

# Aproximación a un modelo predictivo de degeneración discal

Solé Florensa, Pau; Peroy García, Jaume Lluís; Mas Atance, Jaume; Hernández Umaña, Juan Bernardo; Jové Talavera, Ramon.

## Objetivo

- Cuantificar la influencia de las variables recogidas en el grado de degeneración discal objetivado por RM, generar y comparar dos modelos predictivos.

## Material y métodos

- Estudio retrospectivo que incluye a 150 pacientes mayores de 18 años intervenidos de patología vertebral (patología discal y fracturas), se excluyeron pacientes con patología séptica y tumoral.
- Se recogieron datos biométricos (edad, género, IMC), hábitos de vida (trabajo físico, tabaco, consumo crónico de AINE), antecedentes (diabetes, hipertensión arterial, dislipemia), valores analíticos (perfil lipídico, hematies, urato) y grado de degeneración discal en RM (clasificación de Pfirrmann) (Figura1).
- Los grados de Pfirrmann se dividieron en menor degeneración (I y II) y mayor degeneración (III-V).
- Los datos se analizaron mediante el software R-Studio (v.2023.06.0). Se procedió a la realización de un análisis univariante y mutivariante de los datos y se realiza un modelo de regresión logística y un análisis discriminante no paramétrico.

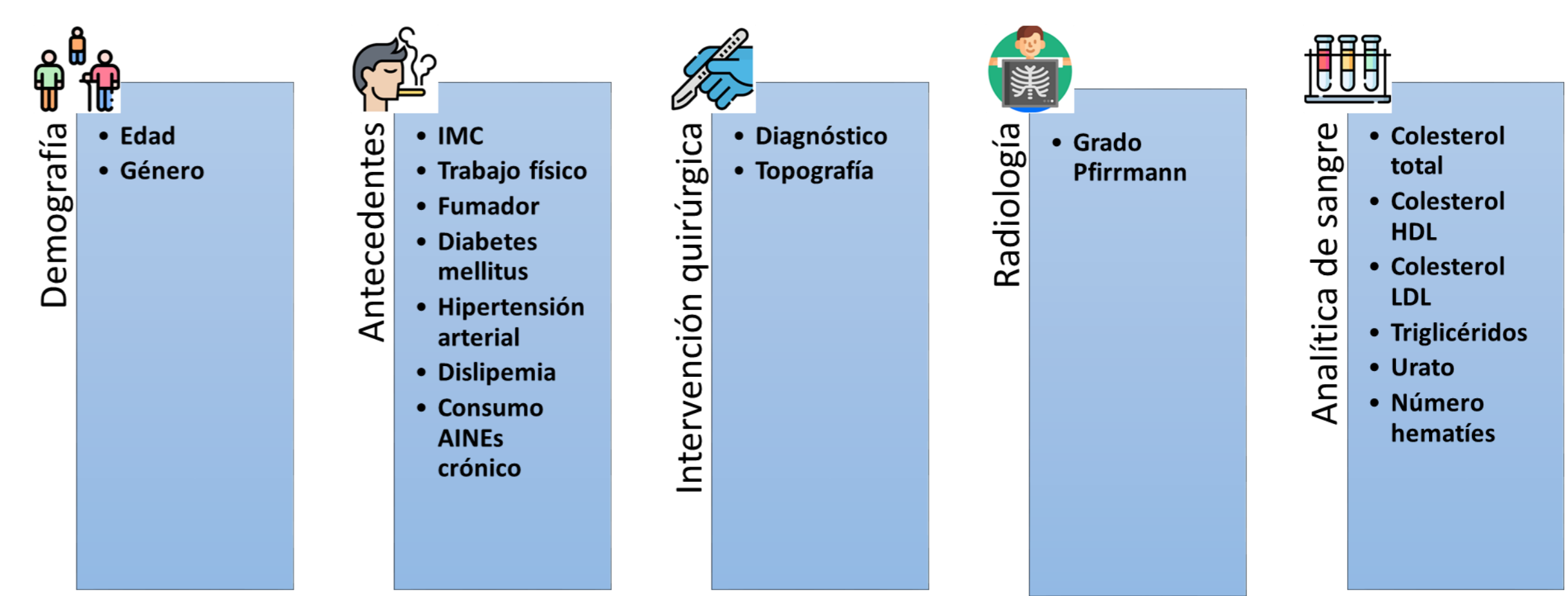


Figura 1. Variables recogidas en el estudio.

