

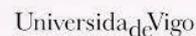
MEMORIA CIENTÍFICA

PROYECTOS COLABORATIVOS GALICIA



MINISTERIO DE CIENCIA E INNOVACIÓN

Plan de Recuperación, Transformación y Resiliencia



Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	1/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Título y acrónimo

Modelos de curación tipo Single-Index: aplicación a pacientes con cardiotoxicidad (MCSICAR)

Resumen *(máx. 300 palabras)*

La investigación en cáncer de mama ha logrado grandes avances en las últimas décadas, lo cual ha derivado en un aumento de la tasa de supervivencia. Sin embargo, algunas de las terapias oncológicas que cuentan con una buena eficacia antitumoral pueden provocar efectos secundarios, como es el caso de la cardiotoxicidad. Poder anticiparse a la aparición de este daño cardiaco resulta de fundamental importancia. A pesar de que se conocen algunos factores de riesgo para la cardiotoxicidad, estos no son suficientes para predecir qué pacientes desarrollarán dicha complicación durante su tratamiento. Por este motivo, se realizan periódicamente ecocardiogramas para chequear el estado del corazón. Con la finalidad de mejorar la calidad de vida de las pacientes, sería interesante conocer qué factores influyen en la probabilidad de padecer este efecto secundario y en el tiempo hasta su aparición. Los factores objeto de estudio pueden ser unidimensionales, vectoriales, funcionales o incluso de tipo imagen. Este es el caso de la TDI (Tissue Doppler Imaging), que muestra la velocidad a la que se relaja y contrae el músculo cardiaco. La utilización de imágenes clínicas está abriendo importantes posibilidades de asistencia al diagnóstico y al tratamiento. Puesto que, en muchas ocasiones, dichas imágenes presentan información no deseada y mucho ruido, requieren un preprocesado previo. En este contexto, desde un punto de vista estadístico, es necesario contar con modelos adecuados para trabajar con tiempos de vida en una situación de alta dimensión.

Líneas de actuación

Línea de Actuación 2. Implementación y análisis de bases de datos en medicina de precisión.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora	
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14	
Observacións		Páxina	2/18	
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==			
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).			

Contexto científico y resultados previos (máx. 1000 palabras)

El análisis de supervivencia es una rama de la estadística que estudia el tiempo que transcurre hasta la ocurrencia de un evento de interés. Si consideramos como evento la cardiotoxicidad, esta disciplina nos podría permitir conocer cómo se distribuye el tiempo hasta la aparición de dicho efecto adverso y qué factores pueden hacer que se adelante o se retrase. Los modelos de supervivencia permiten trabajar con datos censurados, que surgen de manera natural en muchos estudios médicos. El término censura hace referencia a situaciones en las que hay una pérdida de información en la variable de interés, de modo que algunas veces no se observa esta variable sino otra (conocida como variable de censura) que impide observar el tiempo hasta el evento de interés. Por ejemplo, si una paciente llega a la fecha de fin de estudio sin haber padecido cardiotoxicidad, únicamente sabremos que el tiempo hasta el desarrollo de dicho efecto adverso (en caso de ocurrir) es mayor o igual que el tiempo durante el cual la hemos seguido. Lo mismo ocurre si perdemos el seguimiento de las mujeres por otros motivos, como la muerte por una causa ajena al estudio o el cambio de hospital. Los modelos clásicos de supervivencia suponen que eventualmente todos los pacientes sufrirían el evento si pudiesen ser seguidos durante un tiempo suficientemente largo. Esta hipótesis no es realista en el caso de la cardiotoxicidad, pues no todas las pacientes son susceptibles a sufrirla. En situaciones de este tipo, lo apropiado es utilizar modelos de curación, que consideran la posibilidad de que la paciente se *cure*, entendiendo como curación el hecho de que nunca llegue a sufrir el evento de interés (cardiotoxicidad, en este caso).

