

Integración de la inteligencia artificial para la predicción personalizada de supervivencia en Oncología Radioterápica

Héctor Miras del Río

*Especialista en Radiofísica Hospitalaria
Hospital Universitario Virgen Macarena, Sevilla*

JORNADAS SARH

Inteligencia Artificial Aplicada
a la Física Médica



Parador de Antequera
14 de noviembre 2025

Contenido

1. Presentación del proyecto
2. Sistema de información de tratamientos radioterápicos en el HUVVM
3. SEQUOIA-RT, descripción y resultados preliminares.

Proyecto Investigación



Convocatoria: PI-SEFM-2024

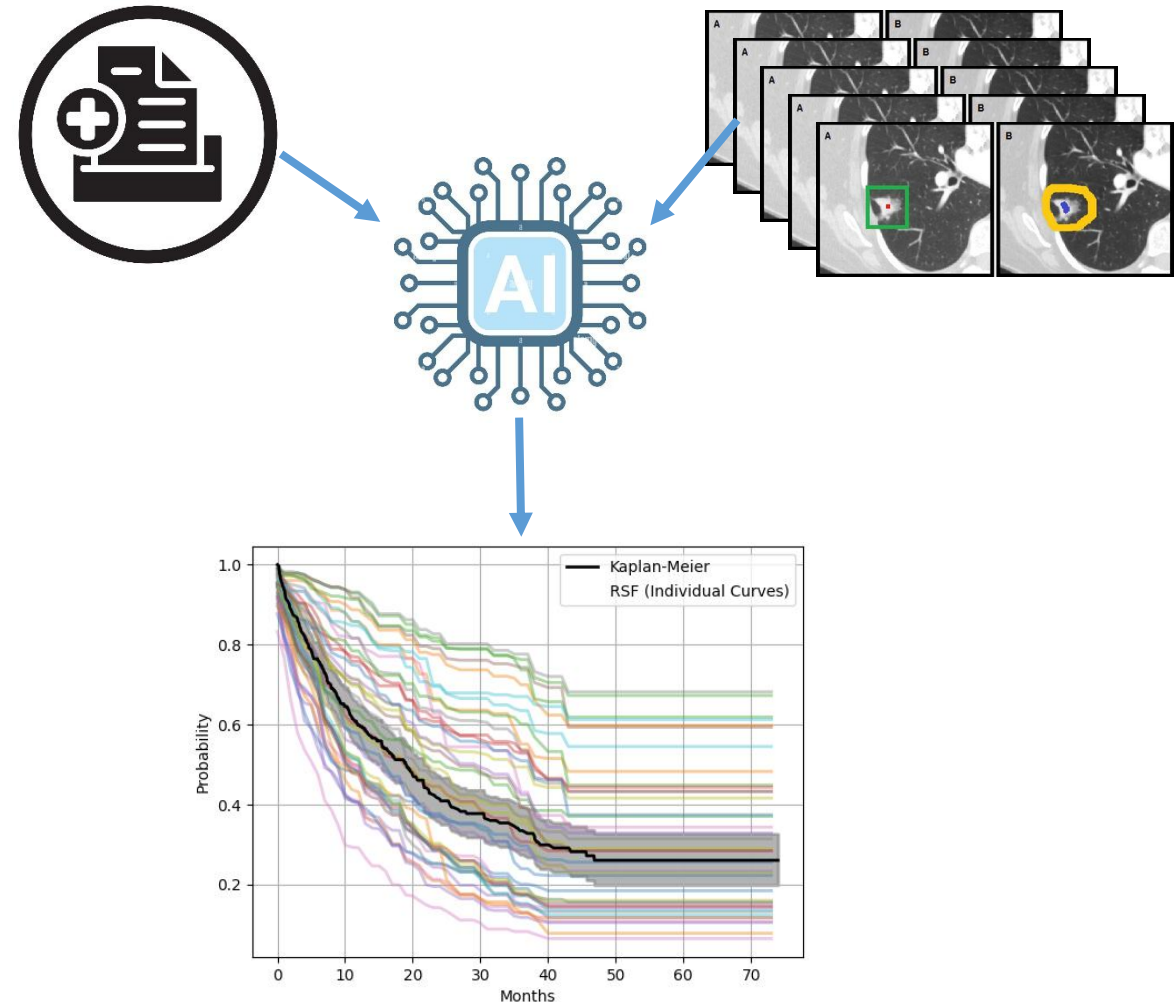
Duración: 2024-2026

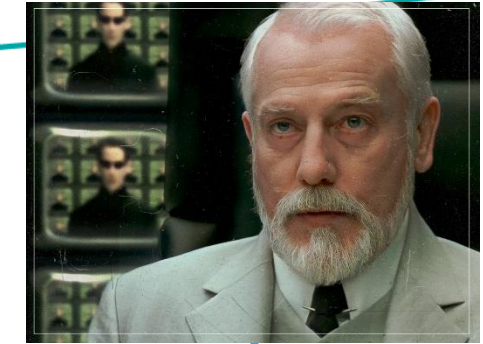
Centro coordinador: Hospital Universitario
Virgen Macarena

Entidades: FISEVI, Biogipuzkoa

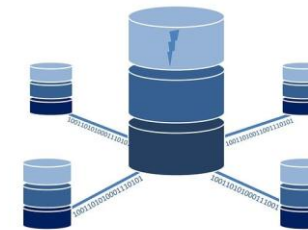
Financiación: 4.000 € → infraestructura
informática

Objetivo: Aplicar inteligencia artificial para la
predicción personalizada de supervivencia en
pacientes tratados con radioterapia mediante la
integración de datos clínicos y radiómicos.