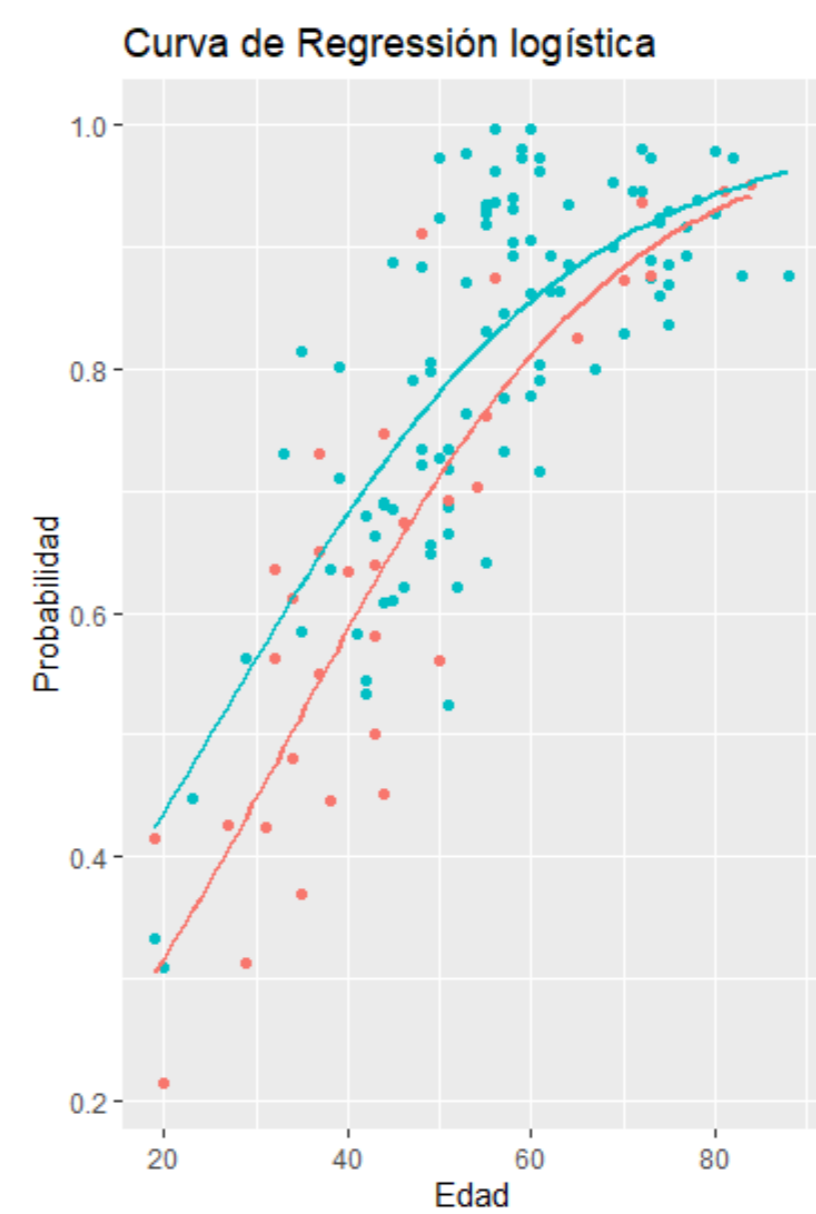


Figura 2. Curva de regresión logística para la variable edad. El RR obtenido es 1,06 (1,03-1,09; p=0). La probabilidad de degeneración se iguala con la edad en ambos grupos.

