56	0,4	0,63	0,33	0,48
Sólo variables clínicas	NO	0,61	0,72	0,4	0,68	0,44	0,56
Sólo variables clínicas	SI	0,57	0,67	0,4	0,67	0,4	0,53

Resultados

Tabla 18. Matriz de confusión obtenida al aplicar el modelo predictivo que incluye variables de textura seleccionadas y las variables sociodemográficas, aplicando PCA, para predecir $RMQ > 4$ al final del seguimiento, a los 6 meses.

		Predicción	
		Positivo	Negativo
Observación	Positivo	15	8
	Negativo	3	2
Sensibilidad (<i>Recall</i>): 0,83		AUC ROC: 0,52	
Especificidad: 0,20		Razón de verosimilitud +: 1,04	
Exactitud (<i>Accuracy</i>): 0,61		Razón de verosimilitud -: 0,85	
VPP (<i>Precision</i>): 0,65		Odds Ratio diagnóstico: 1,33	
VPN: 0,4		F1 score: 0,73	

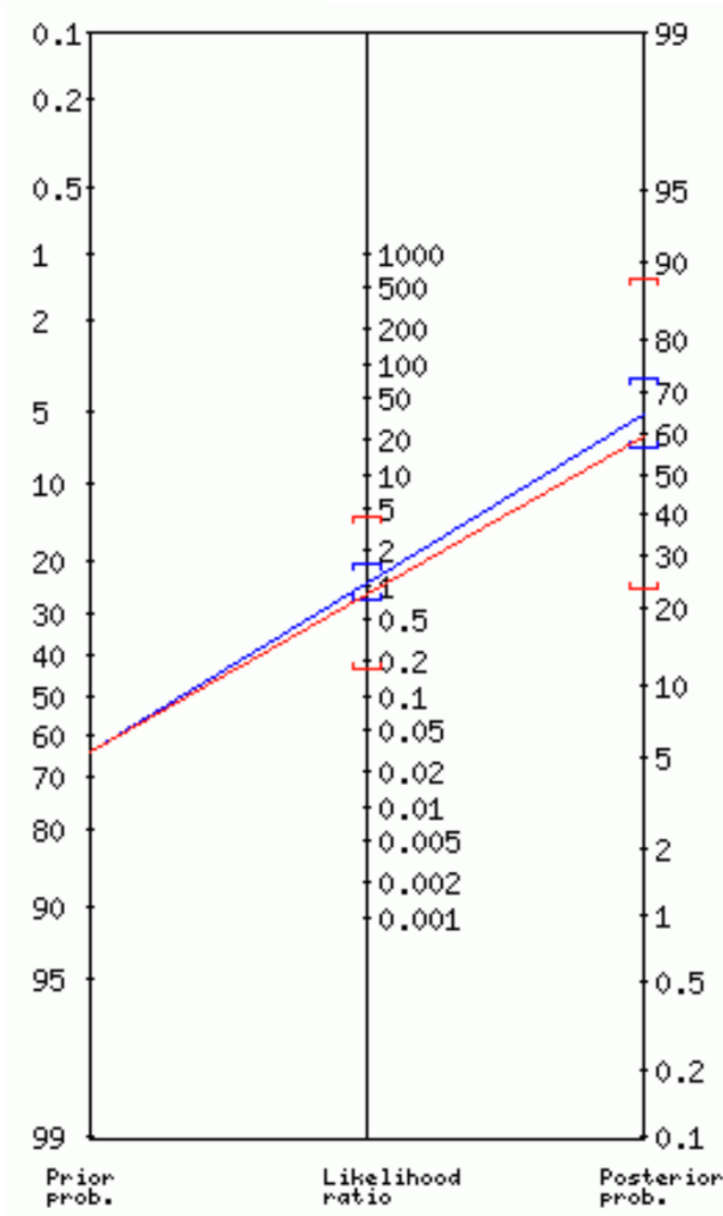


Figura 22. Nomograma de Nolan del modelo predictivo con mejor rendimiento para predecir puntuación en RMQ > 4 al final del seguimiento.

Resultados

7 Discusión

Discusión

7.1 *Hallazgos principales*

Esta Tesis Doctoral se diseñó para evaluar el papel que podría desempeñar el análisis radiómico a la hora de determinar el pronóstico de los pacientes con lumbalgia crónica inespecífica y si dicho análisis podría mejorar el rendimiento de la lectura convencional de la resonancia magnética del raquis lumbar en esta patología tan prevalente.

Como hallazgos principales, los valores de diversos parámetros de textura seleccionados para su estudio (31%) se correlacionaron de forma significativa con el valor de END a los 6 meses, siendo esta asociación leve - moderada (R cercana a 0,4). Además, hasta un 40% de los parámetros de textura mostraron unas diferencias significativas en la distribución de sus valores entre el grupo de pacientes que no mejoraban su dolor y los que sí. Los parámetros de textura del disco intervertebral son los que mostraron mayor asociación con la evolución del dolor: un 43% de los parámetros de disco analizados mostraron asociación significativa frente al 37,5% de los parámetros de platillo vertebral y el 22% de los parámetros calculados en músculo. Algo similar se observó al analizar la distribución de los valores entre los dos grupos de pacientes con mejoría o sin ella: de los parámetros con diferencia significativa en su distribución, 42% eran de disco, 32% de musculatura y 26% de platillo vertebral. Estos hallazgos nos permitirían afirmar que el análisis de textura ofrece mejor correlación con el valor de END a los 6 meses, después del tratamiento rehabilitador, que el análisis convencional, cualitativo, de las imágenes de resonancia magnética.

Basándonos en los resultados obtenidos mediante el algoritmo de inteligencia artificial Random Forest de clasificación, el análisis de textura de las distintas estructuras vertebrales permitiría, junto con los parámetros clínicos básicos incluidos en el estudio, clasificar a los pacientes que no van a mejorar de su dolor tras 6 meses de tratamiento conservador, con una exactitud (accuracy) del 79%, una sensibilidad del 86% y una especificidad del 57%.

En base a estos resultados podríamos decir que el modelo ensayado en este estudio tendría buena fiabilidad en el caso de que el test determine mejoría del paciente (test negativo a no mejoría), en cambio tendría menos valor en caso de que el resultado del test determinase no recuperación (test positivo a no mejoría). En la muestra estudiada la probabilidad pre-test de no mejorar sería del 75% y pasaría a ser del 86% en caso de obtener un

Discusión

resultado positivo en el modelo, bajando al 43% en el caso de que el modelo arrojarase un resultado negativo. Estos datos justifican su inclusión en la práctica clínica.

Sin embargo, los distintos análisis realizados para establecer un pronóstico de la discapacidad del paciente según la escala RMQ no han mostrado resultados favorables. No se ha conseguido demostrar correlaciones significativas de las variables de textura y la discapacidad del paciente a los 6 meses. No se obtuvo ningún modelo predictivo suficientemente capaz de identificar a los pacientes que van a presentar una discapacidad residual ($RMQ > 4$) ni clasificar a aquellos pacientes sin discapacidad. Los modelos de regresión planteados tampoco han permitido predecir los valores de RMQ a los 6 meses ni la proporción de variación de los mismos.

En esta Tesis Doctoral también se ha estudiado la correlación del dolor y la discapacidad con los parámetros radiológicos cualitativos convencionales que se pueden obtener a partir del análisis visual de las imágenes de RM de la columna lumbar. En este estudio no se ha podido demostrar ninguna correlación significativa entre el dolor y la discapacidad del paciente en la visita inicial con el grado de degeneración del disco intervertebral, presencia de zonas de alta intensidad en el mismo, presencia de cambios de Modic en los platillos vertebrales o el grado de degeneración grasa de la musculatura retrosomática a nivel lumbar. También se analizó la posible correlación de estos parámetros cualitativos con el dolor y discapacidad a los 6 meses; solamente se encontró una correlación significativa leve con el grado de degeneración discal a nivel L4-L5, siendo el grado de degeneración mayor en los pacientes que no mejoraban significativamente el dolor que en los que si mejoraban el 30% o más su puntuación END. No se pudo demostrar ninguna correlación de estos parámetros cualitativos con la discapacidad (RMQ) a los 6 meses. No se pudieron plantear modelos de regresión que permitieron construir una ecuación predictiva que explicara el dolor y la discapacidad al final del seguimiento a los 6 meses. Estos resultados nos confirman que la correlación entre la clínica y los hallazgos patológicos obtenidos de la evaluación convencional de las imágenes de RM de la columna lumbar es muy baja.

En base a los resultados obtenidos, el análisis de textura de las imágenes de RM lumbar mejora la capacidad predictiva de la evolución del dolor lumbar inespecífico que ofrece el análisis convencional de dichas

imágenes, sin embargo, estos resultados no se han reproducido en la predicción de la discapacidad. Este estudio no ha permitido obtener un modelo que permita clasificar a los pacientes cuya discapacidad va a persistir tras el tratamiento conservador. Es difícil interpretar esta disparidad de los resultados a la hora de predecir la evolución del dolor y la discapacidad. La discapacidad es una variable multidimensional que describe el grado de limitación de un paciente en la realización de actividades ordinarias. Estudios previos informan que la correlación entre dolor y discapacidad es leve-moderada en pacientes con dolor lumbar (94). Algunos pacientes con mucho dolor apenas están discapacitados, mientras otros con dolor leve presentan discapacidad severa. En el presente estudio se calculó la correlación entre la puntuación en END y puntuación en RMQ a lo 6 meses en nuestra muestra y se observó que, aunque existía correlación estadísticamente significativa, su intensidad era moderada-baja ($R = 0,460$; $p < 0,001$), en línea con otros trabajos previamente publicados (94). Una posible interpretación podría ser que la discapacidad estuviera más relacionada con variables psicosociales (depresión, conductas de catastrofización, creencias y actitudes sobre temor-evitación, aspectos laborales), que no se incluyeron entre las variables de este estudio, y la evolución del dolor estuviera más influenciada por factores estructurales y fisiológicos de los elementos vertebrales que representan las variables de textura.

Estudios recientes han caracterizado la degeneración discal desde el punto de vista histológico, estableciendo una gradación de dicha degeneración en base a la reducción del contenido de glucosaminoglicanos en el núcleo pulposo, microestructura de los platillos vertebrales y la microestructura del anillo fibroso (95). La teoría mecánica de la degeneración discal pone en el centro del proceso al platillo vertebral: la lesión del mismo compromete la nutrición del disco intervertebral y precipita la degradación de su matriz. Se trataría de microfracturas por fatiga debidas traumatismos mínimos o sobrecargas que pasan desapercibidas y que no serían detectables por las técnicas de imagen con convencionales. Tras la lesión del platillo vertebral se inicia la degeneración discal con una reducción del contenido en agua y proteoglicanos, la presión del núcleo pulposo se reduce y las capas internas del anillo fibroso se delaminan (96). Los cambios en la inervación del disco intervertebral también han sido motivo de estudio dentro de la fisiopatología del dolor lumbar. El disco intervertebral normal está escasamente inervado lo que contrasta con el disco degenerado que está densamente inervado. Esta proliferación neural está mediada por

Discusión

neurotrofinas, que además juegan un papel importante en la regulación respuesta inflamatoria y la transmisión del dolor, aumentando la expresión de péptidos relacionados con el dolor y modulando las sinapsis de las neuronas nociceptivas en la médula espinal (17).

Estos cambios a nivel tisular probablemente se expresen en cambios en la imagen de RM, que podrían ser indetectables al ojo humano, pero ser los que permiten al análisis de textura mejorar su rendimiento pronóstico frente a la imagen radiológica convencional.