VPN

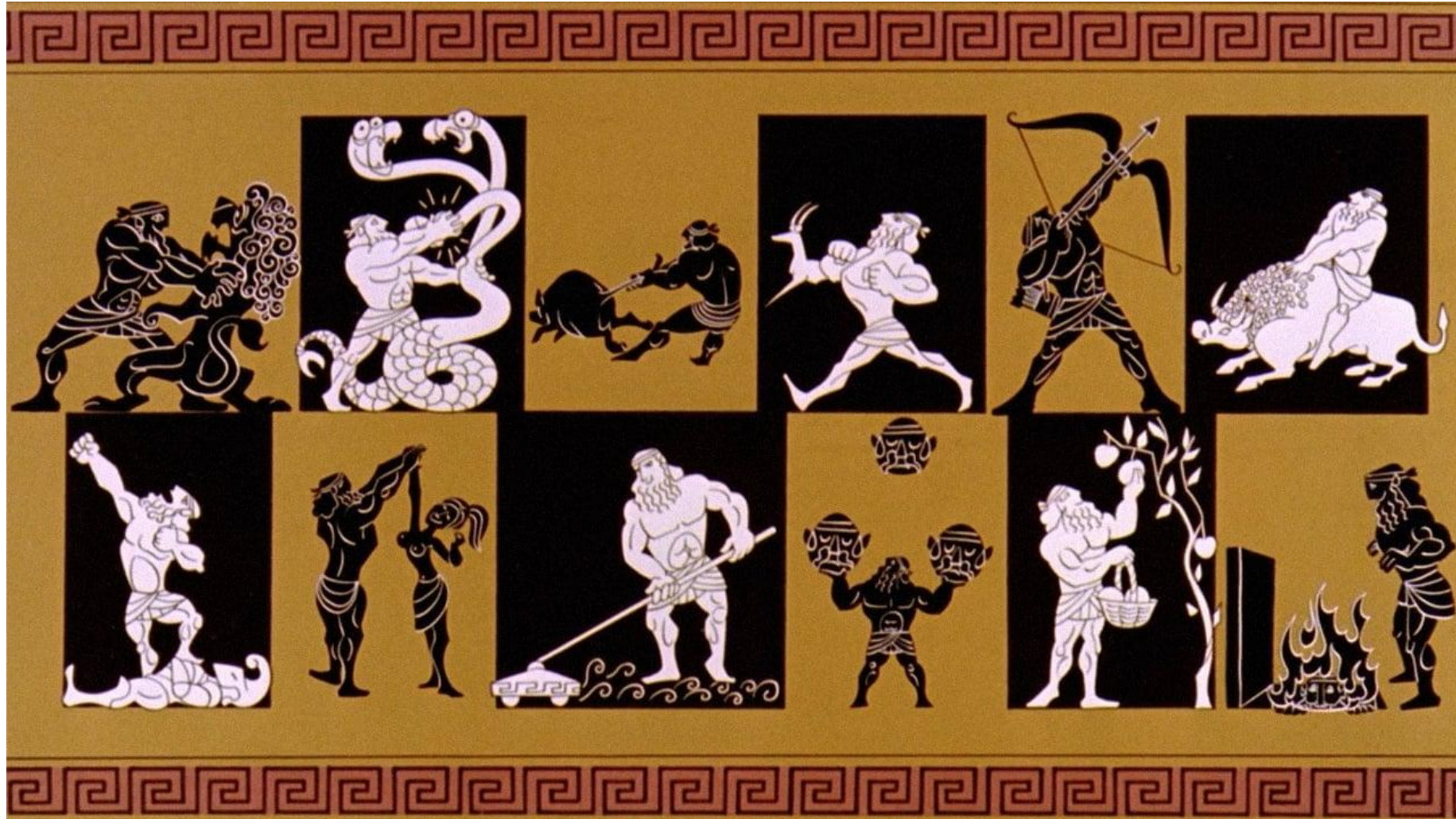


Pasos administrativos y éticos

Este proyecto tiene una serie de particularidades que dificultan todo el proceso burocrático:











- Investigación con datos clínicos
- Necesidad de acceso a sistemas corporativos del SSPA para obtener datos que van a usarse con fines de investigación: DIRAYA, PACS
- Instalación de un equipo informático con configuración no estándar para el SAS (Ubuntu)
- Investigador externo, que además es el IP: Necesidad de acceso a datos y conexión remota al equipo de investigación

Pasos administrativos y éticos



Pasos administrativos y éticos



Etapa	Descripción	Entidad responsable
  Dictamen ético	Evaluación y aprobación del proyecto por el Comité de Ética de la Investigación de Sevilla	CEI Sevilla
 Convenio FISEVI-SEFM	Formaliza la financiación y gestión administrativa del proyecto	FISEVI / SEFM
 Convenio FISEVI-BIOGIPUZKOA	Regula la colaboración con investigador co-IP externo	FISEVI / Biogipuzkoa
 Adquisición de equipamiento	Gestión de compra de equipamiento informático	HUVM / FISEVI
 Instalación y configuración	Instalación equipo en la red del HUVM. Configuración VPN para acceso de investigador externo.	Servicio TICs HUVM
 Acceso a datos SSPA (Res. 1/2021)	Solicitud formal con: protocolo, consentimiento informado, EIPD, declaraciones de confidencialidad y plan de destrucción de datos	HUVM / SAS / Consejería Salud
 Extracción de datos	Solicitud de acceso y extracción de datos de PACS y Historia clínica	Dep. Innovación / TICs HUMV
  Anonimización y control	Anonimización independiente antes del análisis IA	Servicio de Doc. Clínica HUVM

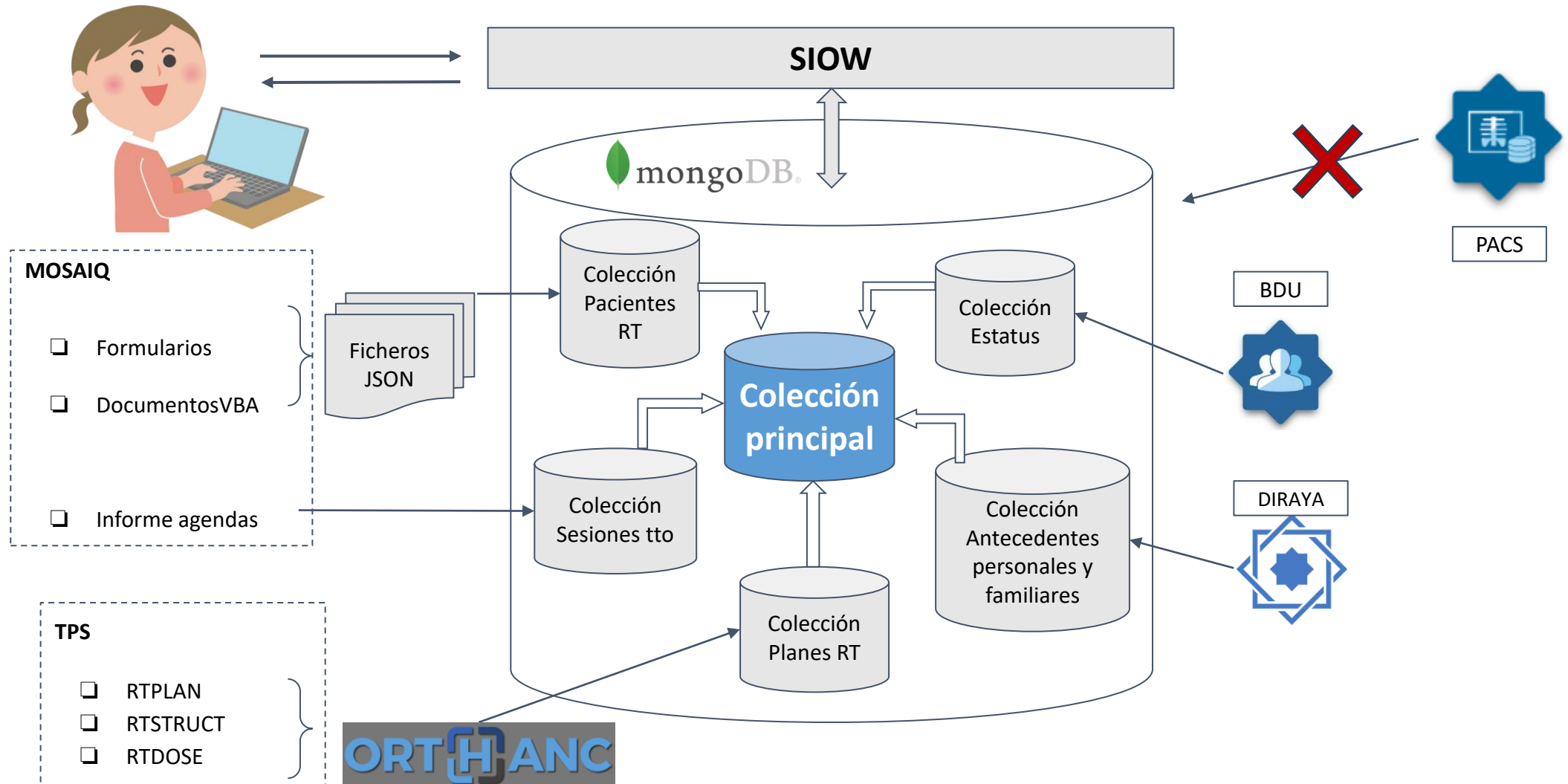
Investigar con datos de salud

Mi experiencia personal:

- Burocracia desproporcionada
- Poca formación de investigadores en temas de protección de datos y seguridad informática
- Falta de apoyo a las unidades para trámites burocráticos
- Alta exposición a riesgo de los investigadores



SIOW: Sistema de Información Oncológica Web



Implementación de flujo paperless en Mosaic

Descripción
00 Sólo primera consulta
000 INTERCONSULTA
01 primera consulta
02 cita TAC Sim
03 Crear hoja Sim
04 Inmovilización y TAC SIM
05 Foto ID y posicionamiento
06 Completar hoja Simulación
07 Exportar TC a TPS y PACS
08 Importar y Traducir TC & Crear hoja de Inicio
08.1 Traducir TAC a Pinnacle
08.2 Delimitación OARs
09 Delimitación de volúmenes
10 Prescripción Dosimetría
11.0 Asignar nivel de prioridad a diseño tto
11.1 DISEÑO TTO
11.2 Revisión del tratamiento
12 Verificación TTO
12.1 Subir informe verificación a MOSAIQ
13.1 Exportación
13.2 Documentación
13.3 Importación MOSAIQ / DIRAYA
14 Aprobar PLAN TTO
15 EN ESPERA TTO
15.1 INFORMACION PARA INICIAR RT
16 Agendar en maquina TTO
16.1 Calendario tratamiento
17 AVISAR paciente para inicio
17.1 GESTION TRANSPORTE
18 Primera consulta enfermería
19 SETUP e INICIO TTO
20 Sesión de tratamiento
21 Evaluar error sistemático
22 Revisión de imagen



ELSEVIER

International Journal of Medical Informatics

Volume 144, December 2020, 104301

Organic generation of real-world real-time data for clinical evidence in radiation oncology

A. Bertolet ^{a, b} ✉, A. Wals ^a, H. Miras ^a, J. Macías ^a

Show more

Share Cite

Flujo tare <https://doi.org/10.1016/j.ijmedinf.2020.104301>

Formularios

Protocolo TTO

Protocolo TTO (PRH) - Próstata Hipofrac.

Se cambiarán Datos de prescripción

25 fx - NUHSA: RF09122018 DEL LAGO, LANZAROTE

Fecha	Día	Hora	Ionado con	Actividad	Duración	Estado
11/05/2018	11	11:00		TTO+CI	0.15	
11/05/2018	11	11:00		TTO+CI	0.15	
11/05/2018	11	11:00		TTO+CI+ES	0.15	
11/05/2018	11	11:00		TTO+CI	0.15	
11/05/2018	11	11:00		TTO	0.15	
11/05/2018	11	11:00		TTO	0.15	
11/05/2018	11	11:00		TTO	0.15	
11/05/2018	11	11:00		TTO+CI	0.15	
11/05/2018	11	11:00		TTO	0.15	
11/05/2018	11	11:00		TTO	0.15	
11/05/2018	11	11:00		TTO	0.15	
11/05/2018	11	11:00		TTO+CI	0.15	
11/05/2018	11	11:00		TTO	0.15	
11/05/2018	11	11:00		TTO	0.15	



OTORRINOLINGÜÍSTICA - RADIOFÍSICA DE SEVILLA

posicionamiento del paciente

PAC000000000 3 Mos Transgener

Amadeo Jesus

Datos planificación	
Fecha	11/05/2018
Hora	11:00
Datos del plan	
Técnica de Tto	IMRT
Aplicador	HiRadus
Bolus	SI
Cabeza Gantry	Exterior Bolus (cm)
118	Campo NC
	Radiofísico Plan
	Dif. Mercaderes cm. Gantry 270
	Dif. Mercaderes cm. Long

5 campos

Int	Error sistemático	Desplazamientos corregidos	Referencia Paciente
Longitudinal (X)	-0.8	-0.8	CERECHE
Vertical (Y)	-1.8	1.2	SUPERIOR
	-1.4	-0.4	POSTERIOR

Documentos
eScribe

Get rights and content

LLMs para extracción de datos de la historia clínica

JMIR MEDICAL INFORMATICS

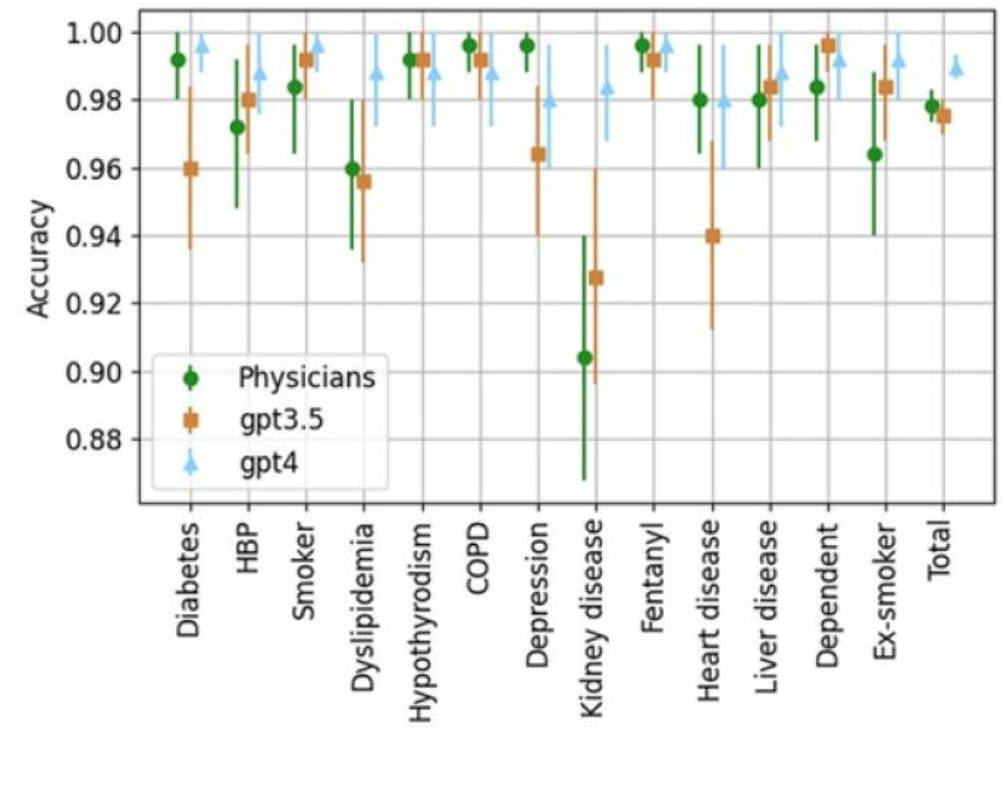
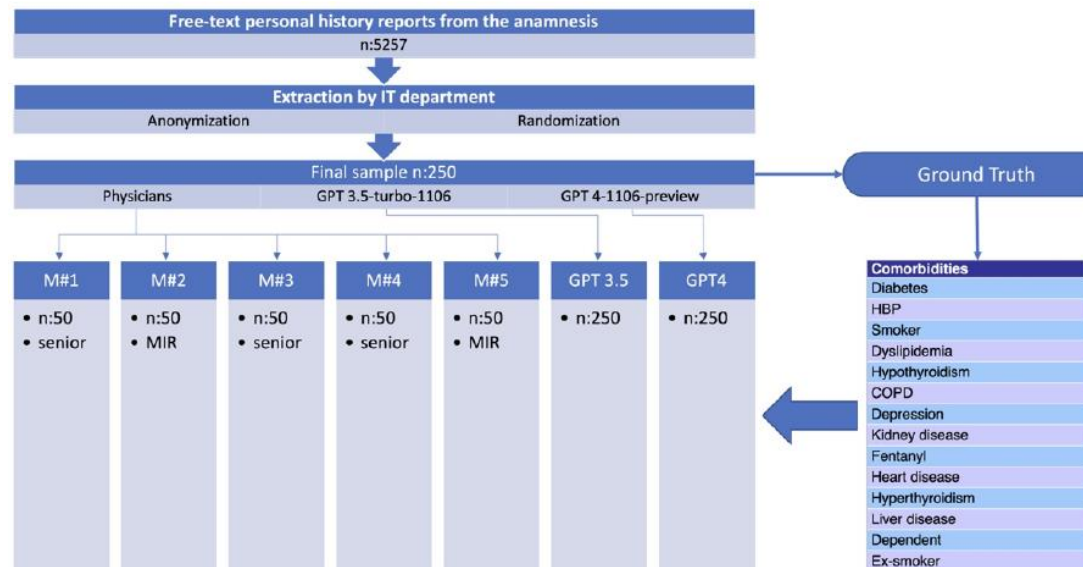
Wals Zurita et al

