

INTEGRACIÓN DE MODELOS PROBIT CON ANÁLISIS DE EFECTOS MARGINALES, VÍA A-RULES Y MODELOS DE META-APRENDIZAJE INTERPRETABLES PARA PREDECIR LA RETIRADA DE FÁRMACOS BIOSIMILARES EN PACIENTES REUMATOLÓGICOS, BASADOS EN INTELIGENCIA ARTIFICIAL

David Castro Corredor¹, Luis Ángel Calvo Pascual², E. C. Garrido Merchán².

¹ Servicio de Reumatología. Hospital General Universitario de Ciudad Real.

² Departamento de Métodos Cuantitativos. ICADE. Universidad Pontificia de Comillas.

Email: d.castrocorredor@gmail.com

INTRODUCCIÓN

Para predecir la retirada de medicamentos biosimilares en enfermos reumatológicos y entender las causas de la retirada, usamos una amplia variedad de enfoques metodológicos, aprovechando las ventajas de cada modelo. Por un lado, el modelo Probit, aunque tiene escaso poder predictivo, proporciona información detallada sobre la influencia de las variables independientes en la variable objetivo y sus efectos marginales. Por otro lado, el método A-rules data mining, a pesar de ser menos común en la literatura reumatológica, destaca por su capacidad para descubrir relaciones causa-efecto entre variables categóricas y predecir resultados experimentales de manera efectiva. Finalmente, exploramos el potencial de las técnicas de meta-aprendizaje explicables, que ofrecen una valiosa oportunidad para mejorar la interpretabilidad de los modelos de Machine Learning habituales.

OBJETIVOS

Analizar las causas de la retirada de medicación biosimilar en pacientes reumatológicos a través de variables demográficas e inflamatorias, calculando los efectos marginales de las variables influyentes en la retirada de biosimilares, identificar las mejores reglas de asociación y trabajando con métodos ML explicables y de meta-aprendizaje.

RESULTADOS

Se ha desarrollado un modelo Probit (Figura 1A) con variables significativas que exhiben un buen ajuste y ausencia de multicolinealidad (Figura 1B). La robustez del modelo se confirma mediante un bajo LLR p-value, indicando una mejora significativa sobre el modelo nulo, y mediante el análisis del Área ROC (Figura 1C).

Los efectos marginales revelan una mayor probabilidad de retirada del biosimilar en pacientes jóvenes con alta actividad inflamatoria a las 6 semanas y baja medicación con corticoides. (Figura 2).

Por otro lado, las reglas más importantes, mostradas en la Tabla 1 (figura 1ª), destacan la repetición de variables en la retirada de biosimilares en adultos de 30 a 60 años con alta actividad a los 6 meses y tratamiento con corticoides, presentando métricas de bondad significativas en ambas direcciones (Figura 3).

Finalmente, tras analizar diversos modelos clasificatorios sobre la variable objetivo RETIRADA del biosimilar, hemos observado la importancia explicativa de la variable actividad final, el empleo de corticoides y la edad. En modelos donde RETIRADA =1 (Fig 5A) o RETIRADA =0 (Fig 5B) vemos como en pacientes a los que se le quita o deja el biosimilar la explicación inflamatoria es la más importante, teniendo diferentes respuestas en función de la edad.