Dado que en la actualidad existen pocas estrategias para determinar el pronóstico a medio plazo de los pacientes con dolor lumbar crónico, el análisis de textura aporta una información valiosa que permite orientar el manejo terapéutico del paciente, lo cual es de gran importancia en una patología tan prevalente y que tantos recursos consume en el sistema sanitario.

7.2 Comparación con otros estudios

Diversos estudios han analizado los factores de mal pronóstico o de riesgo de cronificación del dolor lumbar, destacando la influencia de los factores psicosociales como limitantes de la recuperación de los pacientes (97, 98). Existen pocos estudios cuyo objetivo sea analizar las alteraciones estructurales en imágenes de RM para determinar el pronóstico del dolor lumbar inespecífico tras el tratamiento conservador, lo que contrasta con la existencia de múltiples trabajos recientes que proponen modelos para predecir el resultado del tratamiento quirúrgico o la posibilidad de complicaciones post-operatorias (99, 100).

Además del análisis de textura de los distintos elementos vertebrales, en este trabajo se ha realizado un análisis de la capacidad predictiva de la RM convencional sobre el dolor y la discapacidad a los 6 meses. El grado de degeneración discal medido con la gradación visual de Pfirrmann, la presencia de cambios de Modic, la presencia de ZAI o el grado de degeneración grasa de la musculatura no son capaces de determinar el pronóstico de los pacientes con dolor lumbar a medio plazo. Estudios recientes sobre la correlación entre el análisis convencional de las imágenes de RM del raquis lumbar con la presencia de dolor crónico ofrecen resultados similares a los obtenidos en esta Tesis Doctoral, no encontrando correlación entre la presencia de cambios degenerativos y la intensidad del dolor o la presencia de dolor neuropático (78, 98, 99). Otros, sin embargo, encuentran una mayor prevalencia de dolor lumbar en pacientes con mayor grado de degeneración discal aunque esta correlación se pierde en pacientes con “distrés mental” (103).

Wang *y cols*, en un estudio retrospectivo, desarrollaron un modelo predictivo de la necesidad de tratamiento quirúrgico en pacientes con dolor lumbar comparando el AUC de un modelo que incluía variables clínicas con otro que además incluía la RM, pasando el AUC de 0,75 a 0,77 al incluir la RM entre las variables analizadas (104). Utilizando un método de clusters o agrupaciones no supervisado, Jim *y cols* se basaron en la intensidad de señal del disco en la imagen de RM y la altura del mismo para clasificar a los pacientes en tres agrupaciones; en uno de estos grupos el porcentaje de pacientes que desarrollaban dolor lumbar era mayor (105). En un estudio transversal, Shuim *y cols* evaluaron la capacidad predictiva de diversos algoritmos de inteligencia artificial a la hora de identificar a pacientes con riesgo de desarrollar dolor lumbar crónico (106). En estos modelos incluyeron 13 variables, demográficas y clínicas, sin incluir

Discusión

variables radiológicas. El AUC obtenida con el Random Forest fue de 0,660, inferior a la obtenida con nuestro modelo, y el modelo que obtuvo mejor capacidad predictiva fue el calculado con el algoritmo naïve Bayes, con AUC de 0,712, igual a nuestro modelo.

Liew *y cols* demostraron mejor capacidad predictiva de diversos modelos obtenidos mediante algoritmos de inteligencia artificial frente a modelos de regresión convencionales, aplicados a la evolución de pacientes con dolor cervical. Estos modelos incluían variables de imagen convencional, además de las variables sociodemográficas y clínicas. El mejor resultado se obtuvo con el algoritmo Xgboost, con AUC = 0.726 para predecir la persistencia del dolor cervical a los 3 meses (107).

Varios estudios han aplicado el análisis de textura de las imágenes de RM de la columna lumbar para clasificar a pacientes con dolor lumbar. Abdollah *y cols* encontraron diferencias entre los pacientes con lumbalgia o sin ella en los parámetros homogeneidad y energía a nivel del disco intervertebral L5-S1 así como en contraste, homogeneidad y energía de los platillos vertebrales adyacentes (76). Ketola *y cols* obtuvieron un modelo basado en los parámetros de textura calculados a nivel L4-L5 y L5-S1 con un AUC de 0,91 para distinguir pacientes con lumbalgia de individuos asintomáticos, siendo mayor que si se utilizaban los discos lumbares superiores (AUC = 0,76) o si se utilizaba la clasificación de Pfirmann (AUC = 0,62)(77). Eriksson *y cols* aplicaron el análisis de textura para identificar fisuras del anillo fibroso, en especial aquellas que no se apreciaban a simple vista en la imagen de RM, obteniendo una AUC de 0,97, siendo su valor entre 0,69 y 0,75 para la predicción del dolor, mostrando alta sensibilidad (0,93) y baja especificidad (0,4) (108). Resultados similares han publicado Lagerstrand *y cols*, con una precisión del 71% para predecir el dolor (109). Aunque estos valores son similares a los obtenidos en nuestro estudio no aportan un valor de predicción a medio plazo de la evolución del dolor, solamente correlacionan los hallazgos del análisis de textura con el estado actual del paciente.

La correlación de los cambios degenerativos en la RM lumbar con la discapacidad también ha sido motivo de publicaciones recientes. Gebrewold *y cols* concluyen que no existe correlación entre la presencia de protrusión o hernia discal, estenosis foraminal o cambios de Modic, encontrando correlación débil con la presencia de estenosis de canal lumbar (110); Dunsmuir *y cols* concluyen que no existe correlación entre la discapacidad y el tamaño de la hernia discal (102). Banitalebi *y cols* no

encontraron correlación entre el grado de degeneración grasa de la musculatura paravertebral lumbar y el grado de discapacidad (111).

No se han encontrado estudios que analicen el valor pronóstico del análisis de textura en pacientes con dolor lumbar inespecífico por lo que no es posible comparar nuestros resultados con estudios similares. Si que podemos compararlos con estudios pronósticos basados en variables clínicas y alguno que incorpora imágenes de RM convencional. Una revisión sistemática y metaanálisis de reciente en la que se incluyeron estudios prospectivos en los que se evaluaba la capacidad predictiva de distintas escalas concluyó que la capacidad predictiva del dolor era baja, con AUC de 0,63 y aceptable para la discapacidad con AUC de 0,71 (43). Los mismos autores realizaron un estudio prospectivo en el que se evaluó la capacidad predictiva sobre el dolor lumbar inespecífico de tres cuestionarios. Encontraron una baja capacidad para predecir el mal pronóstico, con AUC de 0,61 para PICKUP tool, 0,66 para Short Form Orebro Musculoskeletal Pain Screening Questionnaire y 0,69 para START back tool (42). Otro estudio prospectivo obtuvo un resultado más favorable con un modelo basado en variables clínicas para predecir buen pronóstico (AUC = 0,77), sin diferencia significativa al incorporar los hallazgos de la imagen de RM vertebral (AUC = 0,78) (46).

Los modelos predictivos basados en variables clínicas requieren la cumplimentación de cuestionarios complejos lo que dificulta su aplicación de forma rutinaria en la clínica. Los modelos estudiados en esta Tesis Doctoral se basan en variables sociodemográficas y clínicas que se incluyen habitualmente en la historia clínica básica. Por otra parte, el análisis de las imágenes de RM lumbar, segmentación y análisis de textura se puede automatizar utilizando algoritmos inteligencia artificial. Esto facilitaría su implementación de forma rutinaria en el estudio de los pacientes con dolor lumbar inespecífico, aportándonos valiosa información sobre el pronóstico del paciente y ayudarnos a establecer un plan de tratamiento más personalizado.

Uno de los principales problemas que se encuentran en los estudios de predicción es la falta de validación de los modelos (112). Una de las fortalezas de este estudio es que se ha llevado a cabo un entrenamiento del método en el 70% de la muestra y posteriormente se ha realizado una validación interna con el resto de la misma.

7.3 Aplicaciones clínicas de los resultados

En la actualidad, existe incertidumbre sobre el pronóstico del paciente con dolor lumbar inespecífico y sobre la respuesta al tratamiento conservador. Las alteraciones estructurales cualitativas observadas en técnicas de imagen convencional (RX, TC y RM) no son capaces de predecir que pacientes mejoraran tras el tratamiento conservador. Dada la prevalencia elevada del dolor lumbar inespecífico, sus consecuencias sobre la calidad de vida de las personas y la repercusión sobre el sistema sanitario es de gran importancia conocer cuál va a ser el pronóstico de los pacientes para, de esta forma, establecer un plan terapéutico personalizado y orientar los recursos disponibles de la forma más eficiente posible. Como ya se ha comentado, existen pocos estudios sobre modelos predictivos del dolor lumbar.

Los resultados obtenidos en esta Tesis Doctoral abren una nueva línea de trabajo que no se había investigado hasta el momento actual. Con el modelo predictivo que incluye las variables sociodemográficas básicas, la puntuación END y RMQ y las variables de textura podríamos estimar cual va a ser la evolución del dolor a medio plazo, identificando a los casos en los que la respuesta favorable al tratamiento rehabilitador fuera más probable y así priorizar el tratamiento de los mismos, estableciendo una estrategia terapéutica individualizada. Antes de su implementación en la práctica clínica será necesaria la validación externa del modelo para confirmar su seguridad y eficacia.

7.4 *Limitaciones del estudio*

Esta Tesis Doctoral presenta algunas limitaciones que se analizan a continuación.

Existe un sesgo de selección de la muestra ya que se incluyen pacientes que son remitidos a una unidad especializada a nivel hospitalario lo que implica que muchos de ellos ya han recibido tratamientos previos que en la mayor parte de los casos han sido inefectivos; probablemente este sea el motivo por el que la edad media de los pacientes se sitúe alrededor de 50 años. Los pacientes en esta muestra presentaban un dolor crónico de larga duración. Es posible que los resultados de este estudio no sean aplicables a pacientes con dolor lumbar agudo o subagudo.

Otro sesgo a considerar es el tipo de tratamiento recibido por los pacientes: aunque el tratamiento rehabilitador a nivel del hospital fue homogéneo, no se han controlado otros tratamientos que el paciente pudiera haber recibido por su cuenta, fisioterápicos o farmacológicos. Este sesgo tendría una importancia relativa ya que estas variaciones de tratamiento no controladas formarían parte del manejo y evolución habitual de estos pacientes.

El tamaño muestral podría considerarse insuficiente para el número de variables incluidas en los modelos, aunque la mitad de los estudios predictivos aplicando inteligencia artificial publicados sobre patología del raquis incluyen menos de 100 pacientes (113). Esto podría originar lo que se conoce como overfitting o sobreajuste del modelo: al aplicar el modelo fuera de la muestra podría empeorar su rendimiento.

Aunque una fortaleza del estudio es la inclusión de variables clínicas sencillas y reproducibles no se incluyeron variables del ámbito psicosocial, que parecen tener un alto peso específico en la evolución del dolor lumbar (98).

La segmentación de las ROI se realizó de forma manual, lo que podría reducir su precisión intra e inter-observador. Para reducir dicho sesgo se realizó el proceso dos veces y se promediaron los resultados.

7.5 Líneas futuras de investigación

A partir del modelo predictivo propuesto, sería imprescindible una validación externa de los resultados, preferiblemente mediante un estudio multicéntrico, previo a considerar su implementación en la clínica. Otra línea de investigación sería estudiar la aplicación del mismo método en pacientes con lumbalgia aguda, comprobando si este modelo predictivo es válido y en caso de que no lo fuera, formular nuevos modelos.