En los últimos años, diversos autores, entre los que se encuentran tres de los investigadores implicados en este proyecto (Ricardo Cao Abad, M. Amalia Jácome Pumar y Ana López Cheda), han propuesto modelos de curación flexibles (no paramétricos). Estos modelos permiten estudiar la probabilidad de cura (ver López-Cheda, Cao, Jácome y Van Keilegom, 2017) y la distribución del tiempo hasta que se desarrolla la cardiotoxicidad en los individuos susceptibles, también llamada latencia (ver López-Cheda, Jácome y Cao, 2017). Asimismo, en López-Cheda, Jácome, Van Keilegom y Cao (2020) se han propuesto métodos para contrastar si la probabilidad de cura depende de cada una de las posibles variables explicativas. Los métodos propuestos están implementados y disponibles para toda la comunidad científica en el paquete `npcure` de R (ver López-de-Ullibarri, López-Cheda y Jácome (2020), y López-Cheda, Jácome y López-de-Ullibarri, 2021). Recientemente, se han aplicado estas técnicas para estimar tiempos de estancia en hospital para pacientes de COVID-19 (ver Pedrosa-Laza, López-Cheda y Cao (2021), y López-Cheda, Jácome, Cao, y De Salazar (2021), entre otros), o en contextos médicos más generales (ver García-Vicuña, López-Cheda, Jácome y Mallor, 2023). Un aspecto a destacar es que estos trabajos se restringen a la presencia de variables explicativas unidimensionales. Sin embargo, en los estudios biomédicos es frecuente encontrar, de manera natural, variables explicativas vectoriales (de dimensión mayor que uno), funcionales o de tipo imagen. Para cubrir ese hueco en la literatura, López-Cheda, Peng y Jácome (2023) han propuesto un método para las situaciones en las que existen diferentes variables que influyen en la probabilidad de cura y en la latencia. Actualmente, Ricardo Cao Abad y Ana López Cheda están dirigiendo la tesis doctoral de Beatriz Piñeiro Lamas, cuyo objetivo es la extensión de las propuestas existentes en la literatura a situaciones en las que existen múltiples variables explicativas (factores de riesgo). A tal efecto, se han propuesto los modelos *single-index* de curación de tipo *mixtura*, tanto vectoriales como funcionales. Además, se ha implementado un preprocesado de las imágenes TDI, que permite extraer una función a partir de cada una de ellas. Las funciones resultantes presentan un comportamiento periódico, pues muestran la velocidad a la que se relaja y contrae el corazón en cada latido. En cada uno de los ciclos se pueden observar tres ondas predominantes: la S (sístole) y la E' y la A' (diástole). A efectos prácticos, se ha seleccionado un único ciclo para cada paciente. Esto ha permitido detectar diferencias relevantes en los alrededores de la onda E' entre las pacientes que han sufrido cardiotoxicidad y las que no. Dicho hallazgo pone de manifiesto la posible capacidad predictiva de la TDI en este contexto.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	3/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Referencias:

- Garcia-Vicuña D, López-Cheda A, Jácome MA, Mallor F (2023). Estimation of patient flow in hospitals using up-to-date data. Application to bed demand prediction during pandemic waves. PLoS ONE 18(2): e0282331.
- López-Cheda A, Cao R, Jácome MA, Van Keilegom I (2017). Nonparametric incidence estimation and bootstrap bandwidth selection in mixture cure models. Computational Statistics & Data Analysis, 105, 144-165.
- López-Cheda A, Jácome MA, Cao R (2017). Nonparametric latency estimation for mixture cure models. TEST, 26(2), 353-376.
- López-Cheda A, Jácome MA, Cao R, De Salazar PM (2021). Estimating lengths-of-stay of hospitalised COVID-19 patients using a non-parametric model: a case study in Galicia (Spain). Epidemiology and Infection, 149, e102, 1-8.
- López-Cheda A, Jácome MA, López-de-Ullibarri I (2021). The R Journal, 13 (1), 21-41.
- López-Cheda A, Jácome A, Van Keilegom I, Cao R (2020). Nonparametric covariate hypothesis tests for the cure rate in mixture cure models. Statistics in Medicine, 39, 2291-2307.
- López-Cheda A, Peng Y, Jácome MA (2023). Nonparametric estimation in mixture cure models with covariates. TEST. Invited paper. <https://doi.org/10.1007/s11749-022-00840-z>
- López-de-Ullibarri I, López-Cheda A, Jácome MA (2020). npcure: Nonparametric Estimation in Mixture Cure Models. R package version 0.1-5. <https://CRAN.R-project.org/package=np cure>
- Pedrosa-Laza M, López-Cheda A, Cao R (2022). Cure models to estimate time until hospitalization due to COVID-19. Applied Intelligence, 52, 794-807.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	4/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Necesidad clínica e impacto social (máx. 1000 palabras)