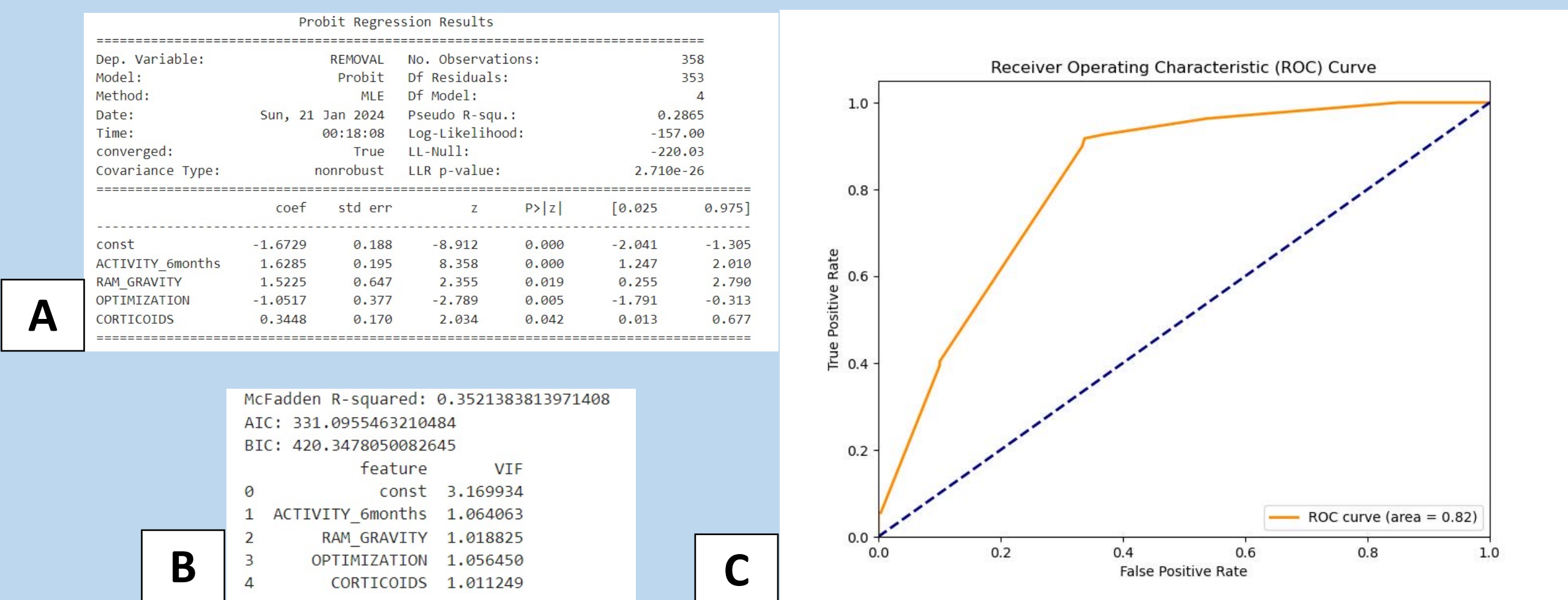


Figura 1. Modelos Probit. (A) Tabla resumen con las variables significativas, sus coeficientes, intervalo de confianza, etc. (B) Métricas y análisis de la multicolinealidad con VIF. (C) Curva ROC.