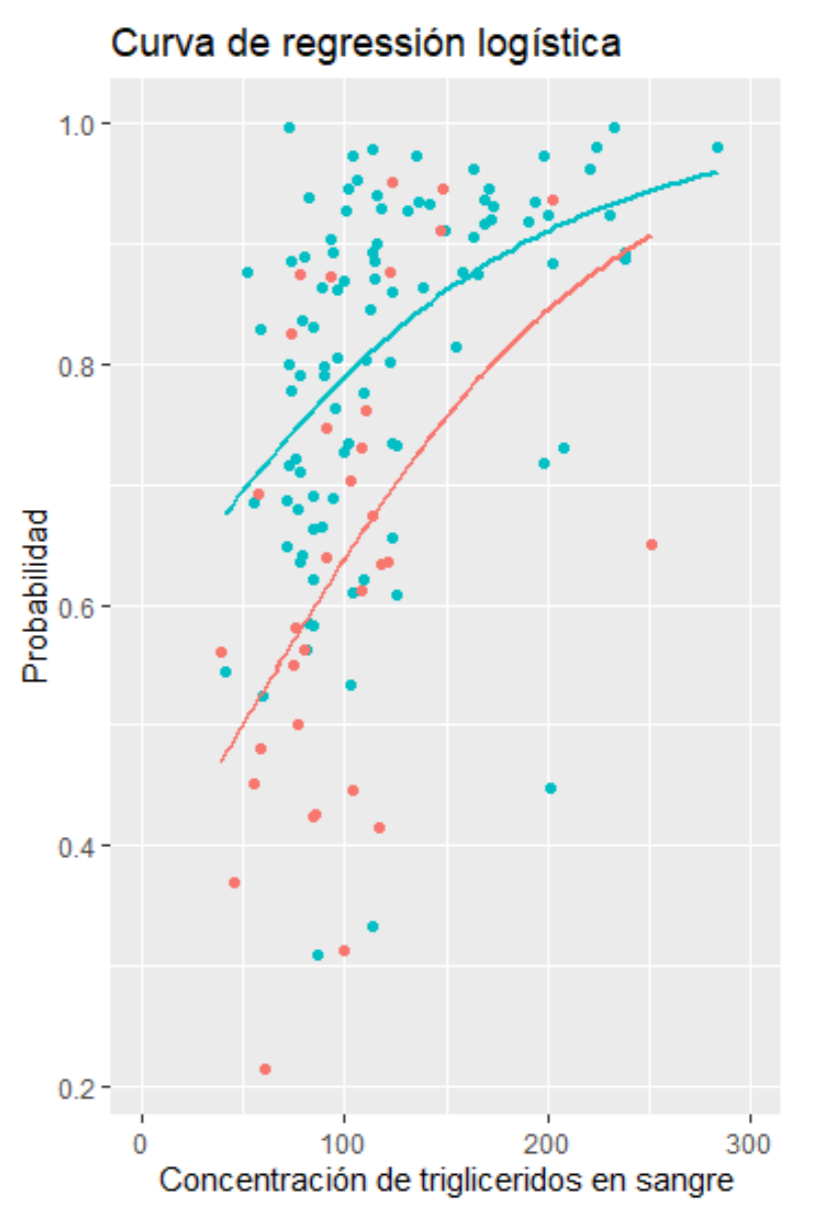


Figura 3. Curva de regresión logística para la variable triglicéridos. El RR es 1,01 (0,99-1,02; p=0,12). Con el aumento de triglicéridos se eleva la probabilidad de degeneración.