[Original Paper](#)

The Transformative Potential of Large Language Models in Mining Electronic Health Records Data: Content Analysis

Amadeo Jesus Wals Zurita, MD; Hector Miras del Rio, MP; Nerea Ugarte Ruiz de Aguirre, MD; Cristina Nebrera Navarro, MD; Maria Rubio Jimenez, MD; David Muñoz Carmona, MD, PhD; Carlos Miguez Sanchez, MD

Servicio Oncología Radioterápica, Hospital Universitario Virgen Macarena, Andalusian Health Service, Seville, Spain



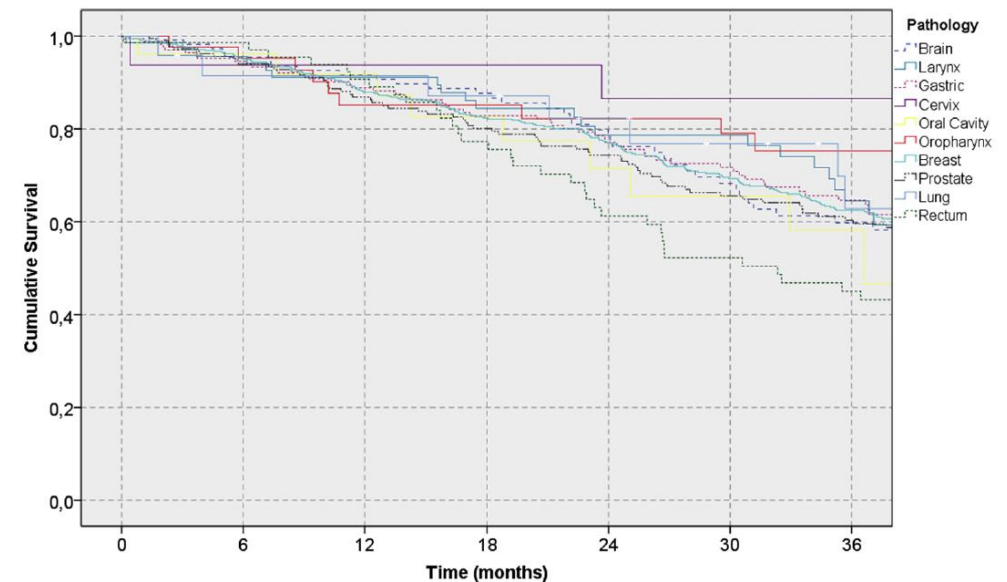
LLMs para extracción de datos de la historia clínica

- Alta **precisión y eficiencia** en la extracción de información estructurada a partir de textos clínicos.
- Gran potencial para **impulsar la investigación basada en datos del mundo real (RWD)**.
- **Ausencia de validación clínica:** cada aplicación debe evaluarse específicamente. Los resultados deben **interpretarse con cautela** y bajo supervisión experta.
- Los modelos de mayor rendimiento suelen **depender de proveedores externos**, con riesgos legales y de privacidad.
- Surgen **alternativas locales:** instalación en hardware accesible, con resultados de calidad aceptable y mayor control sobre los datos.

¿Y ahora qué?

- BD radioterapia con > 10.000 pacientes desde el año 2018 (6.400 con datos dosimétricos)
- Esto ya nos permite generar nuestros resultados en salud, como curvas de supervivencia
- Pregunta que nos hacemos:

¿Puede la IA ayudarnos a hacer modelos predictivos individualizados al paciente?



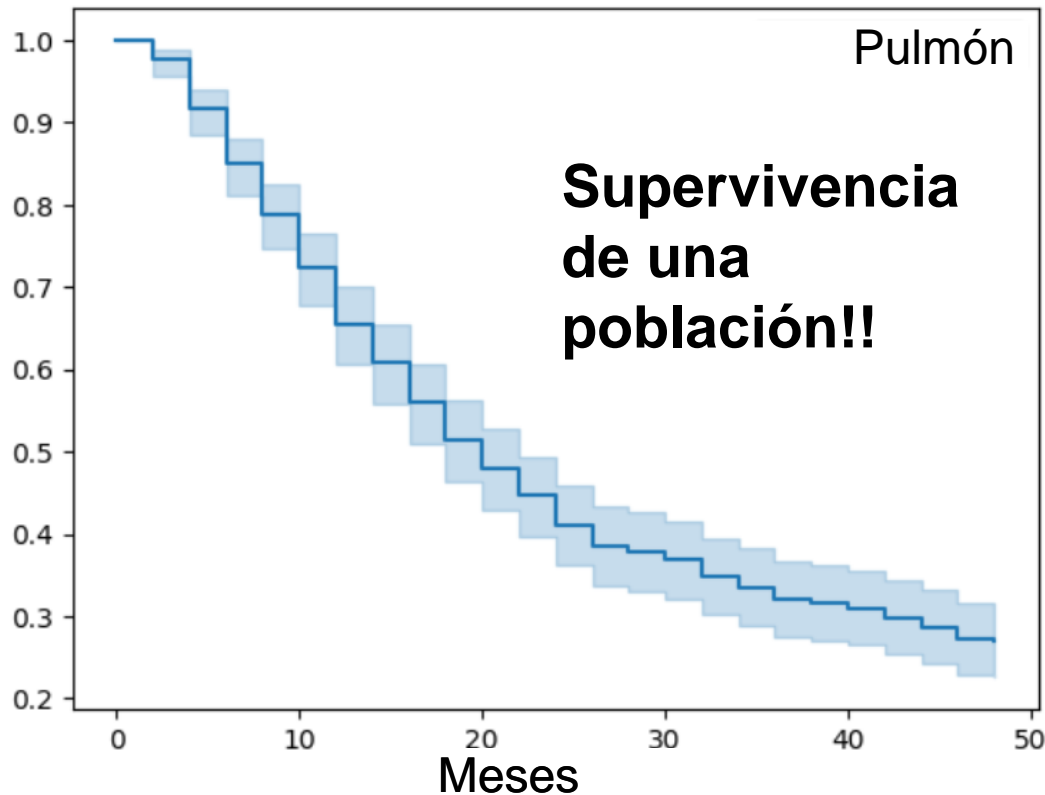


Estimador Kaplan-Meier (método no paramétrico)