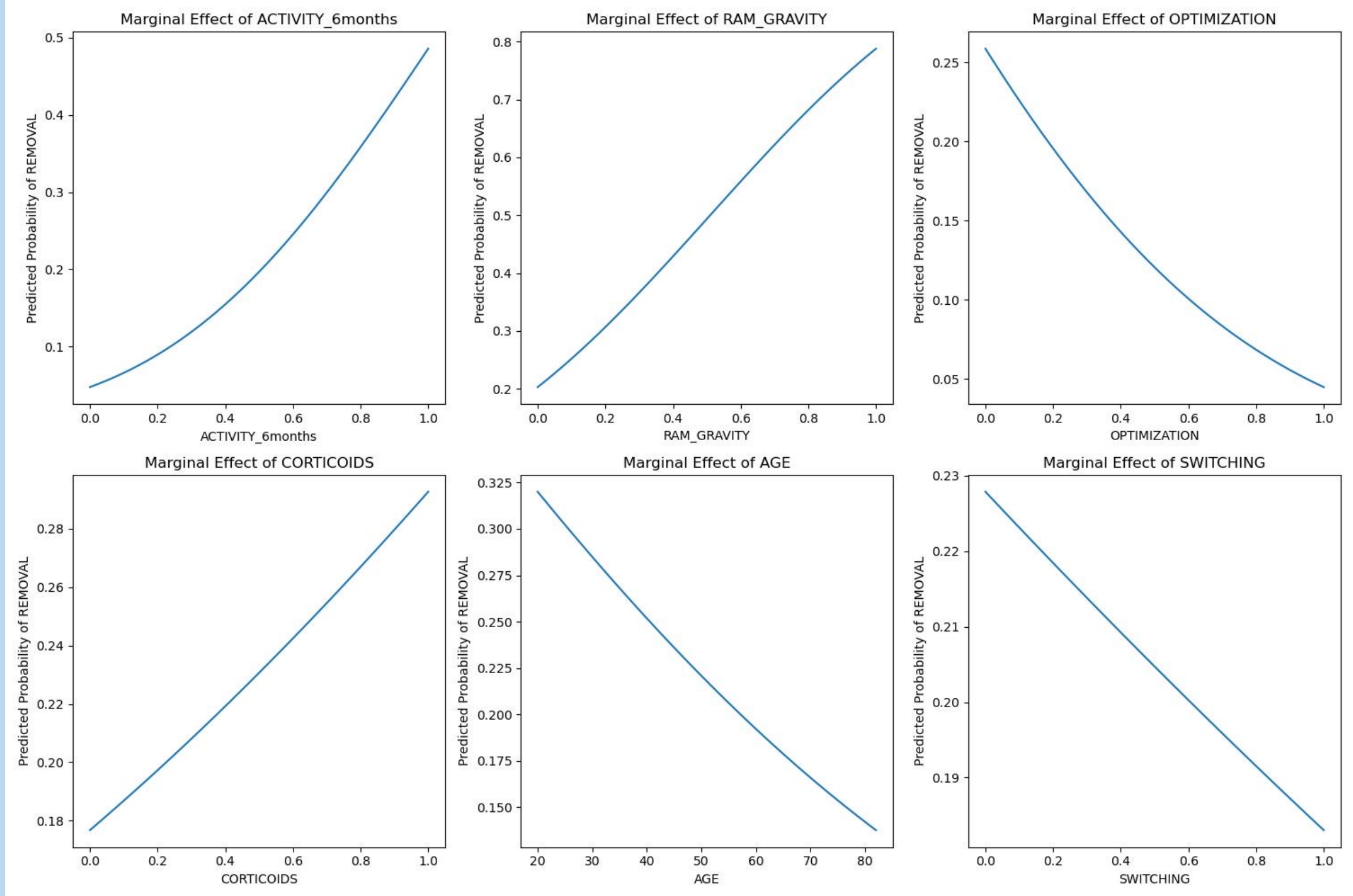


Figura 2. Efectos marginales de las principales variables sobre la retirada de biosimilares.

CONCLUSIONES

- Todos los modelos dan como resultado una relación directa entre la alta actividad inflamatoria en pacientes de mediana edad (30-60) y el tratamiento previo con corticoides, con la retirada del fármaco biosimilar.
- Otra variable que jugó un papel relevante en los diferentes modelos fue la retirada gradual del biosimilar (OPTIMIZATION), que tiene efectos óptimos disminuyendo la inflamación y confirmando que la retirada era la mejor opción.

MATERIAL Y MÉTODOS

Estudio descriptivo observacional y retrospectivo. La población son 365 pacientes con enfermedades reumáticas inflamatorias que cambiaron de un fármaco biológico a un fármaco biosimilar en el período Febrero 2022 a Noviembre de 2023.

En este estudio, implementamos un análisis Probit para evaluar la retirada del biosimilar como variable independiente, evaluando la robustez del modelo a través de métricas como el pseudo R-cuadrado de McFadden, el AIC, el BIC y la precisión de la regla de la mayoría, mientras abordamos la multicolinealidad mediante el Factor de Inflación de la Varianza (VIF). Además, analizamos los efectos marginales de las variables más significativas en relación con la retirada de los biosimilares. Posteriormente, empleamos el algoritmo A-priori para calcular diferentes reglas de asociación (A-rules), utilizando métricas como Soporte, Confianza y Lift, destacando la métrica de Zhang para medir la causalidad entre antecedente y consecuente. Por otro lado, empleamos técnicas de aprendizaje automático interpretable, como los valores de Shapley y LIME, junto con clasificadores de conjunto, optimizados mediante una optimización bayesiana en la validación cruzada, para obtener un modelo más efectivo para predecir y entender la retirada del biosimilar. Finalmente, aplicamos un enfoque data-driven de aumento de datos y utilizamos modelos como Random Forest y métodos bayesianos para estimar la incertidumbre en la versión final del modelo.