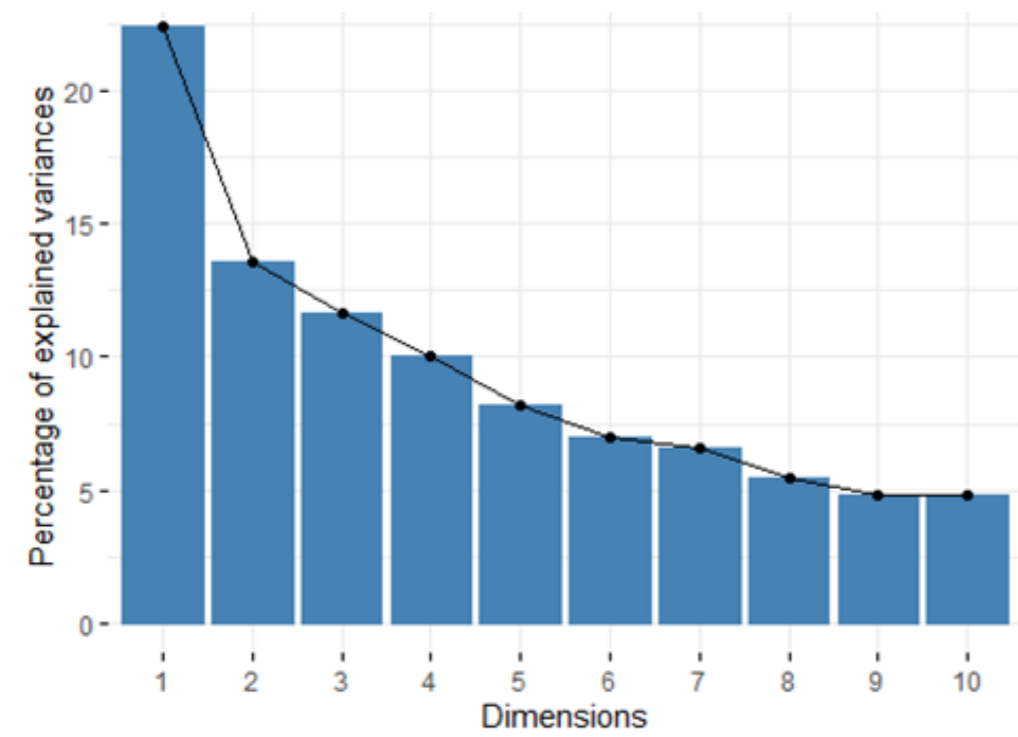


Figura 4. Diagrama de Scree. Reducción de la dimensión de los datos mediante el análisis de componentes principales. Con 8 componentes se explica el 80% de la variabilidad de la muestra.

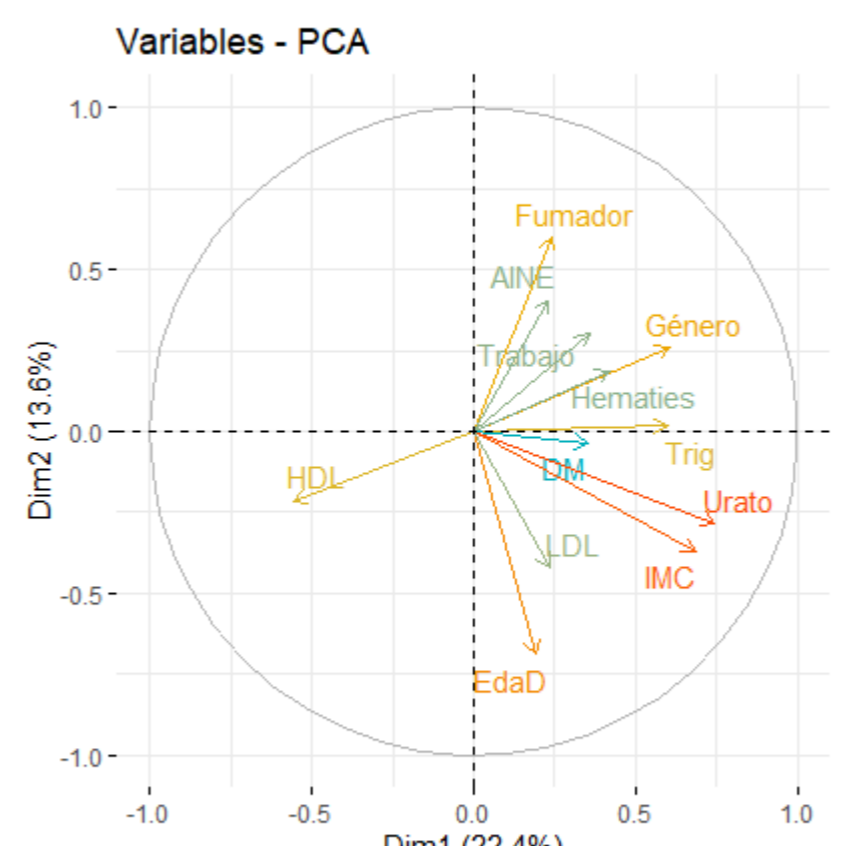


Figura 5. La contribución de las distintas variable de los dos primeros componentes principales a la variabilidad.