El cáncer de mama es el más frecuente en mujeres. De acuerdo con la Organización Mundial de la Salud (2021), en 2020 se diagnosticó cáncer de mama a 2.3 millones de mujeres en todo el mundo, y 685000 fallecieron por esa enfermedad (ver World Health Organization, 2021). Aproximadamente un 20% de los cánceres de mama tienen altos niveles de una proteína llamada HER2, cuya función es transmitir señales a las células para que crezcan y se dividan. Este subtipo de cáncer se conoce como HER2 positivo y su tratamiento por excelencia es el trastuzumab. Estudios previos han confirmado que este anticuerpo monoclonal mejora la supervivencia, así como la progresión libre de enfermedad (ver, entre otros, Mazzotta et al, 2019). Sin embargo, la cardiotoxicidad es un efecto adverso conocido de esta terapia oncológica (Wadhwa et al, 2009 y Dempsey et al, 2021), que sufren aproximadamente el 10% de las pacientes limitando, en muchas ocasiones, una apropiada administración del tratamiento y haciendo disminuir su eficacia. Por este motivo, la detección de los factores de riesgo que hacen que una paciente pueda sufrir cardiotoxicidad es de capital importancia. Entre las pruebas que se llevan a cabo para evaluar la función cardíaca en pacientes sometidas a trastuzumab está el ecocardiograma doppler tisular. En este contexto, y con el objetivo de mejorar la calidad de vida de las pacientes, surgen las siguientes cuestiones:

- ¿Cuál es la probabilidad de sufrir cardiotoxicidad en pacientes que reciben trastuzumab y otros tratamientos potencialmente cardiotoxicos? ¿Qué factores pueden hacer que dicha probabilidad aumente?
- Para las pacientes que sufren dicho efecto adverso, ¿cuál es la distribución del tiempo hasta su ocurrencia? ¿Qué factores podrían retrasar la aparición de la cardiotoxicidad?

Dar respuesta a estas preguntas podría encaminarnos hacia una medicina más personalizada, adaptada a cada paciente. Con este objetivo nació la cardio-oncología. La Sociedad Europea de Cardiología (ESC) elaboró las primeras guías de cardio-oncología, que contienen pautas para la toma de decisiones en pacientes oncológicos con comorbilidades cardiovasculares (ver Lyon et al, 2022).

Por otra parte, la predicción de los periodos de mayor riesgo de cardiotoxicidad podría ser útil para realizar una óptima distribución de las citas médicas y pruebas de control que reciben las pacientes.

En este contexto, sería ideal poder aprovechar al máximo toda la información disponible para cada paciente. Por ejemplo, cada vez que se realiza un ecocardiograma doppler se almacenan automáticamente, además de la imagen TDI, una serie de parámetros relacionados con el corazón. Sacar provecho de toda esta información obtenida mediante pruebas no invasivas podría conducir al descubrimiento de nuevos factores de riesgo.

Referencias:

- Dempsey N et al (2021). Trastuzumab-induced cardiotoxicity: a review of clinical risk factors, pharmacologic prevention, and cardiotoxicity of other HER2-directed therapies. *Breast Cancer Res Treat*, 188, 21-36.
- Lyon AR et al (2022). ESC Guidelines on cardio-oncology developed in collaboration with the European Hematology Association (EHA), the European Society for Therapeutic Radiology and Oncology (ESTRO) and the International Cardio-Oncology Society (IC-OS): Developed by the task force on cardio-oncology of the European Society of Cardiology (ESC), *European Heart Journal*, Volume 43, Issue 41, 1 November 2022, Pages 4229-4361.
- Mazzotta M et al (2019). Long-term safety and real-world effectiveness of trastuzumab in breast cancer. *J. Clin. Medicine* 8(2), 254.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	5/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



- Wadhwa D et al (2009). Trastuzumab mediated cardiotoxicity in the setting of adjuvant chemotherapy for breast cancer: a retrospective study. Breast Cancer Res. Treat, 117, 357-364.
- World Health Organization. Breast cancer (2021). <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	6/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Objetivos *(máx. 700 palabras)*

El objetivo principal de este proyecto es

Mejorar la esperanza de vida y la calidad de vida de las pacientes de cáncer de mama mediante la propuesta de metodologías que den lugar a sistemas más precisos de prevención y diagnóstico no invasivo frente al desarrollo de cardiotoxicidad, dando lugar a una medicina personalizada.