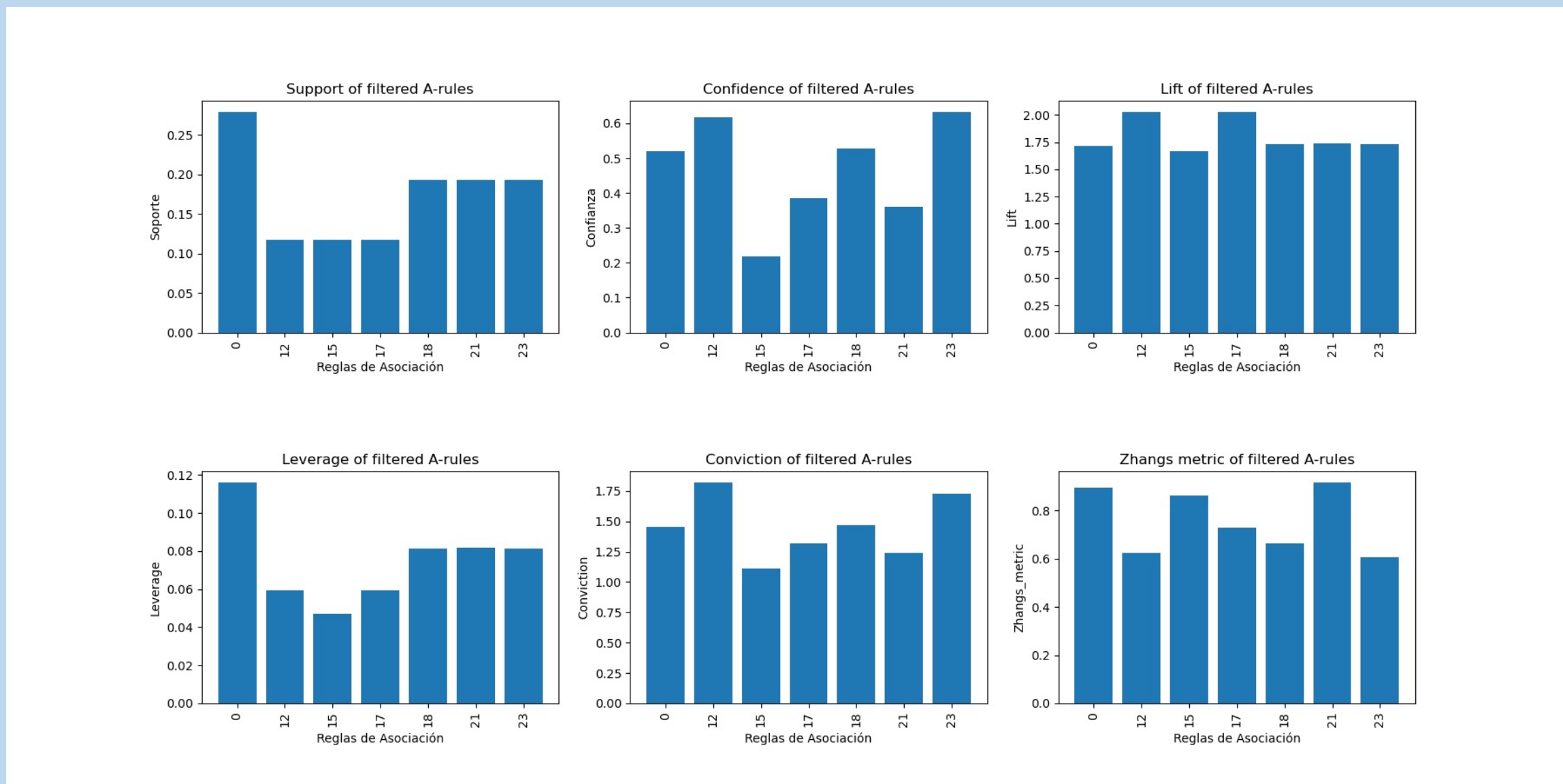


Figura 3. Principales métricas de las reglas de asociación filtradas.