La automatización del proceso de identificación de las distintas estructuras vertebrales en la imagen de RM y la segmentación de las ROI utilizando procesos de inteligencia artificial es un campo en continuo desarrollo en los últimos años. Los estudios más recientes han aplicado procedimientos como redes neuronales artificiales, redes neuronales convolucionales y aprendizaje profundo para analizar de forma automática las imágenes de RM de la columna vertebral. Existen en la actualidad herramientas de software cuya capacidad para identificar y etiquetar las distintas estructuras vertebrales superan la capacidad del ojo humano. Estos mismos procedimientos se están desarrollando para realizar la segmentación automática de las ROI con precisión máxima: las redes neuronales convolucionales se han aplicado con este fin obteniendo una precisión a nivel vóxel. Aunque en el campo de la segmentación de las estructuras vertebrales se han obtenido resultados prometedores, sigue habiendo un campo de investigación abierto y con muchas posibilidades de mejora. La aplicación de la segmentación automática a nuestro modelo podría mejorar su capacidad para establecer correctamente el pronóstico de los pacientes. Por otra parte, la automatización del proceso facilitaría su implementación en la práctica clínica diaria; el informe radiológico convencional de la imagen de RM podría completarse con una información pronóstica, que sería de gran valor para el clínico.

Estos modelos predictivos basados en la inteligencia artificial suponen una herramienta más de ayuda a la toma de decisiones, pero no sustituirán nunca a la capacidad de decisión del médico, que deberá actuar en base al contexto clínico singular de cada paciente.

8 Conclusiones

- 1- Diversos parámetros de textura calculados a partir de las imágenes de RM obtenidas con secuencias potenciadas en T2 en el disco intervertebral, los platillos vertebrales y la musculatura paravertebral lumbar han demostrado tener una correlación estadísticamente significativa con la evolución del dolor lumbar crónico a los 6 meses de la evaluación clínica inicial tras el tratamiento rehabilitador.
- 2- El análisis de textura aplicado a las imágenes de RM de la columna lumbar, junto con las variables clínicas iniciales, permite hacer una predicción de la evolución del dolor del paciente a medio plazo, e identificar a los pacientes que van a tener una mejoría clínicamente relevante del mismo a los 6 meses tras un tratamiento rehabilitador.
- 3- El análisis de textura aplicado a las imágenes de RM de la columna lumbar, junto con las variables clínicas iniciales, no permite obtener un modelo predictivo válido para determinar la evolución de la discapacidad del paciente y conocer el grado de discapacidad a los 6 meses.
- 4- El análisis convencional de las imágenes de RM de la columna lumbar (grado de degeneración discal, presencia de ZAI, cambios de Modic y degeneración grasa de la musculatura paravertebral lumbar) no han mostrado una correlación con el grado de dolor y la discapacidad que presenta el paciente en la visita inicial ni a los 6 meses de evolución. No se ha podido obtener ningún modelo predictivo válido con estos parámetros.
- 5- El análisis de textura de las imágenes de RM potenciadas en T2 mejora la capacidad predictiva de la evolución del dolor lumbar crónico respecto a las variables clínicas y el análisis convencional de estas imágenes.
- 6- El modelo predictivo estudiado para conocer la evolución del dolor lumbar crónico permite obtener información complementaria al informe radiológico convencional que puede ser tomada en cuenta a la hora de establecer un plan terapéutico personalizado en cada paciente

9 Bibliografía

Bibliografía

1. Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. *Lancet*. 2012; 379(9814): 482–91
2. Savigny P, Watson P, Underwood M. Guidelines - Early management of persistent non-specific low back pain: Summary of NICE guidance. Vol. 338, *BMJ (Online)*. British Medical Journal Publishing Group; 2009: 1441-2
3. Deyo RA, Weinstein JN. Low Back Pain. *N Engl J Med*. 2001 Feb 20; 344(5): 363-70
4. Downie A, Williams CM, Henschke N, Hancock MJ, Ostelo RWJG, De Vet HCW, et al. Red flags to screen for malignancy and fracture in patients with low back pain: Systematic review. *BMJ*. 2013; 347
5. Waddell G. 1987 volvo award in clinical sciences: A new clinical model for the treatment of low-back pain. *Spine (Phila Pa 1976)*. 1987; 12(7): 632-44.
6. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. *Lancet*. 2017;3 89(10070): 736–47.
7. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4 European guidelines for the management of chronic nonspecific low back pain On behalf of the COST B13 Working Group on Guidelines for Chronic Low Back Pain Contributors. *Eur Spine J*. 2006;192-300.
8. Chou R, Deyo R, Friedly J, et al. Noninvasive Treatments for Low Back Pain. Rockville, MD: Agency for Healthcare Research and Quality(US); 2016. Report No.: 16-EHC004-EF
9. Popescu A, Lee H. Neck Pain and Lower Back Pain. *Med Clin North Am*. 2020; 104(2): 279-92.
10. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015; 386(9995): 743-800.
11. Humbría Mendiola A, Carmona L, Ortiza A, Peña Sagredo J. Impacto poblacional del dolor lumbar en España: resultados del estudio EPISER. *Rev Española Reumatol*. 2002; 29(10): 471-8.
12. Viedma Gil de Vergara P. Encuesta de Salud de la Comunidad Valenciana 2010. Depósito Legal. 2012.
13. Alonso-García M, Sarría-Santamera A. The Economic and Social Burden of Low Back Pain in Spain: A National Assessment of the

Bibliografía

- Economic and Social Impact of Low Back Pain in Spain. *Spine (Phila Pa 1976)*. 2020; 45(16): E1026-32.
14. Soriano JB, Rojas-Rueda D, Alonso J, Antó JM, Cardona PJ, Fernández E, et al. The burden of disease in Spain: Results from the Global Burden of Disease 2016. *Med Clin (Barc)*. 2018; 151(5): 171-190.
 15. Rodríguez Rieiro C. Utilidad de la Resonancia Magnética en pacientes con dolor lumbar inespecífico. Ministerio de Sanidad, Servicios Sociales e Igualdad. Unidad de Evaluación de Tecnologías Sanitarias de la Comunidad de Madrid. 2013. Informes de Evaluación de Tecnologías Sanitarias: 1-54 https://www.seram.es/images/site/utilidad_de_la_rnm_lumbar.pdf.
 16. Ract I, Meadeb JM, Mercy G, et al. A review of the value of MRI signs in low back pain. *Diagnostic and Interventional Imaging*. 2015; 96(3): 239-49.
 17. García-Cosamalón J, del Valle ME, Calavia MG, et al. Intervertebral disc, sensory nerves and neurotrophins: who is who in discogenic pain? *J Anat*. 2010; 217(1): 1-15.
 18. Pfirrmann CWA, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine (Phila Pa 1976)*. 2001; 26(17): 1873-8.
 19. Urrutia J, Besa P, Campos M, Cikutovic P, Cabezon M, Molina M, et al. The Pfirrmann classification of lumbar intervertebral disc degeneration: an independent inter- and intra-observer agreement assessment. *Eur Spine J*. 2016; 25(9): 2728-33.
 20. Peng B, Hou S, Wu W, Zhang C, Yang Y. The pathogenesis and clinical significance of a high-intensity zone (HIZ) of lumbar intervertebral disc on MR imaging in the patient with discogenic low back pain. *Eur Spine J*. 2006; 15(5): 583-7.
 21. Teraguchi M, Samartzis D, Hashizume H, Yamada H, Muraki S, Oka H, et al. Classification of High Intensity Zones of the Lumbar Spine and Their Association with Other Spinal MRI Phenotypes: The Wakayama Spine Study. *PLoS One*. 2016; 11(9): e0160111.
 22. Fardon DF, Williams AL, Dohring EJ, Murtagh FR, Gabriel Rothman SL, Sze GK. Lumbar disc nomenclature: Version 2.0 Recommendations of the combined task forces of the North American Spine Society, the American Society of Spine Radiology and the American Society of Neuroradiology. *Spine J*. 2014; 14(11): 2525-45.
 23. Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR.

- Degenerative disk disease: Assessment of changes in vertebral body marrow with MR imaging. *Radiology*. 1988; 166(1 I):193-9.
24. Perilli E, Parkinson IH, Truong L-H, Chong KC, Fazzalari NL, Osti OL. Modic (endplate) changes in the lumbar spine: bone micro-architecture and remodelling. *Eur Spine J*. 2015; 24(9): 1926-34.
 25. Rahme R, Moussa R. The modic vertebral endplate and marrow changes: pathologic significance and relation to low back pain and segmental instability of the lumbar spine. *AJNR Am J Neuroradiol*. 2008; 29(5): 838-42.
 26. Hildebrandt M, Fankhauser G, Meichtry A, Luomajoki H. Correlation between lumbar dysfunction and fat infiltration in lumbar multifidus muscles in patients with low back pain. *BMC Musculoskelet Disord*. 2017; 18(1): 12.
 27. Weishaupt D, Zanetti M, Hodler J, Boos N. MR imaging of the lumbar spine: Prevalence of intervertebral disk extrusion and sequestration, nerve root compression, end plate abnormalities, and osteoarthritis of the facet joints in asymptomatic volunteers. *Radiology*. 1998; 209(3): 661-6.
 28. Stadnik TW, Lee RR, Coen HL, Neiryneck EC, Buisseret TS, Osteaux MJC. Annular tears and disk herniation: Prevalence and contrast enhancement on MR images in the absence of low back pain or sciatica. *Radiology*. 1998; 206(1): 49-55.
 29. Tonosu J, Oka H, Higashikawa A, Okazaki H, Tanaka S, Matsudaira K. The associations between magnetic resonance imaging findings and low back pain: A 10-year longitudinal analysis. *PLoS One*. 2017; 12(11): e0188057.
 30. Suri P, Boyko EJ, Goldberg J, Forsberg CW, Jarvik JG. Longitudinal associations between incident lumbar spine MRI findings and chronic low back pain or radicular symptoms: retrospective analysis of data from the longitudinal assessment of imaging and disability of the back (LAIDBACK). *BMC Musculoskelet Disord*. 2014;15:152.
 31. Brinjikji W, Diehn FE, Jarvik JG, Carr CM, Kallmes DF, Murad MH, et al. MRI Findings of Disc Degeneration are More Prevalent in Adults with Low Back Pain than in Asymptomatic Controls: A Systematic Review and Meta-Analysis. *AJNR Am J Neuroradiol*. 2015; 36(12): 2394-9.
 32. Berg AJ, Ahmadje U, Jayanna HH, Trégouët P, Sanville P, Kapoor V. The prevalence of lumbar disc degeneration in symptomatic younger patients: A study of MRI scans. *J Clin Orthop Trauma*.