Se pretende alcanzar este objetivo principal a partir de los siguientes objetivos secundarios:

01. Obtener y depurar las bases de datos clínicos y las imágenes.
02. Proponer métodos para estimar la probabilidad de cura y la latencia, considerando covariables multidimensionales de tipo vectorial, a partir de modelos de curación de tipo single-index. Estudiar las propiedades teóricas de los métodos propuestos.
03. Extender los métodos propuestos:
 - a. para modelos en los que influyen covariables funcionales.
 - b. para permitir que la probabilidad de cura y la latencia dependan de diferentes covariables.
04. Proponer contrastes de significación de covariables para la función de la latencia, basándose en los contrastes de significación de covariables para la probabilidad de cura existentes en la literatura.
05. Proponer un método de clasificación que permita determinar si una mujer es susceptible a sufrir cardiotoxicidad.
06. Implementar los algoritmos propuestos mediante software de libre disposición para toda la comunidad investigadora con el fin de que investigadores de otros ámbitos o tecnólogos puedan aplicar los procedimientos desarrollados en la resolución de problemas reales.
07. Aplicar la metodología propuesta al conjunto de datos de pacientes de cardiotoxicidad:
 - a. Identificar los factores de riesgo que aumentan la probabilidad de que una paciente sufra cardiotoxicidad y/o se acorte el tiempo libre de este efecto adverso.
 - b. Estimar la probabilidad de padecer cardiotoxicidad durante el tratamiento con trastuzumab de forma personalizada, en función de las características clínicas de las pacientes.
 - c. Detectar el periodo de mayor riesgo de sufrir el efecto adverso.
 - d. Estudiar las diferencias en las curvas relativas a la velocidad de contracción y relajación del músculo cardiaco correspondientes a pacientes curadas y no curadas.
08. Difundir los resultados

Este proyecto será clave para conocer el papel de determinados factores en la aparición de cardiotoxicidad, lo que permitirá a los profesionales de la biomedicina decidir si un factor tiene efecto significativo tanto en la probabilidad de sufrir cardiotoxicidad como, para aquellas pacientes que sufrirán cardiotoxicidad, en el tiempo libre de este efecto adverso.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	7/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Plan de trabajo y metodología empleada *(máx. 1300 palabras)*

El proyecto se llevará a cabo con datos relativos a pacientes de la Unidad de Imagen y Función Cardiacas del Complejo Hospitalario Universitario de A Coruña (CHUAC). Específicamente, se trata de un diseño retrospectivo de cohorte restringido a las pacientes con cáncer de mama de tipo HER2 positivo que han sido tratadas con trastuzumab y que acudieron entre el 2007 y 2022 al CHUAC para realizar ecocardiogramas doppler, con la finalidad de chequear el estado de su corazón y detectar la posible aparición de cardiotoxicidad. Para cada paciente, se dispondrá de diferente información relativa a variables clínicas, tomadas al inicio del tratamiento: máximo gradiente de presión en la aorta; ratio E/A MV (marcador de la función del ventrículo izquierdo en las ondas E y A); latidos por minuto; dimensión de la aurícula izquierda; diámetro telesistólico del ventrículo izquierdo; LVEF (fracción de eyección del ventrículo izquierdo); masa del ventrículo izquierdo; PPVItD (pared posterior VI en telediástole); DVIItD (diámetro telediastólico del ventrículo izquierdo); edad; peso; estatura y tiempo hasta desarrollar cardiotoxicidad (en caso de haberla sufrido). Una característica importante de este conjunto de datos es la presencia de sujetos curados, ya que existen pacientes tratadas con trastuzumab y que nunca van a sufrir cardiotoxicidad.

Para alcanzar los objetivos presentados en el apartado anterior, se plantean las siguientes **tareas**:

- O1.T1: Recogida y depuración de la base de datos
- O2.T1: Propuesta de un modelo de curación tipo single-index para estimar la probabilidad de cardiotoxicidad
- O2.T2: Propuesta de un modelo de curación tipo single-index para estimar la latencia
- O2.T3: Estudio de simulación para comprobar el buen comportamiento de los métodos propuestos
- O2.T4: Estudio de las propiedades teóricas de los métodos propuestos.
- O3.T1: Adaptación del método propuesto en O2.T1 y O2.T2 a covariables funcionales
- O3.T2: Adaptación del método propuesto en O2.T1 y O2.T2 para permitir que diferentes variables influyan en la probabilidad de cardiotoxicidad y en la latencia
- O4.T1: Propuesta de contrastes de significación de covariables para la latencia
- O4.T2: Estudio de simulación para comprobar el buen comportamiento del método
- O5.T1: Propuesta de un método que permita clasificar a una mujer como susceptible o no a sufrir cardiotoxicidad
- O6.T1: Diseño e implementación de los métodos propuestos, adaptados para que se puedan aplicar a cualquier conjunto de datos
- O6.T2: Preparación de ficheros de ayuda, para que cualquier investigador pueda entender y utilizar el código fácilmente
- O6.T3: Publicación de software en el CRAN de R
- O7.T1: Aplicación de la metodología propuesta al conjunto de datos de cardiotoxicidad
- O7.T2: Recogida e interpretación de resultados
- O8.T1: Difusión de resultados con el envío a revistas de alto impacto y con la asistencia a congresos nacionales y/o internacionales

El proyecto aplicará todas las políticas implementadas por la UDC y el CITIC en relación con la gestión de datos. Todos los datos que puedan hacerse públicos estarán disponibles a través de un repositorio de acceso abierto como Zenodo, e indexados en el portal opendata del CITIC. Los datos y metadatos generados en el proyecto cumplirán los principios FAIR (Findable, Accessible, Interoperable, and Reusable). Las publicaciones y los programas informáticos también se compartirán mediante políticas de acceso libre y abierto.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	8/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Dado que se analizarán datos clínicos, el tratamiento, comunicación y cesión de los datos se hará conforme a lo dispuesto por el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016 y la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. El estudio se llevará a cabo siguiendo la normativa vigente en la Ley 14/2007, de 3 de julio, de investigación biomédica, y de acuerdo con los principios derivados de la última versión de la Declaración de Helsinki y del Convenio de Derechos Humanos y Biomedicina (Convenio de Oviedo). El acceso a la historia clínica y al consentimiento informado se realizará de acuerdo con la normativa vigente. El estudio cuenta con la aprobación del Comité de ética de la investigación con medicamentos de Galicia (CEIm-G).

La **metodología** a usar será la siguiente. Una vez recogida y depurada la base de datos, se realizará el preprocesado de las variables de tipo imagen con la finalidad de seleccionar las regiones de interés. Específicamente, la curva TDI está representada con color amarillo en la imagen médica, por lo que interesa almacenar únicamente los píxeles con diferentes tonalidades amarillas. Para eso, se tendrá en cuenta el color en formato HSV (de inglés, Hue-Saturation-Value) de cada pixel. El motivo es que, en HSV, los píxeles amarillos pueden ser identificados fácilmente al controlar el rango del primer parámetro (Hue), mientras que en formato RGB (Red-Green-Blue), diferentes combinaciones de los tres parámetros corresponden a colores amarillos similares, lo que dificultaría considerablemente recoger por completo la curva de interés. Una vez que los píxeles amarillos han sido detectados, se propondrá un algoritmo para convertir esa imagen en función. A continuación, considerando que, para cada paciente, se habrá almacenado un número diferente de ciclos, se seleccionará un único ciclo por individuo y se estandarizará en un intervalo, para poder comparar a todos los sujetos.

Una vez preparada la información clínica y personal de las pacientes, el efecto de esta en la probabilidad de sufrir cardiotoxicidad se estudiará mediante los modelos single-index de curación de tipo mixtura. Estos modelos generalizarán dos importantes líneas de investigación. Por una parte, los modelos introducidos en López-Cheda, Cao, Jácome y Van Keilegom (2017) y López-Cheda, Jácome y Cao (2017) al caso multidimensional. Por otra parte, la propuesta generaliza el trabajo de Strzalkowska-Kominiak y Cao (2013) a un contexto en el cual, además de individuos censurados, hay individuos curados. Es relevante destacar que, mediante los modelos single-index, tanto en el caso vectorial como funcional, se evita la conocida maldición de la dimensión, presente en los modelos no paramétricos. Sin embargo, el precio que hay que pagar es la incorporación al modelo de un vector (o función) de parámetros a estimar. Dicha estimación se hará por máxima verosimilitud, y en la práctica será necesario recurrir a métodos de optimización numéricos. En el caso funcional, se estudiará reducir previamente la dimensión mediante una representación con bases de Fourier o con componentes principales funcionales.

Asimismo, partiendo del método propuesto en López-Cheda, Jácome, Van Keilegom y Cao (2020), se propondrá uno nuevo para estudiar si la latencia depende de cada variable explicativa.