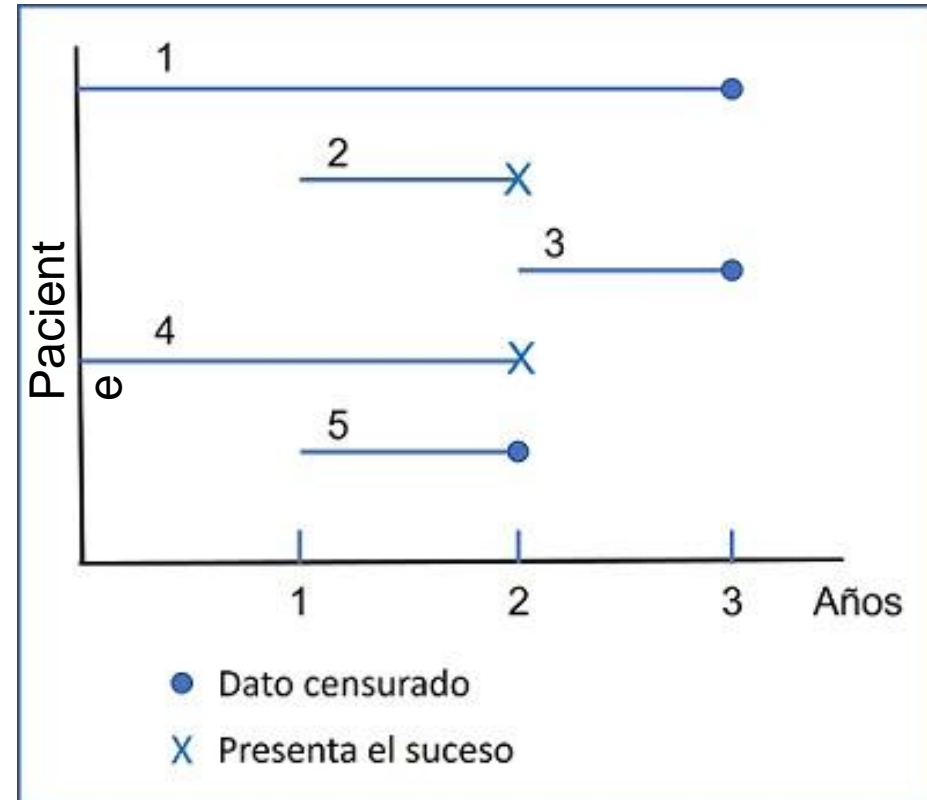
$$\hat{S}(t) = \prod_{i:t_i < t} \left(1 - \frac{d_i}{n_i}\right)$$

d_i : num. eventos en t

n_i : num. pacientes en t



Regresión de Cox (método paramétrico)



Análisis de Supervivencia Individualizada

- No es posible mediante métodos habituales
- No es posible conocer la distribución de probabilidad de supervivencia de individuos específicos de forma empírica
- Problema altamente no lineal
- Dependiente de un gran número de parámetros, a priori, desconocidos.

Análisis de Supervivencia Individualizada

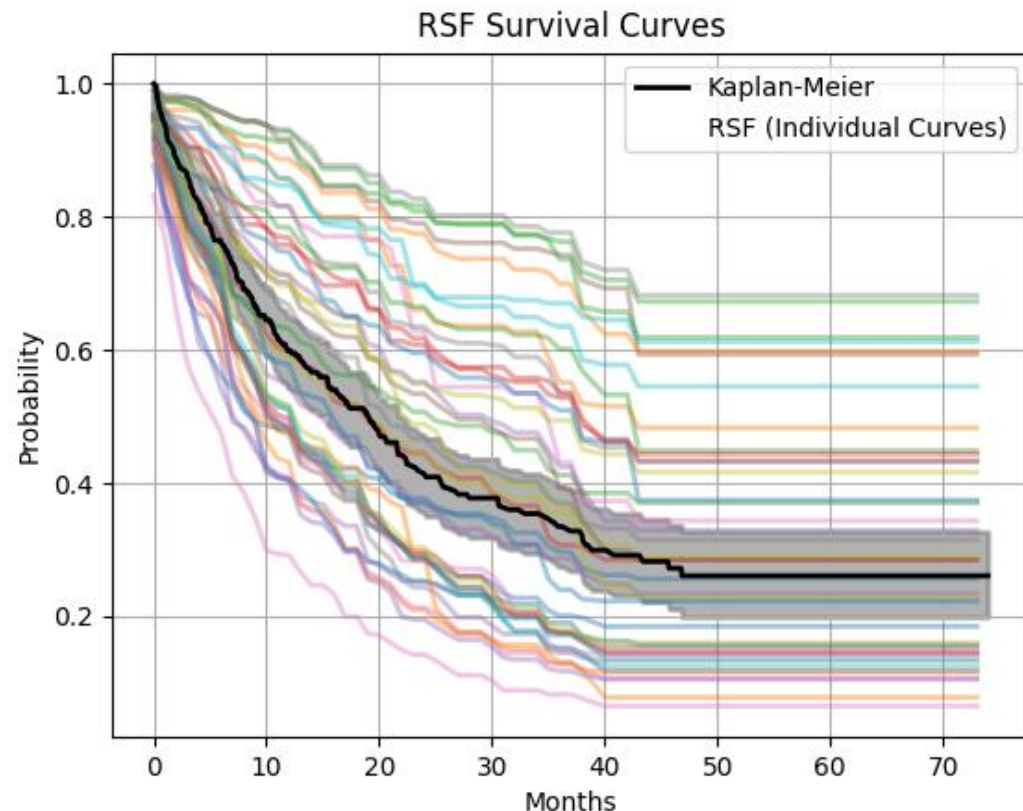
- No es posible mediante métodos habituales
- No es posible conocer la distribución de probabilidad de supervivencia de individuos específicos de forma empírica
- Problema altamente no lineal
- Dependiente de un gran número de parámetros, a priori, desconocidos.



**Solución mediante IA:
no requiere conocer las
curvas de supervivencia
individualizadas!!**

- Predicción de tiempo de evento: DeepHit, Deep AFT, SurvFormer...
- Clasificación en paciente equivalente: RSF, KNN survival, GradientBoosting survival...