Bibliográfia

- 2020; 11(5): 932-6.
33. Vagaska E, Litavcova A, Srotova I, Vlckova E, Kerkovsky M, Jarkovsky J, et al. Do lumbar magnetic resonance imaging changes predict neuropathic pain in patients with chronic non-specific low back pain? *Medicine (Baltimore)*. 2019; 98(17): e15377.
 34. Sääksjärvi S, Kerttula L, Luoma K, Paajanen H, Waris E. Disc Degeneration of Young Low Back Pain Patients: A Prospective 30-year Follow-up MRI Study. *Spine (Phila Pa 1976)*. 2020; 45(19): 1341-7.
 35. Rahyussalim AJ, Zufar MLL, Kurniawati T. Significance of the Association between Disc Degeneration Changes on Imaging and Low Back Pain: A Review Article. *Asian Spine J*. 2020; 14(2): 245-57.
 36. Niemeläinen R, Videman T, Dhillon SS, Battié MC. Quantitative measurement of intervertebral disc signal using MRI. *Clin Radiol*. 2008; 63(3): 252-5.
 37. Salamat S, Hutchings J, Kwong C, Magnussen J, Hancock MJ. The relationship between quantitative measures of disc height and disc signal intensity with Pfirrmann score of disc degeneration. *Springerplus*. 2016; 5(1): 829.
 38. Pandit P, Talbott JF, Padoia V, Dillon W, Majumdar S. T1 ρ and T2-based characterization of regional variations in intervertebral discs to detect early degenerative changes. *J Orthop Res*. 2016; 34(8): 1373-81.
 39. Menezes-Reis R, Salmon CEG, Carvalho CS, Bonugli GP, Chung CB, Nogueira-Barbosa MH. T1 ρ and T2 mapping of the intervertebral disk: comparison of different methods of segmentation. *AJNR Am J Neuroradiol*. 2015; 36(3): 606-11.
 40. Messner A, Stelzeneder D, Trattng S, Götz, Welsch H, Schinhan M, et al. Does T2 mapping of the posterior annulus fibrosus indicate the presence of lumbar intervertebral disc herniation? A 3.0 Tesla magnetic resonance study. *Eur Spine J*. 2017; 26(3): 877-83.
 41. Jiang Y, Yu L, Luo X, Lin Y, He B, Wu B, et al. Quantitative synthetic MRI for evaluation of the lumbar intervertebral disk degeneration in patients with chronic low back pain. *Eur J Radiol*. 2020; 124.
 42. Karran EL, Traeger AC, McAuley JH, Hillier SL, Yau YH, Moseley GL. The Value of Prognostic Screening for Patients With Low Back Pain in Secondary Care. *J Pain*. 2017; 18(6): 673-86.
 43. Karran EL, McAuley JH, Traeger AC, Hillier SL, Grabherr L,

- Russek LN, et al. Can screening instruments accurately determine poor outcome risk in adults with recent onset low back pain? A systematic review and meta-analysis. *BMC Med.* 2017; 15(1): 673-86.
44. Unsgaard-Tøndel M, Kregnes IG, Nilsen TIL, Marchand GH, Askim T. Risk classification of patients referred to secondary care for low back pain. *BMC Musculoskelet Disord.* 2018; 19(1): 166.
 45. Steffens D, Hancock MJ, Maher CG, Williams C, Jensen TS, Latimer J. Does magnetic resonance imaging predict future low back pain? A systematic review. *Eur J Pain.* 2014; 18(6): 755-65.
 46. de Schepper EIT, Koes BW, Oei EHG, Bierma-Zeinstra SMA, Luijsterburg PAJ. The added prognostic value of MRI findings for recovery in patients with low back pain in primary care: a 1-year follow-up cohort study. *Eur Spine J.* 2016; 25(4): 1234-41.
 47. Parekh V, Jacobs MA. Radiomics: a new application from established techniques. *Expert Rev Precis Med drug Dev.* 2016; 1(2): 207-26.
 48. Marti-Bonmati L. Radiomics and imaging biomarkers in observational clinical studies with retrospective data. *An RANM.* 2019; 136: 34-42.
 49. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology.* 2016; 278(2): 563-77.
 50. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp.* 2018; 2(1): 36.
 51. Beers B, Leporq B, Doblaz S, Garteiser P. Imaging Biomarker Measurements. In: *Imaging Biomarkers: Development and Clinical Integration.* Springer. 2017. 87-99.
 52. Strzelecki M, Szczypinski P, Materka A, Klepaczko A. A software tool for automatic classification and segmentation of 2D/3D medical images. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip.* 2013; 702: 137-40.
 53. Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017; 77(21):e104–7.
 54. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. *Sci Rep.* 2015; 5: 13087.
 55. Chitalia RD, Kontos D. Role of texture analysis in breast MRI as a cancer biomarker: A review. *J Magn Reson Imaging.* 2019; 49(4):

Bibliografía

- 927-38.
56. Stelzer PD, Steding O, Raudner MW, Euller G, Clauser P, Baltzer PAT. Combined texture analysis and machine learning in suspicious calcifications detected by mammography: Potential to avoid unnecessary stereotactical biopsies. *Eur J Radiol.* 2020; 132.
 57. Nadrljanski MM, Milosevic ZC. Tumor texture parameters of invasive ductal breast carcinoma in neoadjuvant chemotherapy: early identification of non-responders on breast MRI. *Clin Imaging.* 2020; 65: 119-23.
 58. Skogen K, Schulz A, Dormagen JB, Ganeshan B, Helseth E, Server A. Diagnostic performance of texture analysis on MRI in grading cerebral gliomas. *Eur J Radiol.* 2016; 85(4): 824-9.
 59. Ditmer A, Zhang B, Shujaat T, Pavlina A, Luibrand N, Gaskill-Shipley M, et al. Diagnostic accuracy of MRI texture analysis for grading gliomas. *J Neurooncol.* 2018; 140(3): 583-9.
 60. Tian Q, Yan LF, Zhang X, et al. Radiomics strategy for glioma grading using texture features from multiparametric MRI. *J Magn Reson Imaging.* 2018; 48(6): 1518-28.
 61. Masokano IB, Liu W, Xie S, Marcellin DFH, Pei Y, Li W. The application of texture quantification in hepatocellular carcinoma using CT and MRI: a review of perspectives and challenges. *Cancer Imaging.* 2020;20(1):67.
 62. Lewis S, Hectors S, Taouli B. Radiomics of hepatocellular carcinoma. *Abdom Radiol.* 2021; 46(1): 111-23.
 63. Liu X, Khalvati F, Namdar K, Fischer S, Lewis S, Taouli B, et al. Can machine learning radiomics provide pre-operative differentiation of combined hepatocellular cholangiocarcinoma from hepatocellular carcinoma and cholangiocarcinoma to inform optimal treatment planning? *Eur Radiol.* 2021; 31(1): 244-55.
 64. Crombé A, Marcellin PJ, Buy X, Stoeckle E, Brouste V, Italiano A, et al. Soft-tissue sarcomas: Assessment of MRI features correlating with histologic grade and patient outcome. *Radiology.* 2019; 291(3): 710-21.
 65. Hong JH, Jee WH, Jung CK, Chung YG. Tumor grade in soft-tissue sarcoma: Prediction with magnetic resonance imaging texture analysis. *Medicine (Baltimore).* 2020; 99(27): e20880.
 66. Pereira RR. Texture Analysis Shows Promise in Differentiating Pancreatic Neoplasms. *Acad Radiol.* 2020 ;27(6):824.
 67. Li X, Zhu H, Qian X, Chen N, Lin X. MRI Texture Analysis for Differentiating Nonfunctional Pancreatic Neuroendocrine

- Neoplasms From Solid Pseudopapillary Neoplasms of the Pancreas. *Acad Radiol.* 2020; 27(6):815-823
68. Cai JH, He Y, Zhong XL, Lei H, Wang F, Luo GH, et al. Magnetic Resonance Texture Analysis in Alzheimer's disease. Vol. 27, *Academic Radiology.* Elsevier Inc.; 2020. p. 1774-83.
 69. Mannil M, Eberhard M, von Spiczak J, Heindel W, Alkadhi H, Baessler B. Artificial Intelligence and Texture Analysis in Cardiac Imaging. *Curr Cardiol Rep.* 2020; 22(11): 131.
 70. Tameem HZ, Selva LE, Sinha US. Texture measure from low resolution MR images to determine trabecular bone integrity in osteoporosis. *Annu Int Conf IEEE Eng Med Biol Soc.* 2007; 2007: 2027-30.
 71. MacKay JW, Murray PJ, Kasmai B, Johnson G, Donell ST, Toms AP. Subchondral bone in osteoarthritis: association between MRI texture analysis and histomorphometry. *Osteoarthr Cartil.* 2017 M; 25(5): 700-7.
 72. Lisson CS, Lisson CG, Flosdorf K, Mayer-Steinacker R, Schultheiss M, von Baer A, et al. Diagnostic value of MRI-based 3D texture analysis for tissue characterisation and discrimination of low-grade chondrosarcoma from enchondroma: a pilot study. *Eur Radiol.* 2018; 28(2): 468-77.
 73. Mayerhoefer ME, Stelzeneder D, Bachbauer W, Welsch GH, Mamisch TC, Szczypinski P, et al. Quantitative analysis of lumbar intervertebral disc abnormalities at 3.0 Tesla: value of T(2) texture features and geometric parameters. *NMR Biomed.* 2012; 25(6): 866-72.
 74. Michopoulou S, Costaridou L, Vlychou M, Speller R, Todd-Pokropek A. Texture-based quantification of lumbar intervertebral disc degeneration from conventional T2-weighted MRI. *Acta radiol.* 2011; 52(7): 91-8.
 75. Huber FA, Stutz S, Vittoria de Martini I, Mannil M, Becker AS, Winklhofer S, et al. Qualitative versus quantitative lumbar spinal stenosis grading by machine learning supported texture analysis. Experience from the LSOS study cohort. *Eur J Radiol.* 2019; 114: 45-50.
 76. Abdollah V, Parent EC, Dolatabadi S, Marr E, Croutze R, Wachowicz K, et al. Texture Analysis in the Classification of T₂ Weighted Magnetic Resonance Images in Persons with and without Low Back Pain. *J Orthop Res.* 2021; 39(10): 2187-2196.
 77. Ketola JHJ, Inkinen SI, Karppinen J, Niinimäki J, Tervonen O,

Bibliografía

- Nieminen MT. T₂-weighted magnetic resonance imaging texture as predictor of low back pain: A texture analysis-based classification pipeline to symptomatic and asymptomatic cases. *J Orthop Res.* 2021; 39(11): 2428-2438.
78. Mannil M, Burgstaller JM, Thanabalasingam A, Winklhofer S, Betz M, Held U, et al. Texture analysis of paraspinal musculature in MRI of the lumbar spine: analysis of the lumbar stenosis outcome study (LSOS) data. *Skeletal Radiol.* 2018; 47(7): 947-54.
79. Dieckmeyer M, Inhuber S, Schlaeger S, Weidlich D, Rama M, Mookiah K, et al. Texture Features of Proton Density Fat Fraction Maps from Chemical Shift Encoding-Based MRI Predict Paraspinal Muscle Strength. *Diagnostics (Basel).* 2021;11(2):2392021.
80. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. <https://doi.org/107326/M14-0697>. 2015 Jan 6; 162(1): 55-63.
81. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. *Ann Rheum Dis.* 1978; 37(4): 378-81.
82. Vicente Herrero MT, Delgado Bueno S, Bandrés Moyá F, Ramírez Iñiguez de la Torre M V, Capdevila García L, Teófila Vicente Herrero M. Valoración del dolor. Revisión comparativa de escalas y cuestionarios. *Rev Soc Esp Dolor.* 2018; 25(4): 228-36.
83. Dworkin RH, Turk DC, Peirce-Sandner S, Baron R, Bellamy N, Burke LB, et al. Research design considerations for confirmatory chronic pain clinical trials: IMMPACT recommendations. *Pain.* 2010; 149(2): 177-193.
84. Kongsted A, Kent P, Axen I, Downie AS, Dunn KM. What have we learned from ten years of trajectory research in low back pain? *BMC Musculoskelet Disord.* 2016; 17(1): 220.
85. Ostelo RWJG, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. *Spine (Phila Pa 1976).* 2008; 33(1): 90-4.
86. Kovacs FM, Llobera J, Gil del Real MT, Abairra V, Gestoso M, Fernández C, et al. Validation of the Spanish version of the Roland-Morris Questionnaire. *Spine (Phila Pa 1976).* 2002; 27(5): 538-42.
87. Roland M, Morris R. The 1982 Volvo award in clinical science: A study of the natural history of back pain: Part I: Development of a

- reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)*. 1983; 8(2): 141-4.
88. Roland M, Fairbank J. The Roland-Morris disability questionnaire and the Oswestry disability questionnaire. *Spine (Phila Pa 1976)* 2000;25(24):3115-24. Erratum in: *Spine (Phila Pa 1976)*. 2001; 26(7): 847.
89. Kovacs FM, Abraira V, Zamora J, Gil Del Real MT, Llobera J, Fernández C, et al. Correlation between Pain, Disability, and Quality of Life in Patients with Common Low Back Pain. *Spine (Phila Pa 1976)*. 2004; 29(2) :206-10.
90. Lao J, Chen Y, Li ZC, Li Q, Zhang J, Liu J, et al. A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. *Sci Rep*. 2017; 7(1): 1-8.
91. Pedregosa F, Michel V, Grisel O, Blondel M, Prettenhofer P, Weiss R, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2011; 12: 2825-30.
92. McKinney W. Data Structures for Statistical Computing in Python. *Proc 9th Python in science conf. (SCIPY 2010)*.
93. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013 ;310(20): 2191-4.
94. Alamam DM, Moloney N, Leaver A, Alsobayel HI, MacKey MG. Pain Intensity and Fear Avoidance Explain Disability Related to Chronic Low Back Pain in a Saudi Arabian Population. *Spine (Phila Pa 1976)*. 2019; 44(15): E889-98.
95. Rutges JPHJ, Duit RA, Kummer JA, Bekkers JEJ, Oner FC, Castelein RM, et al. A validated new histological classification for intervertebral disc degeneration. *Osteoarthr Cartil*. 2013; 21(12): 2039-47.
96. Hadjipavlou AG, Tzermiadianos MN, Bogduk N, Zindrick MR. The pathophysiology of disc degeneration: A critical review. *J Bone Joint Surg Br*. 2008; 90(10): 1261-70.
97. Hayden JA, Wilson MN, Riley RD, Iles R, Pincus T, Ogilvie R. Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. *Cochrane Database Syst Rev*. 2019 Nov 25; 2019(11): CD011284.
98. Otero-Ketterer E, Peñacoba-Puente C, Ferreira Pinheiro-Araujo C, Valera-Calero JA, Ortega-Santiago R. Biopsychosocial Factors for Chronicity in Individuals with Non-Specific Low Back Pain: An Umbrella Review. *Int J Environ Res Public Health*. 2022; 19(16):

Bibliografía

- 10145.
99. Pedersen CF, Andersen MØ, Carreon LY, Eiskjaer S. Applied Machine Learning for Spine Surgeons: Predicting Outcome for Patients Undergoing Treatment for Lumbar Disc Herniation Using PRO Data. *Global Spine J.* 2022; 12(5): 866-76.
 100. Ren GR, Yu K, Xie ZY, Wang PY, Zhang W, Huang Y, et al. Current Applications of Machine Learning in Spine: From Clinical View. *Glob spine J.* 2022; 12(8). Paginas ???
 101. Vagaska E, Litavcova A, Srotova I, Vlckova E, Kerkovsky M, Jarkovsky J, et al. Do lumbar magnetic resonance imaging changes predict neuropathic pain in patients with chronic non-specific low back pain?. *Medicine (Baltimore).* 2019; 98(17): e15377.
 102. Dunsmuir RA, Nisar S, Cruickshank JA, Loughenbury PR. No correlation identified between the proportional size of a prolapsed intravertebral disc with disability or leg pain. *Bone Joint J.* 2022; 104-B(6): 715-20.
 103. Mertimo T, Karppinen J, Niinimäki J, Blanco R, Määttä J, Kankaanpää M, et al. Association of lumbar disc degeneration with low back pain in middle age in the Northern Finland Birth Cohort 1966. *BMC Musculoskelet Disord.* 2022; 23(1): 359.
 104. Wang H, Liu C, Meng Z, Zhou W, Chen T, Zhang K, et al. Real-world study for identifying the predictive factors of surgical intervention and the value of magnetic resonance imaging in patients with low back pain. *Quant Imaging Med Surg.* 2022; 12(3): 1830-43.
 105. Jin L, Jiang C, Gu L, Jiang M, Shi Y, Qu Q, et al. Predictive Classification System for Low Back Pain Based on Unsupervised Clustering. *Glob spine J.* 2021: 21925682211001813.
 106. Shim JG, Ryu KH, Cho EA, Ahn JH, Kim HK, Lee YJ, et al. Machine learning approaches to predict chronic lower back pain in people aged over 50 years. *Medicina (Kaunas).* 2021; 57(11): 1230.
 107. Liew BXW, Kovacs FM, Rügamer D, Royuela A. Machine learning versus logistic regression for prognostic modelling in individuals with non-specific neck pain. *Eur Spine J.* 2022; 31(8): 2082-91.
 108. Eriksson S, Waldenberg C, Torén L, Grimby-Ekman A, Brisby H, Hebelka H, et al. Texture Analysis of Magnetic Resonance Images Enables Phenotyping of Potentially Painful Annular Fissures. *Spine (Phila Pa 1976).* 2022; 47(5): 430-7.
 109. Lagerstrand K, Hebelka H, Brisby H. Identification of potentially painful disc fissures in magnetic resonance images using machine-

- learning modelling. *Eur Spine J.* 2022; 31(8): 1992-9.
110. Gebrewold Y, Tesfaye B. Does lumbar MRI predict degree of disability in patients with degenerative disc disease? A prospective cross-sectional study at University of Gondar comprehensive specialized hospital, North West Ethiopia, 2020. *BMC Med Imaging.* 2022; 22(1): 138.
 111. Banitalebi H, Aaen J, Storheim K, Negård A, Myklebust TÅ, Grotle M, et al. A novel MRI index for paraspinal muscle fatty infiltration: reliability and relation to pain and disability in lumbar spinal stenosis: results from a multicentre study. *Eur Radiol Exp.* 2022; 6(1): 38.
 112. McIntosh G, Steenstra I, Hogg-Johnson S, Carter T, Hall H. Lack of Prognostic Model Validation in Low Back Pain Prediction Studies. *Clin J Pain.* 2018; 34(8): 748-54.
 113. Tagliaferri SD, Angelova M, Zhao X, Owen PJ, Miller CT, Wilkin T, et al. Artificial intelligence to improve back pain outcomes and lessons learnt from clinical classification approaches: three systematic reviews. *NPJ Digit Med.* 2020; 3: 93.

Bibliografía

10 Anexos

Anexos

Anexo 1. Hoja de información al paciente

Título del Estudio: Aplicación del análisis de textura sobre imágenes de RM convencional al estudio de pacientes con lumbalgia mecánica crónica

Investigador Principal: Dr Vicente Climent Peris. Médico especialista en cirugía ortopédica y traumatología.

Estimado Sr/Sra.:

Se le ha invitado a participar en un estudio sobre dolor lumbar crónico. En este documento se le ofrece información detallada sobre la naturaleza del estudio.

Tras leer esta información y solicitar de su médico las aclaraciones oportunas deberá firmar el documento de consentimiento informado si desea participar en el estudio.

Naturaleza del estudio y objetivo del estudio

Se trata de un estudio observacional en el que se pretende aplicar nuevos métodos de estudio de las imágenes de resonancia magnética para obtener más información de las mismas y que en un futuro se pueda aplicar este análisis para determinar el pronóstico de los pacientes con lumbalgia mecánica crónica.

Procedimientos del estudio

Se analizarán los datos recogidos en su historia clínica obtenidos en las visitas realizadas y se estudiarán distintos parámetros de su resonancia magnética.

Para este fin, el investigador principal accederá a su historia clínica y a sus imágenes de RM. Todos estos datos serán anonimizados de forma que en ningún momento se conocerá la identidad de los pacientes.

Los pacientes que acepten participar en el estudio no se verán sometidos a ninguna prueba o tratamiento distinto a los protocolos habituales de los servicios de COT / RHB del hospital Arnau de Villanova de Valencia.

Riesgos y molestias

La participación en este estudio no conlleva ningún riesgo o molestia adicional para el paciente ya que no se realizarán más pruebas ni se administrarán tratamientos adicionales a los que forman parte de la práctica clínica habitual.

Posibles beneficios del estudio

Es posible que usted no obtenga ningún beneficio directo del estudio, pero ayudará a que desarrollen nuevos procedimientos diagnóstico que puedan aplicarse a otros pacientes en el futuro.

Participación y retirada voluntarias

Su participación es totalmente voluntaria. Puede negarse a participar en el estudio sin que ello afecte a su atención médica futura.

Política de confidencialidad

De acuerdo con la legislación actual sobre protección de datos, usted otorga expresamente su consentimiento para incluir los datos de su historia clínica y los datos derivados de su participación en el estudio en un archivo de datos personales bajo la responsabilidad del investigador principal.

El acceso a la información quedará restringido al estudio de investigación, este acceso se realizará manteniendo una estricta confidencialidad de manera que no se viole la privacidad de los pacientes que participan en el estudio.

El investigador, durante el procesamiento y tratamiento de sus datos adoptará las medidas adecuadas para proteger y evitar el acceso a ellos por parte de terceros no autorizados. Únicamente se le identificará por el número de paciente. Su identidad tampoco

se revelará en ningún informe, resultado o publicación relacionada con el estudio.

Recibirá una copia firmada de esta hoja de información para que la guarde o para que la comente con su médico, familia o cualquier persona con la que lo quiera comentar.

Anexos

Anexo 2. Documento de consentimiento informado

Consentimiento Informado

Título del estudio: Aplicación del análisis de textura sobre imágenes de RM convencional al estudio de pacientes con lumbalgia mecánica crónica

Investigador Principal: Dr Vicente Climent Peris. Médico especialista en Cirugía Ortopédica y Traumatología.

Yo,(nombre del paciente)

He leído la hoja de información que se me ha facilitado

He podido preguntar acerca del estudio

He recibido información suficiente sobre el estudio

He hablado con el Dr.

Comprendo que mi participación es voluntaria

Comprendo que puedo retirarme del estudio cuando lo desee, sin tener que dar explicaciones y sin que esto repercuta en mis cuidados médicos.

Del mismo modo, declaro que se me ha informado de las medidas que se adoptarán para garantizar la confidencialidad y disociación de la información sobre mi persona obtenida durante el estudio y la posibilidad de ejercer mi derecho de acceso, rectificación, cancelación y oposición por medio de una solicitud formal al investigador principal.

Comprendo que puedo retirar mi consentimiento y finalizar mi participación en cualquier momento y sin necesidad de indicar el motivo.

Por la presente, acepto voluntariamente participar en el estudio.

Firma del paciente

Fecha

Nombre del paciente

Firma del investigador

Fecha

Nombre del investigador

Anexo 3. Cuestionario de discapacidad Rolland Morris

Cuando su espalda le duele, puede que le sea difícil hacer algunas de las cosas que habitualmente hace. Esta lista contiene algunas de las frases que la gente usa para explicar cómo se encuentra cuando le duele la espalda (o los riñones). Cuando las lea, puede que encuentre algunas que describan su estado de *hoy*. Cuando lea la lista, piense en cómo se encuentra usted *hoy*. Cuando lea una frase que describa como se siente hoy, póngale una señal. Si la frase no describe su estado de hoy, pase a la siguiente frase. Recuerde, tan solo señale la frase si está usted seguro de que describe cómo se encuentra usted hoy.