Para determinar si una paciente es susceptible o no a la cardiotoxicidad, se propondrá una regla de clasificación que tenga en cuenta la estimación de la probabilidad de cura.

Se prevén tres grandes aspectos en el plan de difusión de los resultados de este proyecto:

1. Publicación de artículos de investigación en revistas de ámbito estadístico y médico, reconocidas internacionalmente.
2. Presentación de ponencias en congresos de ámbito nacional e internacional. Esto permite avanzar los resultados obtenidos a la comunidad estadística y médica internacional cuando aún están en una etapa de consolidación, con la consecuente retroalimentación obtenida en los mismos.
3. Puesta a disposición de documentos de investigación y software con libre acceso. Los documentos de investigación estarán a disposición en la web

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	9/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



del grupo MODES (ver <http://dm.udc.es/modes/>). El software estará accesible en el CRAN de R (ver <http://cran.r-project.org/>).

Referencias:

- López-Cheda A, Cao R, Jácome MA, Van Keilegom I (2017). Nonparametric incidence estimation and bootstrap bandwidth selection in mixture cure models. *Computational Statistics & Data Analysis*, 105, 144-165.
- López-Cheda A, Jácome MA, Cao R (2017). Nonparametric latency estimation for mixture cure models. *TEST*, 26(2), 353-376.
- López-Cheda A, Jácome MA, Van Keilegom I, Cao R (2020). Nonparametric covariate hypothesis tests for the cure rate in mixture cure models. *Statistics in Medicine*, 39, 2291-2307.
- Strzalkowska-Kominiak E, Cao R (2013). Maximum likelihood estimation for conditional distribution single-index models under censoring. *Journal of Multivariate Analysis*, 114, 74-98.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	10/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Consortio (entidades gallegas participantes y de otras CCAA, describir la entidad y qué va a aportar al proyecto cada una de ellas) *(máx. 1000 palabras)*

Las entidades participantes en este proyecto son la Universidade da Coruña (UDC) y más concretamente el Centro de Investigación TIC (CITIC), y el Instituto de Investigación Biomédica A Coruña (INIBIC).

- **Centro de Investigación en Tecnologías de la Información y las Comunicaciones (CITIC)**

El **CITIC** es un Centro de Investigación creado en 2008 que potencia el avance y la excelencia en I+D+i aplicada a las TIC. Impulsado por la UDC, el CITIC es un punto de encuentro entre la Universidad y la empresa en el que confluyen departamentos de I+D de empresas del sector TIC con investigadores procedentes del ámbito universitario. En el año 2016, el CITIC recibe el distintivo de Centro Singular de Investigación de Galicia 2016-2019. Es el único centro TIC de Galicia reconocido con este distintivo, cofinanciado por la Xunta de Galicia y por la Unión Europea a través del Fondo Europeo de Desarrollo Regional. En 2019 renueva esta acreditación, pasando a denominarse Centro de Investigación del Sistema Universitario de Galicia 2019-2022. La actividad científica del CITIC se estructura en cuatro áreas de investigación principales, Inteligencia Artificial, Ciencia e Ingeniería de Datos, Computación de Altas Prestaciones y Servicios y Redes Inteligentes; y un área de investigación transversal a todas ellas, la de Ciberseguridad. En este proyecto participarán investigadores del área de Ciencia e Ingeniería de Datos y del área de Inteligencia Artificial.