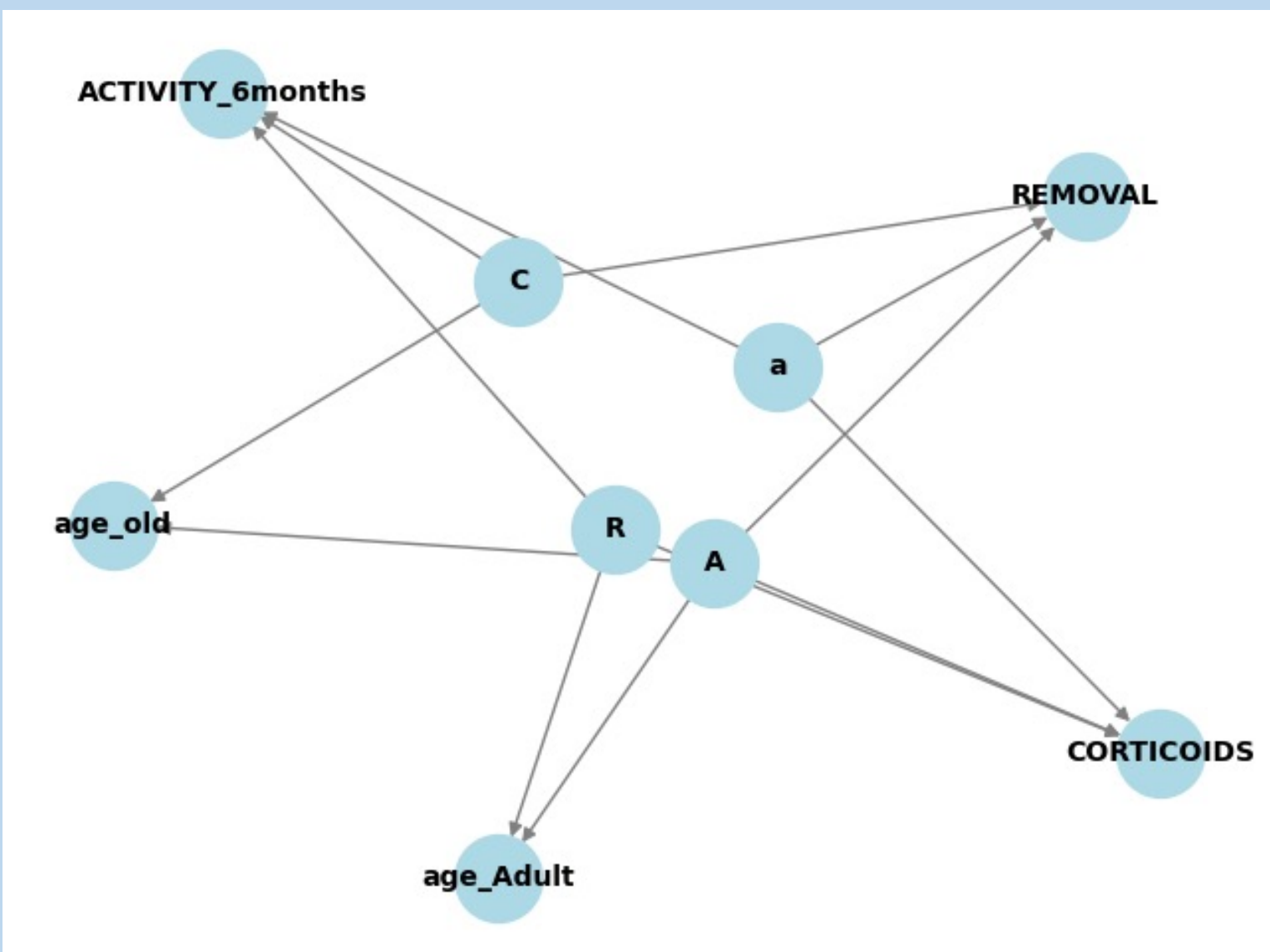


Figura 4. Association rule network, gráfico explicativo de cómo se relacionan las reglas de la figura 3.