1. Me quedo en casa la mayor parte del tiempo por mi dolor de espalda.
 2. Cambio de postura con frecuencia para intentar aliviar la espalda.
 3. Debido a mi espalda, camino más lentamente de lo normal.
 4. Debido a mi espalda, no puedo hacer ninguna de las faenas que habitualmente hago en casa.
 5. Por mi espalda, uso el pasamanos para subir escaleras.
 6. A causa de mi espalda, debo acostarme más a menudo para descansar.
 7. Debido a mi espalda, necesito agarrarme a algo para levantarme de los sillones o sofás.
 8. Por culpa de mi espalda, pido a los demás que me hagan las cosas.
 9. Me visto más lentamente de lo normal a causa de mi espalda.
 10. A causa de mi espalda, sólo me quedo de pie durante cortos períodos de tiempo.
 11. A causa de mi espalda, procuro evitar inclinarme o arrodillarme.
 12. Me cuesta levantarme de una silla por culpa de mi espalda.
 13. Me duele la espalda casi siempre.
 14. Me cuesta darme la vuelta en la cama por culpa de mi espalda.
 15. Debido a mi dolor de espalda, no tengo mucho apetito.
 16. Me cuesta ponerme los calcetines - o medias - por mi dolor de espalda.
 17. Debido a mi dolor de espalda, tan solo ando distancias cortas.
 18. Duermo peor debido a mi espalda.
 19. Por mi dolor de espalda, deben ayudarme a vestirme.
 20. Estoy casi todo el día sentado a causa de mi espalda.
 21. Evito hacer trabajos pesados en casa, por culpa de mi espalda.
 22. Por mi dolor de espalda, estoy más irritable y de peor humor de lo normal.
 23. A causa de mi espalda, subo las escaleras más lentamente de lo normal.
 24. Me quedo casi constantemente en la cama por mi espalda.
-

Anexo 4. Informe CEI



CEI: 10/2017

**DR. JUAN CARLOS ANDREU BALLESTER, PRESIDENTE DEL COMITÉ
ÉTICO DE INVESTIGACIÓN DEL DEPARTAMENTO DE SALUD ARNAU DE
VILANOVA – LLIRIA**

INFORMA

Reunidos los miembros del Comité Ético de Investigación del Hospital Arnau de Vilanova, en su sesión ordinaria del día **26 de abril de 2017**, y una vez estudiada la documentación presentada por Vicent Climent y Julio Domenech como Investigador Principal y colaborador respectivamente, del Servicio COT del H. Lluís Alcanyís de Xativa y que se realizará en el servicio de traumatología del H. Arnau de Vilanova de Valencia, el proyecto de investigación, titulado: **"Valor diagnóstico y pronóstico de la textura de los elementos de la unidad vertebral en pacientes con lumbalgia mecánica crónica"**, se ajusta a las normas deontológicas establecidas para tales casos.

Se emite **dictamen favorable** a la realización de dicho proyecto en el Servicio de Traumatología en el H. Arnau de Vilanova de Valencia.

Y para que conste, lo firma en Valencia con fecha 26 de abril de dos mil diecisiete

Fdo. Dr. Juan Carlos Andreu-Ballester
Presidente del CEI



Anexo 5. Publicación

European Spine Journal
<https://doi.org/10.1007/s00586-023-07936-6>

ORIGINAL ARTICLE



Predictive value of texture analysis on lumbar MRI in patients with chronic low back pain

Vicente-Jose Climent-Peris¹ · Luís Martí-Bonmati² · Alejandro Rodríguez-Ortega³ · Julio Doménech-Fernández⁴

Received: 19 March 2023 / Revised: 2 August 2023 / Accepted: 30 August 2023
 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Purpose The aim of this study was to determine whether MRI texture analysis could predict the prognosis of patients with non-specific chronic low back pain.

Methods A prospective observational study was conducted on 100 patients with non-specific chronic low back pain, who underwent a conventional MRI, followed by rehabilitation treatment, and revisited after 6 months. Sociodemographic variables, numeric pain scale (NPS) value, and the degree of disability as measured by the Roland–Morris disability questionnaire (RMDQ), were collected. The MRI analysis included segmentation of regions of interest (vertebral endplates and intervertebral disks from L3–L4 to L5–S1, paravertebral musculature at the L4–L5 space) to extract texture variables (PyRadiomics software). The classification random forest algorithm was applied to identify individuals who would improve less than 30% in the NPS or would score more than 4 in the RMDQ at the end of the follow-up. Sensitivity, specificity, and the area under the ROC curve were calculated.

Results The final series included 94 patients. The predictive model for classifying patients whose pain did not improve by 30% or more offered a sensitivity of 0.86, specificity 0.57, and area under the ROC curve 0.71. The predictive model for classifying patients with a RMDQ score 4 or more offered a sensitivity of 0.83, specificity of 0.20, and area under the ROC curve of 0.52.

Conclusion The texture analysis of lumbar MRI could help identify patients who are more likely to improve their non-specific chronic low back pain through rehabilitation programs, allowing a personalized therapeutic plan to be established.

Keywords Radiomics · Texture analysis · MRI · Back pain

Introduction

Non-specific mechanical low back pain is currently one of the most serious health concerns in developed societies, with a high prevalence and a high percentage of chronification [1]. Magnetic resonance imaging (MRI) has become the most commonly used technique for diagnosing low

back pain, frequently replacing conventional X-ray as the first imaging test to be performed. However, its widespread use has not been reflected in the health outcomes of patients suffering from non-specific low back pain [2]. It is difficult to establish a causal relationship between low back pain and findings such as disk degeneration, facet joint arthropathy, disk protrusion, hernia, or spine alignment alteration. Several studies have shown the existence of degenerative changes in the lumbar MRI in asymptomatic patients: Up to 60% of the general population without low back pain can present disk degenerative signs, and this percentage can exceed 80–90% in asymptomatic people over the age of 55 [3–5]. One of the top priorities in spinal pathology research is the search for prognostic markers of low back pain resolution, permanence, and aggravation. The MRI, although being the imaging technique of choice in spinal pathology, presents a low clinical correlation and therefore provides a limited prognostic value [6].

✉ Vicente-Jose Climent-Peris
 vicentcp@gmail.com

¹ Orthopaedic Surgery Department, Hospital Mare de Déu Dels Liris, Alcoi (Alicante), Spain

² Medical Imaging Department, Hospital Universitario y Politécnico La Fe, Valencia, Spain

³ Graphic Engineering Department, Universidad Politécnica de Valencia, Valencia, Spain

⁴ Orthopaedic Surgery Department, Clínica Universidad de Navarra, Pamplona, Spain

Published online: 16 September 2023

Springer

Radiomics, or texture analysis, emerges as a new paradigm in medical image analysis, allowing the extraction of a series of quantifiable parameters that are objective and non-visible to the human eye, and which are contained in the images obtained through the techniques used (including CT scan and MRI). They represent a potential source of biomarkers that have already proven their effectiveness in cancer diagnosis and phenotyping, as well as prognosis, response to treatment, and disease's evolution [7, 8].

Some studies have explored the possibility of applying texture analysis on conventional MR images to diagnose intervertebral disk pathology. Texture analysis appears to be more sensitive than conventional MRI analysis in T2-weighted images for detecting a pathological disk [9] and has proven to be more capable of identifying patients with back pain than conventional MRI analysis [10, 11]. No prognostic studies have been found that allow us to predict how pain and disability will evolve in the medium term.

Methods

Our purpose research was to establish a predictive model based on texture analysis of conventional MRI images of the lumbar spine to estimate the medium-term evolution of patient with non-specific chronic low back pain.

A prospective, observational, non-interventional study was designed. A series of 100 consecutive patients diagnosed with chronic mechanical low back pain were invited to participate. Patients were recruited from a public Rehabilitation and Orthopedic Surgery Department serving a population of 320,000 people. The following inclusion criteria were established: age between 18 and 65 years old, non-specific mechanical low back pain of at least 6 months of duration [12], lumbar spine MRI exploration request, receiving a rehabilitation treatment for 6 months and attending a subsequent checkup visit and signing the informed consent document. The following exclusion criteria were considered: source-specific pain labeled as vertebral fracture, spine or neural structure neoplasm, spondyloarthropathy or spondylodiscitis, having received a rehabilitation treatment previous to the first visit, previous spine surgery, intellectual disability, mental disorder, substance abuse or dependence and analphabetism, as it hinders the completion of the disability questionnaires.

Patients had two visits: the initial, in which the study of the lumbar spine MRI was requested, and the follow-up visit 6 months later, right after having completed the rehabilitation program.

The following variables were included:

- Demographic variables, age, sex, basic anthropometrics (weight and height), and the body mass index (BMI) calculated in Kg/cm^2 .
- Clinical variables: the score of the numeric rating scale for pain (NRS) and the score of the Roland–Morris disability questionnaire (RDMQ) obtained in both the initial and the 6-month visit.
- Radiomic variables: basic texture features from the PyRadiomics library.

The procedure for the radiomic analysis of the MRI started selecting the lumbar spine MRI examinations performed. The sagittal and the axial T2-weighted sequences were selected for analysis. Images were obtained in DICOM format and converted into NIFTI using *dicom2nii* software. Manual segmentation of the following regions of interest (ROI) was performed using the ITK-SNAP 3.6.0 software:

- The intervertebral disks L3–L4, L4–L5, L4–S1, the upper and lower vertebral endplates of each one of these disks in the mid-sagittal section of the T2-weighted image. A ROI located in the cerebrospinal fluid was segmented for image intensity normalization (Fig. 1).
- Two ROIs corresponding to the right and left lumbar paravertebral musculature were segmented in the transverse T2-weighted section over the intervertebral disk L4–L5.

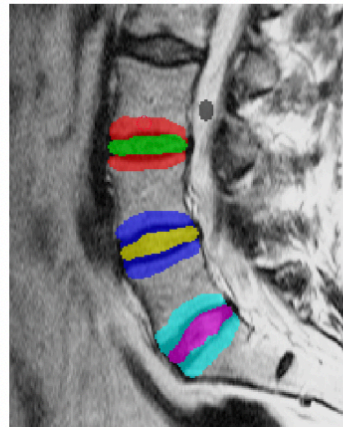


Fig. 1 ROIs segmented in the sagittal T2-weighted sequence: L3–L4, L4–L5, and L5–S1 intervertebral disks and upper and lower vertebral endplates at these levels, in the mid-sagittal image

- Two ROIs located symmetrically at the midline in the subcutaneous adipose tissue that were used for the normalization of the image intensity were also segmented (Fig. 2).

The manual segmentation was performed twice in each patient, separated by weeks, with the objective of verifying the procedure's reproducibility and removing any negative effects that may arise when performing manual segmentation, averaging the values obtained for each ROI.

A total of 107 basic texture features were calculated for each ROI using the PyRadiomics software: shape parameters, first-order parameters, and GLCM, GLSZM, GLRLM, NGTDM, and GLDM matrices. The texture features obtained from disks and the intervertebral endplates were averaged, obtaining a single value per patient for each disk and endplate texture features.

A features reduction was done prior to the statistical analysis: those features that could provide the most information to our study were selected, and those whose contribution to a statistical model was negligible or redundant and was likely to lead to multicollinearity problems were removed. For that, a previously described 4-step method was implemented [13]: assessment of the robustness of the variable ($ICC > 0.85$), assessment of the information (median absolute deviation $MAD > 0$), assessment of the prognostic performance ($C\text{-Index} > 0.58$), and reduction of redundancy (correlation index > 0.9) selecting the one with the higher C-Index.

Statistical analysis

Descriptive statistics were obtained stating the continuous variables in mean (with standard deviation) or median

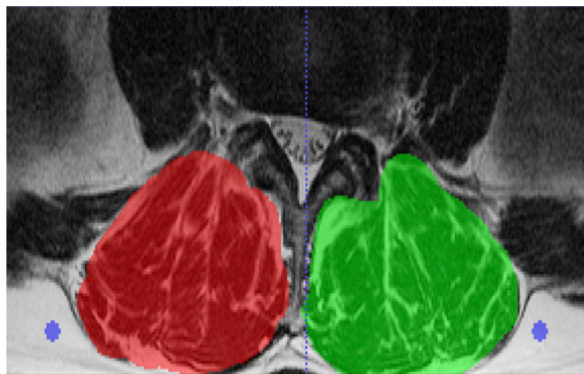
(with interquartile range) according to its normality or non-normality and the categorical variables in number (with percentage). The normal distribution was checked using the Kolmogorov–Smirnov test for determining the use of parametric or nonparametric tests. These analyses were performed with the IBM SPSS Statistics software (version 25).