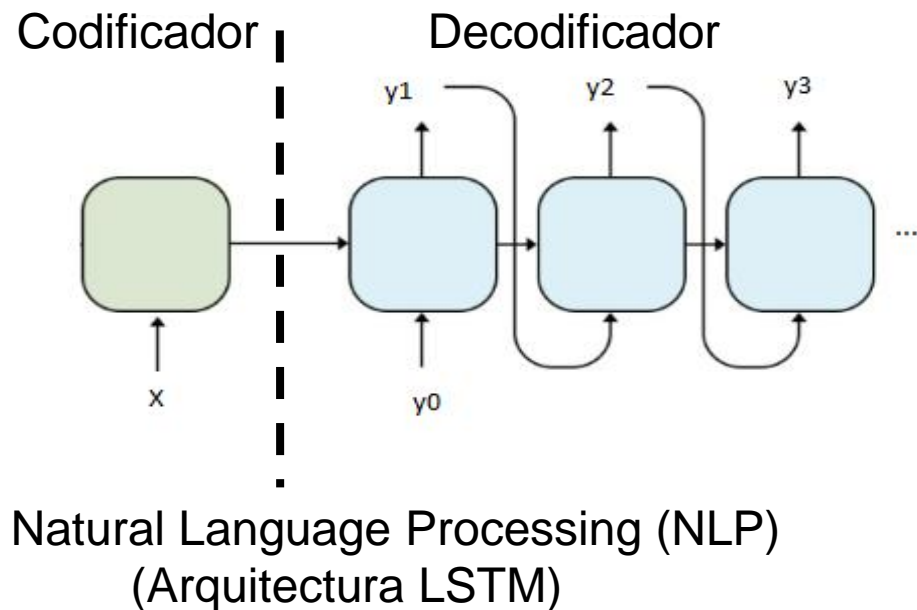
Árboles de Supervivencia



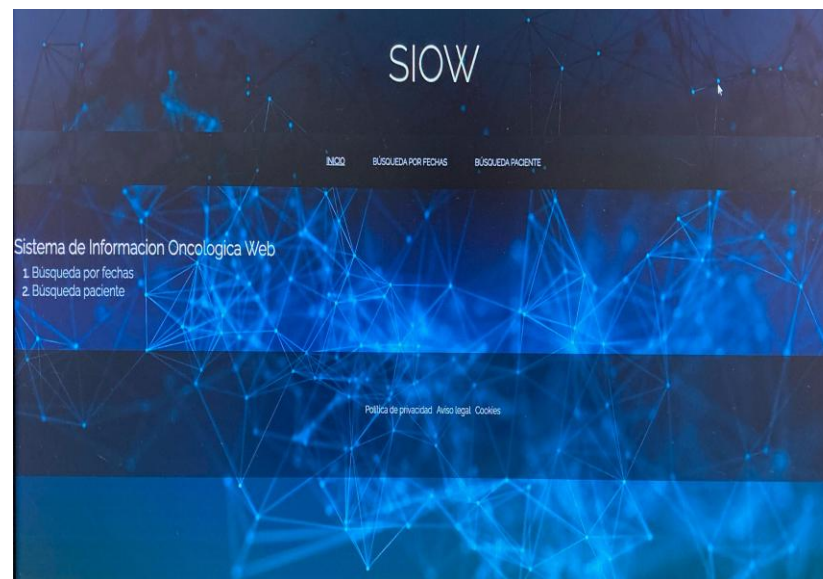
- Promedio de pacientes “equivalentes” tiende a suavizar no linealidades
- Inestables ante datos ruidosos (radiómica)
- Tienden al sobreajuste
- No manejan multicolinealidad adecuadamente
- No predicen factores pronóstico conocidos

SEQUOIA-RT:

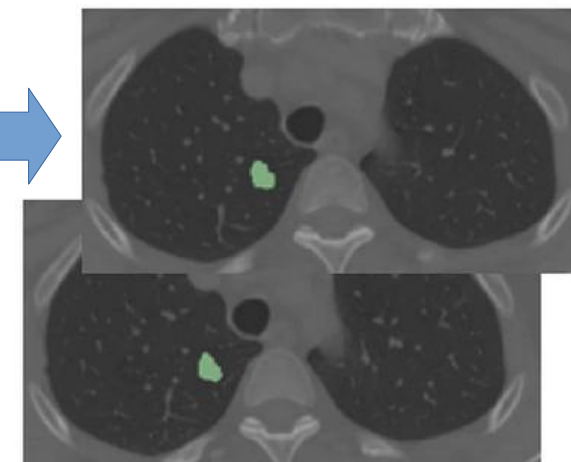
Survival Estimation using QUantitative Oncologic data with Individualized Analysis for Radiotherapy



- Base de Datos



- Imágenes Radiológicas



Datos Clínicos

	Edad	Genero	Estatus	Diabetes	Fentanilo	Hepatopatía	Hipotiroidismo	Dislipemia	Renal	Cardiopatía	...	VolPTV1	VolPTV2	VolPTV3	NumeroSesiones	OverallTime	Event	DiasDemora	DiasSupervivencia	DiasRastreo
848	74	1	0	0	0	0	0	0	0	0	1 ...	127.494000	127.494000	127.494	20	33	1	52	88	1370
849	75	1	0	0	0	0	0	1	0	0	...	457.574986	439.519956	0.000	10	15	1	33	181	1407
854	58	0	0	0	0	0	0	0	0	0	...	457.574986	439.519956	0.000	5	7	1	19	87	1429
860	69	1	0	1	0	0	0	0	0	0	...	457.574986	0.000000	0.000	30	51	1	36	866	1371
861	52	0	1	0	0	0	0	0	0	0	...	291.303000	0.000000	0.000	20	29	0	21	1408	1408
862	61	1	0	0	0	0	0	0	0	1	...	457.574986	0.000000	0.000	10	15	1	23	63	1420
863	61	1	0	0	0	0	0	0	0	1	...	457.574986	0.000000	0.000	3	7	1	49	45	1402
867	72	1	0	1	1	0	0	0	0	0	...	276.507000	0.000000	0.000	14	29	1	72	449	1413

Selección de la muestra

- Consulta en SLOW de pacientes con diagnóstico de **cáncer de pulmón** tratados completamente entre los años 2018 y 2024 → **342**
- Búsqueda y descarga manual de TACs en el PACS
- Exclusiones:
 - Casos con diagnóstico de pulmón que realmente se tratan de otra cosa: holocráneo, metástasis, ...
 - Casos sin datos dosimétricos
 - Casos en los que no se encuentra el TAC en el PACS
- Muestra final: **271**

El problema de la validación

Para cada individuo:

- El modelo predice una distribución de probabilidad
- El evento observado sólo ocurre una vez



El problema de la validación

Capital: Antequera (altitud: 570 m)

Latitud: 37° 0' 51" N - **Longitud:** 4° 33' 28" O - **Posición:** Ver localización

Zona de avisos: Antequera

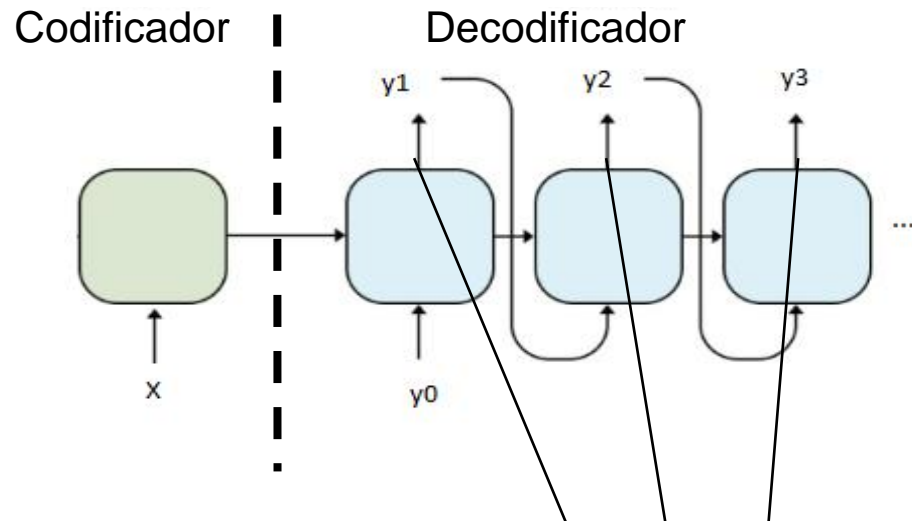
Descargar XML de la predicción detallada de Antequera [XML](#)

jue. 13		vie. 14				sáb. 15		dom. 16		lun. 17	mar. 18	mié. 19
12-18 h	18-24 h	00-06 h	06-12 h	12-18 h	18-24 h	00-12 h	12-24 h	00-12 h	12-24 h			
												