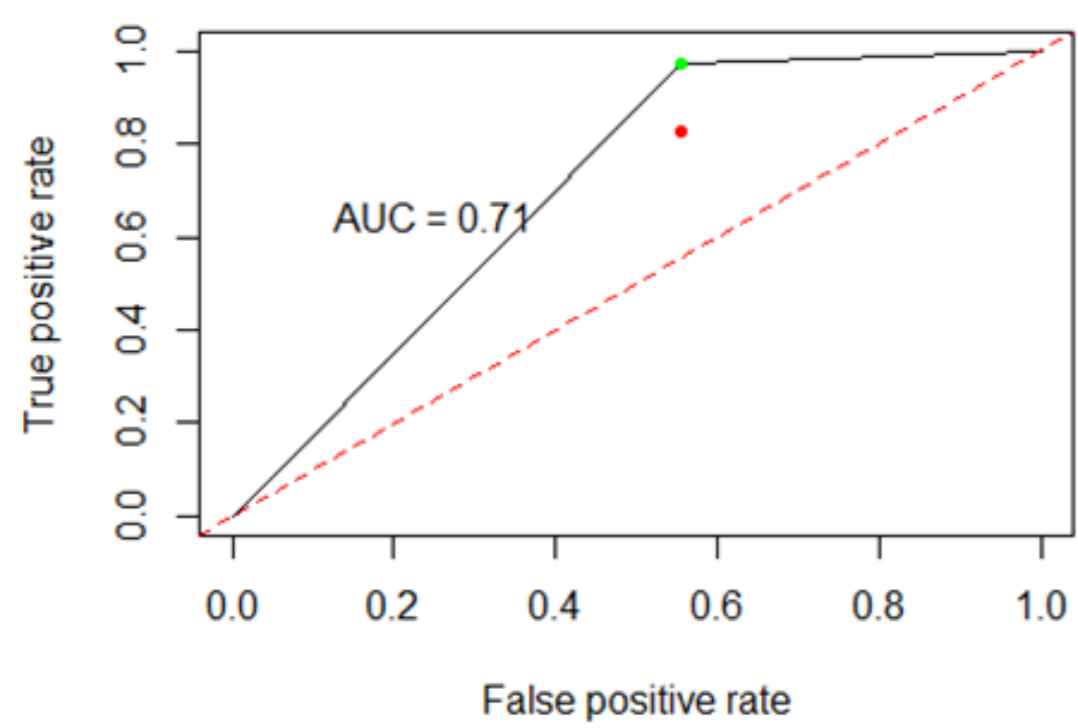


Figura 6. Curva ROC del modelo KNN con componentes principales. El área bajo la curva es 0,71. En rojo el punto óptimo y en verde el punto del modelo.

El modelo KNN es un algoritmo de machine-learning supervisado utilizado en problemas de clasificación y regresión.

## Resultados

- El análisis univariante muestra diferencias significativas en los valores de triglicéridos y en el número de hematies entre los grupos de degeneración discal. También se hallaron diferencias en la proporción de pacientes diagnosticados de dislipemia entre grupos ( $p<0.05$ ).
- El modelo de regresión logística muestra un aumento del riesgo relativo de degeneración discal en las variables triglicéridos, fumadores y edad. Con este modelo, se obtiene una sensibilidad del 83% y una especificidad del 78.6% en la clasificación de pacientes con discos degenerados (Figuras 2 y 3).
- Se realiza un modelo discriminante no paramétrico, mostrando una sensibilidad del 97% y una especificidad del 44% en la clasificación de los pacientes con degeneración discal según las variables recogidas, con un valor de área bajo la curva de 0.71.
- Se comparan las métricas de ambos modelos, consiguiendo mejores resultados en el análisis discriminante no paramétricos para clasificar los pacientes con las variables del estudio.

## Conclusiones

- Los dos modelos obtenidos permiten clasificar los pacientes con degeneración discal en función de datos obtenidos de análisis de sangre rutinarios, demográficos y antecedentes, presentando mejores métricas el modelo discriminante.
- Entre las limitaciones del estudio se encuentra que todos los pacientes presentan algún tipo de patología vertebral (traumática o no), el tamaño de la muestra, el número de variables predictoras y la heterogenicidad de la muestra.

	Reg. Logística	KNN-PCA
Exactitud	0,82	<b>0,88</b>
Precisión	<b>0,97</b>	0,89
Sensibilidad	0,83	<b>0,98</b>
Especificidad	<b>0,79</b>	0,44
F1-score	0,90	<b>0,93</b>

Figura 7. Métricas obtenidas en ambos modelos. El modelo KNN-PCA presenta una sensibilidad una exactitud y un F1-score muy elevados, permitiendo predecir los pacientes con degeneración a partir de las variables estudiadas sin pruebas de imagen.