A classification analysis was conducted based on the texture features selected, with the objective of developing a predictive model that would allow patients to be classified into different groups, according to their response to the treatment:

- Patients with less than 30% (minimum relevant clinical change) of pain improvement (NRS) at 6 months.
- Patients with a score higher than 4 in the disability questionnaire (RMDQ) at 6 months; the individuals with a score equal or lower than 4 are considered as having no significant disability.

In order to perform this analysis, the machine learning supervised classification random forest algorithm was used, and two predictive models were developed including the different variables listed in the study: age, sex, BMI, initial values of NRS and RMDQ, and the selected texture features. Prior to running the random forest algorithm, a principal component analysis (PCA) was performed; this is a statistical method for simplifying the complexity of sample spaces with multiple dimensions while retaining their information. Each model was trained with 70% of the sample and was subsequently applied in the remaining 30% for its internal validation. A confusion matrix was generated for each model using the predictions obtained in 30% of the sample, and the following parameters were calculated: accuracy, sensitivity (recall), specificity, positive predictive

Fig. 2 ROIs segmented in the axial T2-weighted sequence: paravertebral muscle at the level of L4–L5 intervertebral disk



value (precision), negative predictive value, and area under the ROC curve. Positive and negative likelihood ratios, the Nolan nomogram, the diagnostic odds ratio and the F1 score were also calculated.

The statistical models and analyses were developed using Python 3.6.10 (Python software foundation, <https://www.python.org>). The random forest models were built using the packages Sklearn [13] and Pandas [14].

The project was adjusted to comply with current Spanish guidelines for patient protection in clinical trials regarding the collection, storage, and custody of anonymized personal data. Personal passwords and data encryption were used to protect the participants' identities and respect their right to confidentiality. In addition, researchers adhered to the declaration of Helsinki on medical investigation [15]. Informed consent was obtained from all individual participants included in the study. The present project was approved by the investigation ethics committee of the hospital (HAV CIE 10/2017).

Results

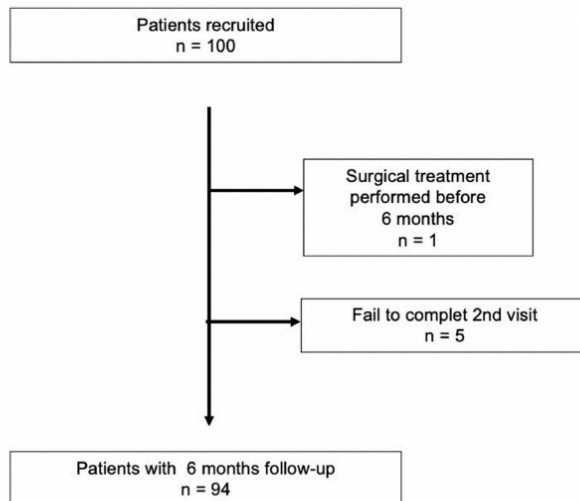
A total of 94 patients finally completed the study; the recruitment and follow-up processes of patients are gathered in Fig. 3. Overall, the sample consisted of patients of both sexes, but with a slight predominance of women (61%)

with an average age of 49 years old. The demographic and clinical features of patients are shown in Table 1. If we analyze the individual evolution of patients at 6 months, 21 of them (20.3%) had worse scores in the NRS, 19 (22.3%) had no changes, and 54 (57.4%) described improvement; 31 patients (33%) reported a pain improvement of at least 30%, which is considered to be a relevant clinical improvement. If we analyze the disability evolution in the same period, 18 patients (19.15%) had worse RMDQ scores, 5 (5.32%) had no changes, and 71 (75.73%) improved their disability degree; 40 of them (42.4%) presented an RMDQ lower or

Table 1 Demographic and clinical features of the sample. Categorical variables age and height expressed as mean (\pm SD), others expressed as median (interquartile range)

Age (years)	49.57 (\pm 8.92)
Sex (%)	
Men	37 (39.4)
Women	57 (60.6)
Weight (Kg)	76.4 (20.5)
Height (m)	1.68 (\pm 0.084)
BMI (Kg/m ²)	25.86 (5.17)
Initial NRS (0–10)	7 (3)
Final NRS (0–10)	6 (4)
Initial RMDQ (0–24)	10 (7)
Final RMDQ (0–24)	5 (7)

Fig. 3 Flowchart of the recruitment and follow-up process



equal to 4, which states that they did not present disability at the end of the follow-up.

After implementing the variable reduction method, 45 texture features were selected and included in the analysis (Table 2).

Predictive model for pain

The results obtained with the predictive model calculated on 70% of the sample to identify patients that will improve less than 30% in the NRS scores after 6 months of rehabilitation treatment, when applied to the remaining 30%, were as follows: accuracy 0.79, recall (sensitivity) 0.86, specificity 0.57, precision (positive predictive value) 0.86, negative predictive value 0.57, F1 score 0.86, positive likelihood ratio 2, negative likelihood ratio 0.25, and area under the ROC curve 0.71 (Fig. 4). The nomogram of Nolan for this model is represented in Fig. 5.

Predictive model for disability

When applied to the remaining 30% of the sample, the predictive model calculated on 70% of the sample to identify patients with an RMDQ score greater than 4 after 6 months of rehabilitation treatment produced the following results:

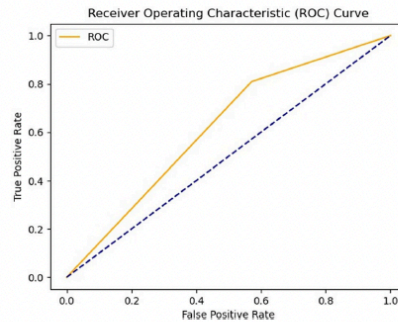


Fig. 4 ROC curve of the predictive model to predict NRS (numerical pain rating scale) improvement lower or equal to 30%. AUC=0.71

accuracy 0.61, recall (sensitivity) 0.83, specificity 0.20, precision (positive predictive value) 0.65, negative predictive value 0.4, F1 score 0.73, positive likelihood ratio 1.04, negative likelihood ratio 0.85, and area under the ROC curve 0.52.

Table 2 Selected texture features from the basics extracted with PyRadiomics

<i>ROI vertebral end plates</i>	<i>ROI muscle</i>
Shape surface area	Shape major axis length
Shape major axis length	Shape minor axis length
Shape maximum 2D diameter column	Shape voxel volume
First-order percentile 90	Shape sphericity
First-order total energy	Shape surface volume ratio
GLCM Imc2	Shape maximum 2D diameter column
GLZCM gray-level variance	Shape surface volume ratio
GLSZM zone variance	First-order energy
<i>ROI intervertebral disk</i>	First-order kurtosis
Shape maximum 2D diameter column	First-order 10 percentile
Shape surface volume ratio	First-order median
Shape voxel volume	First-order maximum
First-order root mean squared	GLCM ldn
First-order mean absolute deviation	GLCM ldm
First-order percentile 10	GLSZM gray-level non-uniformity
First-order total energy	GLSZM low gray-level zone emphasis
First-order skewness	GLSZM large area high gray-level emphasis
First-order kurtosis	GLSZM zone variance
GLCM Imc1	NGTDM strength
GLCM cluster shade	NGTDM coarseness
GLSZM large area high gray-level emphasis	NGTDM busyness
NGTDM busyness	GLDM dependence non-uniformity
GLDM large dependence low gray-level emphasis	GLDM large dependence low gray-level emphasis

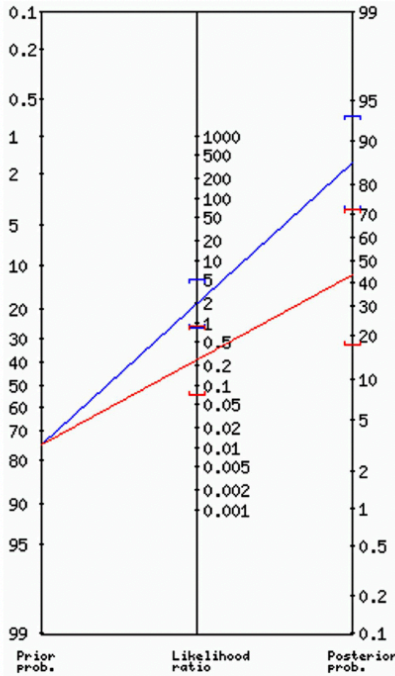


Fig. 5 Nolan nomogram of the predictive model to predict NRS (numerical rating scale for pain) improvement lower or equal to 30%

Discussion

Random forest algorithm using texture analysis of the different vertebral structures allows, together with the basic clinical parameters included in the study, to classify patients without non-specific chronic low back pain improvement after 6 months of conservative treatment, with an acceptable sensitivity (86%) but a low specificity (57%). According to these findings, the tested model in this study would have good reliability if the test determined patient improvement (a negative test to non-improvement), but it would have less value if the test result established non-recovery (a positive test to non-improvement). The pre-test probability of non-improving pain in the studied sample was 75%, which increased to 86% if the model generated a positive result and decreased to 43% if the model produced a negative result. These data would justify its inclusion in clinical practice as

a screening tool in patients with non-specific chronic low back pain.

The analysis conducted to assess the prognosis of the patient's disability based on the RMDQ scale yielded unfavorable results; a predictive model capable of identifying patients with residual disability (RMDQ > 4) or classifying those patients with no disability was not obtained.

Several studies have examined the risk factors for poor prognosis or lumbar pain chronification, emphasizing the role of psychosocial factors as limitations to patient recovery [16, 17]. There are few studies whose objective is to analyze the structural alterations of MRI images in order to determine the non-specific lumbar pain prognosis after a conservative treatment, in contrast to the existence of multiple recent studies that propose models for predicting the results of a surgical treatment or the possibility of postoperative complications [18, 19].

Recent studies show that degenerative changes identified in the lumbar spine MRI are not capable of determining the medium-term prognostic of patients with low back pain. However, others find a higher prevalence of low back pain in patients with higher degree of disk degeneration, although this correlation disappears in patients with mental disorders [20].

Wang et al. performed a retrospective study in which they developed two predictive models of the need for surgical treatment in patients with low back pain and compared the AUC of one model that included clinical variables to another that also included an MRI; the AUC increased from 0.75 to 0.77 when the MRI was included among the variables analyzed [21]. Using a non-supervised clustering method, Jim et al. relied on the disk signal strength and its height in the MRI to classify patients in three groups; in one of these groups, the percentage of patients that developed low back pain was higher [22]. In a cross-sectional study, Shuim et al. assessed the predictive capacity of different artificial intelligence algorithms when identifying patients with risk of developing chronic low back pain [23]. Thirteen demographic and clinical variables were included in these models without radiological. The AUC obtained with random forest was 0.660, lower to the one obtained with our model, and the model with better predictive capacity was the one calculated with the Naive Bayes algorithm, with an AUC of 0.712, which was equal to our model.

Liew et al. showed that various models obtained through artificial intelligence algorithms outperformed conventional regressive models in terms of predictive capacity when applied to the evolution of patients with neck pain. These models included conventional images variables, in addition to clinical and sociodemographic variables. The best result was obtained with the XGBoost algorithm, with AUC = 0.726, to predict the persistence of neck pain at 3 months [24].