16°C	14°C	13°C	16°C	13°C	12°C							
Probabilidad de precipitación												
90%	100%	100%	100%	100%	100%	100%	100%	95%	90%	100%	30%	45%

Reglas de puntuación estrictamente propias

- Función que da una **puntuación** a una predicción probabilística según lo que realmente ocurrió
- Es estrictamente propia si la **única** manera de alcanzar la puntuación óptima es predecir exactamente la distribución verdadera

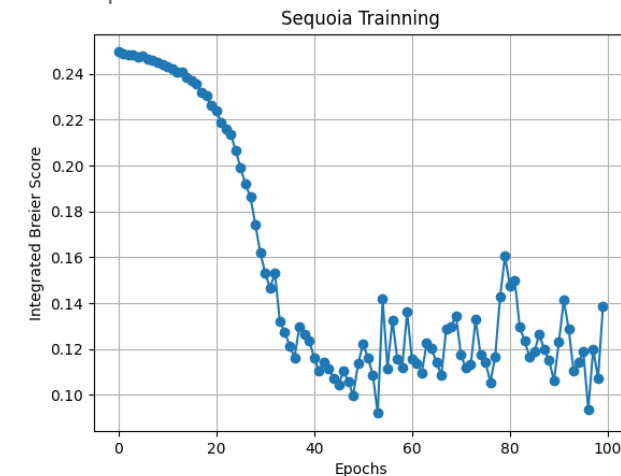
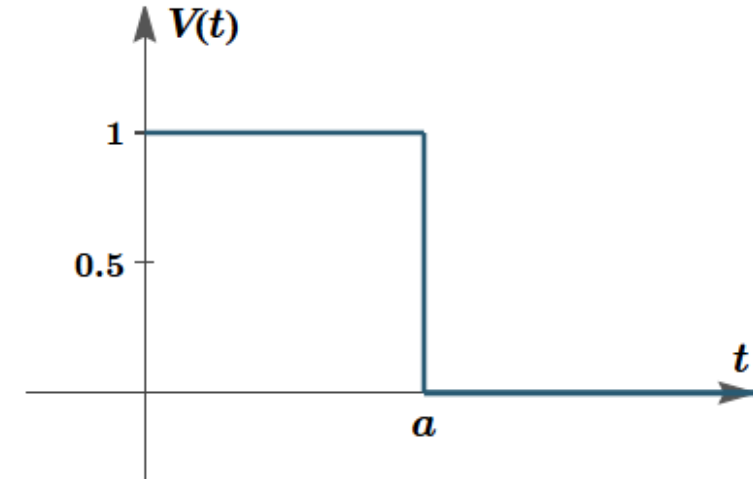
Metricas de Validación



- Metric: Integrated Brier Score (IBS)

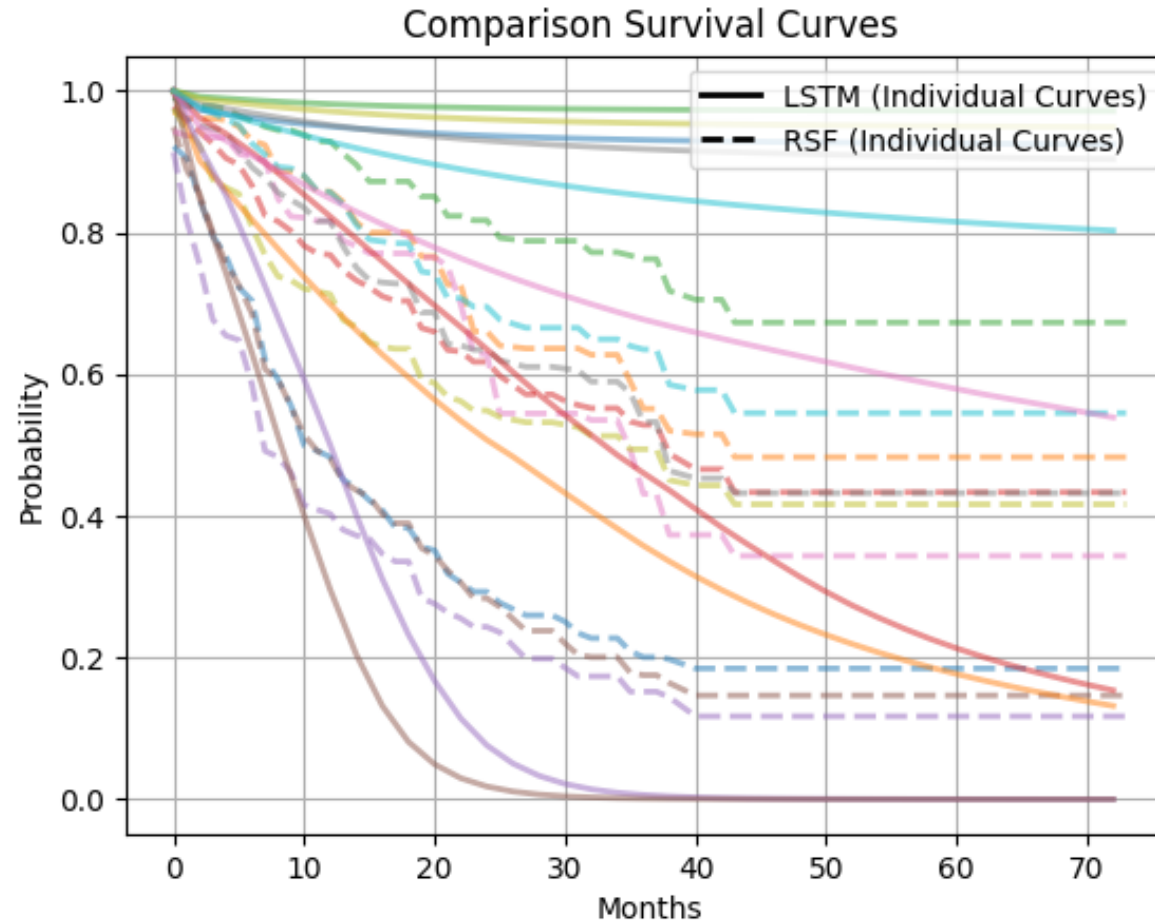
$$IBS = \frac{1}{t_{max} - t_0} \int_{t_0}^{t_{max}} BS(t') dt'$$

Arrows point from t_0 , t_{max} , and t' in the equation to the corresponding labels in the diagram above.