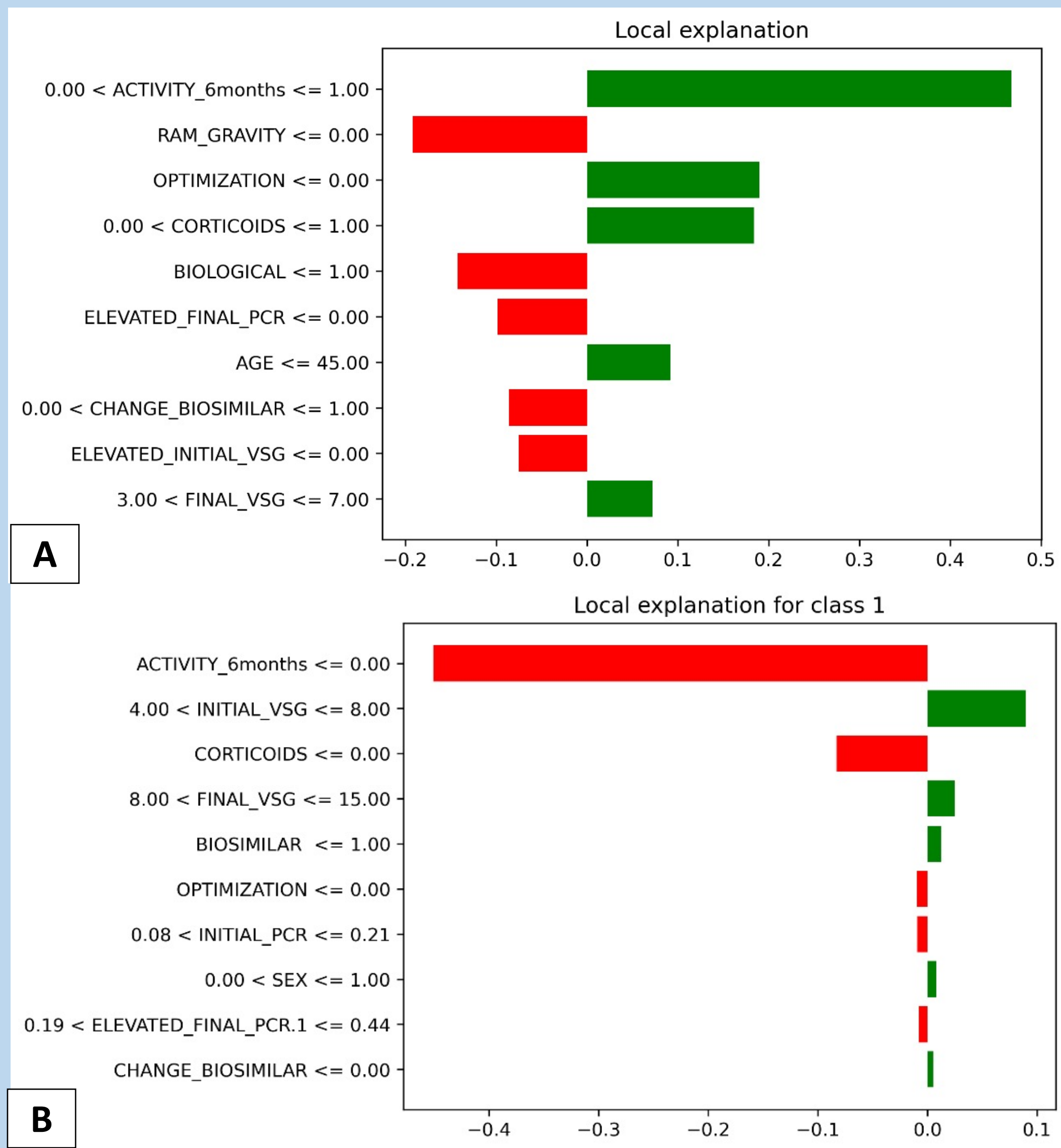


Figura 5. Valores LIME explicativos de RETIRADA. (A) Para una regresión logística y un paciente aleatorio que tuvo REMOVAL del biosimilar. (B) Para un modelo Random over sampler + Decision tree y un paciente aleatorio que NO tuvo REMOVAL del biosimilar.