Several studies have applied the texture analysis of lumbar spine MRI in order to classify patients with low back pain. Abdollah et al. detected differences between patients with low back pain or without it in the homogeneity and energy parameters at the level of the intervertebral disk L5–S1, as well as in contrast, homogeneity, and energy of the adjacent vertebral endplates [10]. Ketola et al. obtained an AUC of 0.91 for distinguishing between patients with low back pain and asymptomatic individuals using texture features calculated at the L4–L5 and L5–S1 levels, which was lower when the superior lumbar disks were used (AUC=0.76) or when the Pfirrmann classification was used (AUC=0.62) [25]. Eriksson et al. used texture analysis to identify annulus fibrosus fissures, particularly those that were missed at first glance in the MRI, obtaining an AUC of 0.97, its value for pain prediction being between 0.69 and 0.75, and showing high sensitivity (0.93) but low specificity (0.4) [26].

Lagerstrand et al. reported similar results, with a pain prediction precision of 71% [27]. Although these values are similar to those obtained in our study, they do not provide a medium-term predictive value of pain evolution; instead, they only correlate the texture analysis findings with the patient's state at the time the MRI was performed.

The correlation of lumbar MRI degenerative changes with disability has also been a reason for recent publications. Gebrewold et al. conclude that there is no correlation with the presence of protrusion or disk herniation, foraminal stenosis, or Modic changes, although they found a weak correlation with the presence of lumbar spinal stenosis [28]; Dunsmuir et al. conclude that there is no correlation between disability and size of a disk herniation [19]. Banitalebi et al. did not find correlation between the degree of adipose tissue degeneration of the lumbar paravertebral musculature and the disability degree [29].

No studies were found that analyzed the prognostic value of the texture analysis in patients with non-specific low back pain, and hence, it is not possible to compare our results with similar studies. We can compare them with prognostic studies based on clinical variables and some studies that incorporate conventional MRI. A systematic review and meta-analysis including prospective studies that evaluated the predictive capacity of different scales concluded that the predictive capacity for pain was low, with an AUC of 0.63, and acceptable for disability with an AUC of 0.71 [30]. A prospective study to assess the predictive capacity of three questionnaires on non-specific low back pain found a low capacity to predict poor prognosis, with an AUC of 0.61 for the PICKUP tool, 0.66 for the short form Orebro Musculoskeletal pain screening questionnaire, and 0.69 for the STaRT back tool [31]. Another prospective study obtained a more favorable result with a

model based on clinical variables to predict good prognosis (AUC=0.77), with no significant difference when incorporating vertebral MRI findings (AUC=0.78) [32].

This study presents some limitations. There is a sample selection bias because we included patients who were referred to a hospital-level specialized unit, which implies that many of them had previously received mostly ineffective treatment; this group of patients had long-term chronic pain. The findings of this study are unlikely to be applicable to patients suffering from acute or subacute low back pain. Another bias to be considered in external validity is the type of treatment received by patients: although the hospital-level rehabilitation treatment was homogeneous, other treatments that the patient may have received on their own (physiotherapeutic, pharmacological) had not been controlled. This bias would not have a relevant importance, since these non-controlled treatment variations would take part of the management and typical evolution of these patients. The sample size could be considered insufficient for the number of variables included in the models, even though half of the predictive studies published applying artificial intelligence on spine pathology include less than 100 patients [33]. This could lead to what is known as model overfitting: when applying the model outside the sample, it could deteriorate its performance. Although the study's inclusion of simple clinical and reproducible variables is regarded as a strength, psychosocial variables were not included, despite the fact that they appear to have a high specific weight in the evolution of low back pain [34]. The ROIs were manually segmented, which may have reduced intra- and inter-observer precision. To reduce these biases, the segmentation process was repeated twice and the results were averaged.

In conclusion, a predictive artificial intelligence classification random forest model, integrating sociodemographic, clinical, and radiomic features from the intervertebral disk, vertebral endplates, and the retrosomatic musculature, is able to identify chronic low back pain patients that will have at least 30% of pain improvement after 6 months of rehabilitation treatment. Lumbar MRI analysis, segmentation, and texture analysis can be automated using artificial intelligence algorithms, allowing for routine use in the study of patients with non-specific chronic low back pain. It would also provide useful information on the medium-term evolution of the patient's pain and help us establish a more personalized treatment plan.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

- Vos T, Barber RM, Bell B et al (2015) Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013. *Lancet* 386(9995):743–800. [https://doi.org/10.1016/S0140-6736\(15\)60692-4](https://doi.org/10.1016/S0140-6736(15)60692-4)
- Rodríguez Reiro C (2013) Utilidad de la resonancia magnética en pacientes con dolor lumbar inespecífico. Ministerio de Sanidad, Servicios Sociales e Igualdad. Unidad de Evaluación de Tecnologías Sanitarias de la Comunidad de Madrid. Informes de evaluación de tecnologías sanitarias Tecnologías Sanitarias: 1–54. https://www.seram.es/images/site/utilidad_de_la_rmm_lumbar.pdf
- Weishaupt D, Zanetti M, Hodler J, Boos N (1998) MR imaging of the lumbar spine: prevalence of intervertebral disk extrusion and sequestration, nerve root compression, end plate abnormalities, and osteoarthritis of the facet joints in asymptomatic volunteers. *Radiology* 209(3):661–666. <https://doi.org/10.1148/radiology.209.3.9844656>
- Tonosu J, Oka H, Higashikawa A, Okazaki H, Tanaka S, Matsudaira K (2017) The associations between magnetic resonance imaging findings and low back pain: a 10-year longitudinal analysis. *Plos One* 12(11):e0188057. <https://doi.org/10.1371/journal.pone.0188057>
- Rahyussalim AJ, Zufar MLL, Kurniawati T (2020) Significance of the association between disc degeneration changes on imaging and low back pain: a review article. *Asian Spine J* 14(2):245–57. <https://doi.org/10.31616/asj.2019.0046>
- Steffens D, Hancock MJ, Maher CG, Williams C, Jensen TS, Latimer J (2014) Does magnetic resonance imaging predict future low back pain? A systematic review. *Eur J Pain* 18(6):755–765. <https://doi.org/10.1002/j.1532-2149.2013.00427.x>
- Parekh V, Jacobs MA (2016) Radiomics: a new application from established techniques. *Expert Rev Precis Med drug Dev* 1(2):207–226. <https://doi.org/10.1080/23808993.2016.1164013>
- Marti-Bonmati L (2019) Radiomics and imaging biomarkers in observational clinical studies with retrospective data. *An RANM* 136:34–42. <https://doi.org/10.32440/ar.2019.136.01.rev07>
- Mayerhoefer ME, Stelzener D, Bachbauer W, Welsch GH, Mamsich TC, Szczypinski P et al (2012) Quantitative analysis of lumbar intervertebral disc abnormalities at 3.0 Tesla: value of T2 texture features and geometric parameters. *NMR Biomed* 25(6):866–872. <https://doi.org/10.1002/nbm.1803>
- Abdollah V, Parent EC, Dolatabadi S, Marr E, Croutze R, Wachowicz K, Kawchuk G (2021) Texture analysis in the classification of T2-weighted magnetic resonance images in persons with and without low back pain. *J Orthop Res* 39(10):2187–2196. <https://doi.org/10.1002/jor.24930>
- Ketola JHJ, Inkinen SI, Karppinen J, Niinimäki J, Tervonen O, Nieminen MT (2021) T2-weighted magnetic resonance imaging texture as predictor of low back pain: a texture analysis-based classification pipeline to symptomatic and asymptomatic cases. *J Orthop Res* 39(11):2428–2438
- Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CC, Chenot JF, van Tulder M, Koes BW (2018) Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. *Eur Spine J* 27(11):2791–2803. <https://doi.org/10.1007/s00586-018-5673-2>
- Pedregosa F, Michel V, Grisel O, Blondel M, Prettenhofer P, Weiss R et al (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830. <https://doi.org/10.48550/arXiv.1201.0490>
- Mckinney W (2010) Data structures for statistical computing in Python. *Proc Python Sci Conf* 445(1):56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>
- World Medical Association (2013) World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 310(20):2191–2194. <https://doi.org/10.1001/jama.2013.281053>
- Pedersen CF, Andersen MØ, Carreon LY, Eiskjaer S (2022) Applied machine learning for spine surgeons: predicting outcome for patients undergoing treatment for lumbar disc herniation using PRO data. *Global Spine J* 12(5):866–876. <https://doi.org/10.1177/2192568220967643>
- Ren GR, Yu K, Xie ZY, Wang PY, Zhang W, Huang Y et al (2022) Current applications of machine learning in spine: from clinical view. *Glob Spine J* 12(8):1827–1840. <https://doi.org/10.1177/21925682211035363>
- Vagaska E, Litavcova A, Srotova I, Vlckova E, Kerkovsky M, Jarkovsky J et al (2019) Do lumbar magnetic resonance imaging changes predict neuropathic pain in patients with chronic non-specific low back pain? *Med (Baltim)* 98(17):e1537. <https://doi.org/10.1097/md.00000000000015377>
- Dunsmuir RA, Nisar S, Cruickshank JA, Loughenbury PR (2022) No correlation identified between the proportional size of a prolapsed intravertebral disc with disability or leg pain. *Bone Joint J* 104-B(6):715–720. <https://doi.org/10.1302/0301-620X.104B6.BJ3-2021-1725.R2>
- Mertimo T, Karppinen J, Niinimäki J, Blanco R, Määttä J, Kankaanpää M et al (2022) Association of lumbar disc degeneration with low back pain in middle age in the northern Finland birth cohort 1966. *BMC Musculoskelet Disord* 23(1):359. <https://doi.org/10.1186/s12891-022-05302-z>
- Wang H, Liu C, Meng Z, Zhou W, Chen T, Zhang K et al (2022) Real-world study for identifying the predictive factors of surgical intervention and the value of magnetic resonance imaging in patients with low back pain. *Quant Imaging Med Surg* 12(3):1830–1843. <https://doi.org/10.21037/qims-21-584>
- Jin L, Jiang C, Gu L, Jiang M, Shi Y, Qu Q et al (2023) Predictive classification system for low back pain based on unsupervised clustering. *Glob Spine J* 13(3):630–635. <https://doi.org/10.1177/21925682211001813>
- Shim JG, Ryu KH, Cho EA, Ahn JH, Kim HK, Lee YJ et al (2021) Machine learning approaches to predict chronic lower back pain in people aged over 50 years. *Medicina (Kaunas)* 57(11):1230. <https://doi.org/10.3390/medicina5711230>
- Liew BXW, Kovacs FM, Rügamer D, Royuela A (2022) Machine learning versus logistic regression for prognostic modelling in individuals with non-specific neck pain. *Eur Spine J* 31(8):2082–2091. <https://doi.org/10.1007/s00586-022-07188-w>
- Ketola JHJ, Inkinen SI, Karppinen J, Niinimäki J, Tervonen O, Nieminen MT (2021) T2-weighted magnetic resonance imaging texture as predictor of low back pain: a texture analysis-based classification pipeline to symptomatic and asymptomatic cases. *J Orthop Res* 39(11):2428–2438. <https://doi.org/10.1002/jor.24973>
- Eriksson S, Waldenberg C, Torén L, Grimby-Ekman A, Brisby H, Hebelka H et al (2022) Texture analysis of magnetic resonance images enables phenotyping of potentially painful annular fissures. *Spine (Phila Pa 1976)* 47(5):430–437. <https://doi.org/10.1097/brs.0000000000004160>
- Lagerstrand K, Hebelka H, Brisby H (2022) Identification of potentially painful disc fissures in magnetic resonance images using machine-learning modelling. *Eur Spine J* 31:1992–1999. <https://doi.org/10.1007/s00586-021-07066-x>
- Gebrewold Y, Tesfaye B (2022) Does lumbar MRI predict degree of disability in patients with degenerative disc disease? A prospective cross-sectional study at University of Gondar comprehensive