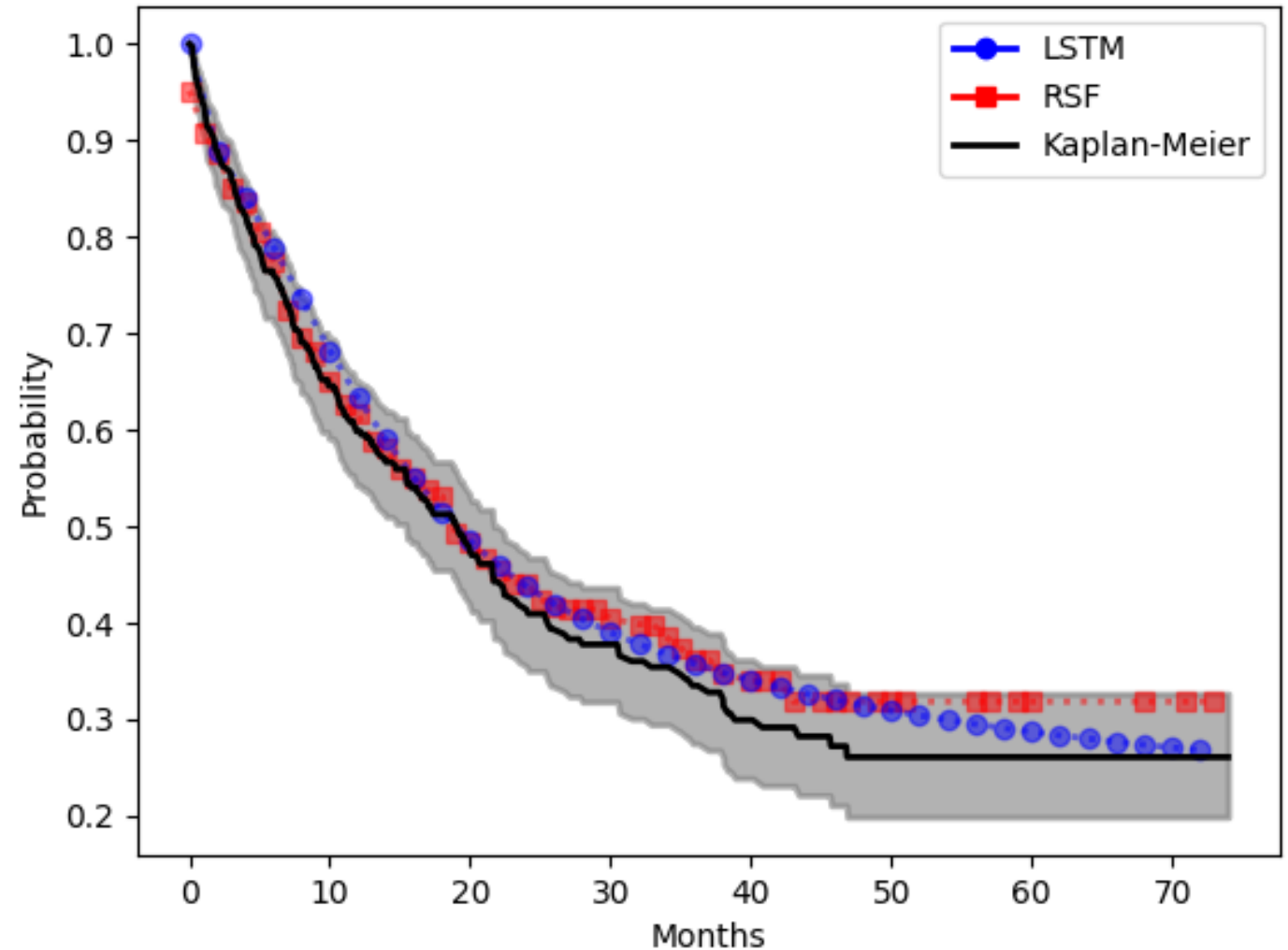
- Entrenamiento con 229 pacientes
- Validacion con 41 pacientes

Resultados Preliminares: Sequoia vs RSF

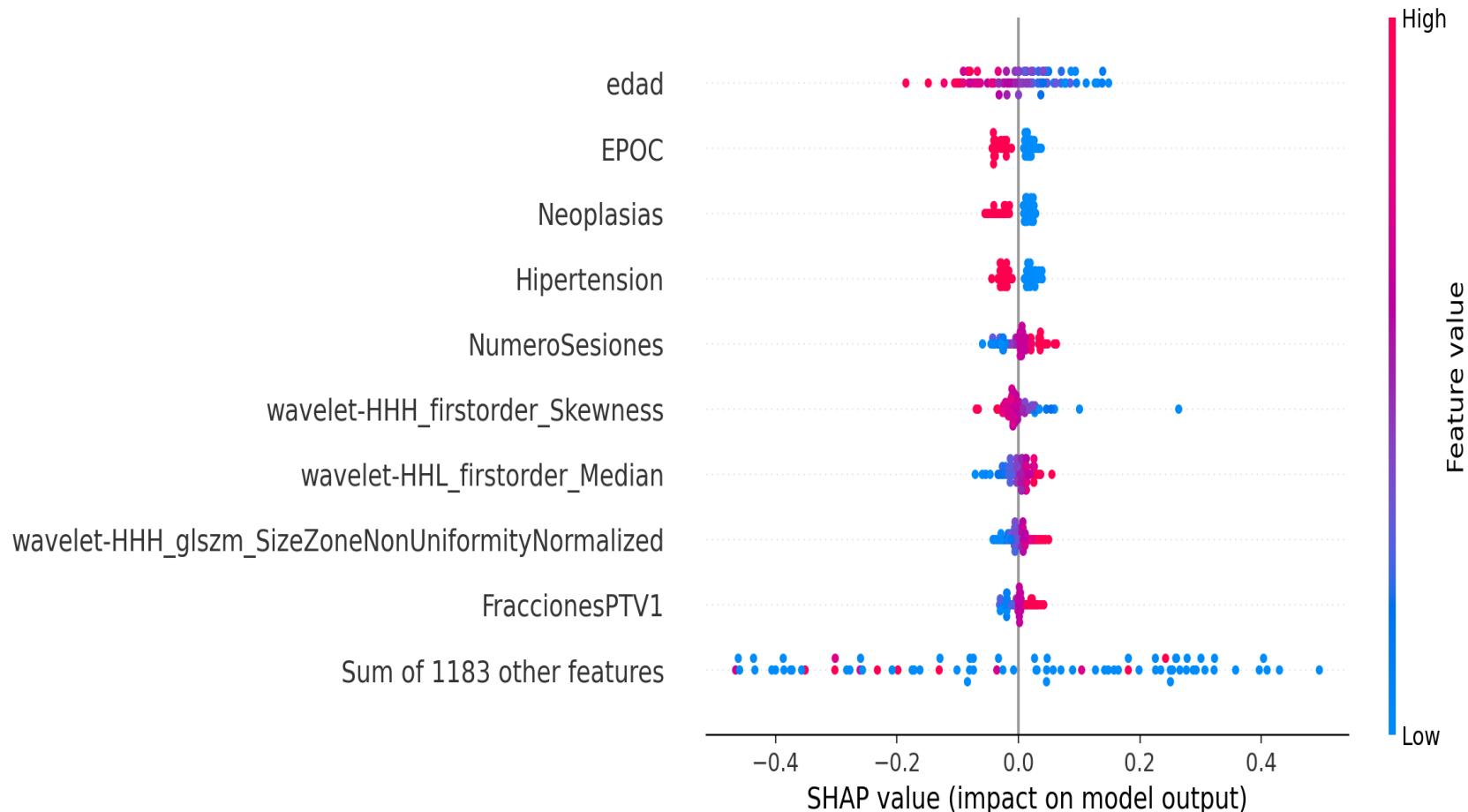


Resultados Preliminares: Sequoia vs RSF

Poblacional Survival



Importancia de Variables



Variables identificadas

- Edad
- Epoc
- Neoplasias
- Hipertensión
- Heterogenidad del tumor
- Ramificaciones del tumor

Conclusiones

- SEQUOIA-RT proporciona supervivencia individualizada y riesgo acumulado con uso de datos clínicos y radiómicos
- Convergencia a población promedio (Kaplan-Meier)
- Predicción de factores pronóstico ya conocidos y descubrimiento de nuevos

Hoja de ruta

- Profundizar en el desarrollo del algoritmo y del marco teórico
- Intercomparación con otros modelos de predicción de supervivencia
- Ampliación de la población de estudio



¡Gracias!

JORNADAS SARH

Inteligencia Artificial Aplicada
a la Física Médica



Parador de Antequera
14 de noviembre 2025