Dentro del CITIC, el proyecto lo liderará el **Grupo de Modelización, Optimización e Inferencia Estadística (MODES)**, siendo el investigador principal Ricardo Cao Abad, quien cuenta con una sólida experiencia en liderazgo de proyectos científicos. Participarán además los siguientes miembros de este grupo: Ana López Cheda y María Amalia Jácome Pumar. El grupo MODES está formado por 41 integrantes, siendo la mayoría de ellos investigadores senior con una larga trayectoria. Su investigación se centra en dos grandes áreas: Modelización e Inferencia Estadística, e Investigación Operativa. En el área de Modelización e Inferencia Estadística, donde se enmarca este proyecto, la investigación se centra en la propuesta y desarrollo de nuevos métodos estadísticos en diferentes contextos, siendo de interés: (i) el estudio de sus propiedades teóricas, (ii) la evaluación de su correcto comportamiento mediante simulación, y (iii) su aplicación a un gran número de problemas relevantes en otras áreas de investigación científica, destacando la Medicina, la Biología y la Ingeniería. El grupo de investigación ha sido financiado parcialmente en los últimos 25 años por los programas de investigación MINECO-MICINN-MEC y por las Axudas do programa de consolidación e estructuración de unidades de investigación competitivas (modalidad de grupos de referencia competitiva, GRC). El grupo MODES contribuye a la comunidad científica con la publicación de una media de 30-40 artículos al año, la gran mayoría en revistas incluidas en el Journal of Citation Reports de la Web of Science. Los avances en las líneas de investigación del grupo frecuentemente conducen a colaboraciones con empresas e instituciones interesadas en la resolución de problemas relacionados con la Ciencia de Datos, susceptibles de ser abordados mediante algoritmos de análisis de datos o con modelos de investigación operativa. Dichas colaboraciones se materializan mediante proyectos de investigación conjuntos con empresas, así como con contratos de transferencia de tecnología. Entre estos, cabe destacar la colaboración con ABANCA, ECOMT, EDAR Bens, EMALCSA, Estrella Galicia, EXTRACO, FINSA, GABADI, INDITEX, INVERDAMA, La Voz de Galicia, MISTURAS, NAVANTIA, NEODYN, PROMECAN, REGANOSA, TECMAN y WESTAT. Asimismo, los miembros de MODES participan activamente en el desarrollo y publicación en abierto de un gran número de paquetes

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora	
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14	
Observacións		Páxina	11/18	
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==			
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).			

de software, como son las librerías R actualmente disponibles en el CRAN (binnednp, baggedcv, ILS, kerdienst, MPBoost, npcure, npsp, survPresmooth, TSclust, entre otras).

Además del grupo MODES, colaborarán otros miembros del CITIC como investigadores del grupo **Laboratorio de Bases de Datos (LBD)**, liderados por la investigadora Susana Ladra, junto con José R. Paramá, Fernando Silva-Coira, Guillermo de Bernardo, Diego Seco y Elsa Martín-de Arribas, quienes aportarán su experiencia en el desarrollo de algoritmos para el análisis de datos genómicos, diseño y desarrollo de bases de datos y explotación de datos; del grupo de investigación **Redes de Neuronas Artificiais e Sistemas Adaptativos -Informática Médica e Diagnóstico Radiolóxico (RNAS-IMEDIR)**, liderados por Carlos Fernández-Lozano, quien aportará al proyecto su experiencia en técnicas de inteligencia artificial y *machine learning*, así como sus conocimientos previos de biología computacional aplicados al dominio de CCR; y el investigador del programa Ramón y Cajal **José A. Iglesias Guitián**, experto en visualización científica, que podrá aportar sus conocimientos para la gestión de las imágenes generadas por los ecocardiogramas doppler. Todos los grupos de investigación participantes tienen experiencia previa en proyectos del ámbito biotecnológico y en colaboración con entidades biosanitarias.

- **Instituto de Investigación Biomédica de A Coruña (INIBIC, Fundación Profesor Novoa Santos)**

El INIBIC es un Instituto de Investigación Sanitaria (IIS), constituido por la Consellería de Sanidade, el Servicio Galego de Saúde, la Universidade da Coruña y la Fundación Profesor Novoa Santos. Fue creado en enero de 2008 mediante la firma de un convenio de colaboración entre las entidades firmantes. El INIBIC surge con la necesidad de promover, desarrollar, gestionar y difundir la investigación en el ámbito de las ciencias de la salud. Con esta finalidad, el Instituto promueve y estrecha las relaciones de intercambio de conocimiento entre todos los investigadores y grupos de investigación que pertenecen a los distintos centros (Gerencia de Gestión Integrada de A Coruña y UDC) y promueve la colaboración con otras instituciones y entidades públicas y privadas, priorizando la realización de actividades para potenciar las alianzas para llevar a cabo una investigación traslacional con repercusión en la actividad clínica.

Alberto Bouzas Mosquera y Cayetana Barbeito Caamaño son cardiólogos de la Unidad de Imagen y Función Cardiacas del Servicio de Cardiología del CHUAC y miembros del INIBIC. En este proyecto, Alberto Bouzas será el responsable de suministrar los datos totalmente anonimizados y Cayetana Barbeito formará parte del equipo investigador, aportando su conocimiento experto en cardiología y, específicamente, en tratamiento de pacientes con cardiotoxicidad.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora	
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14	
Observacións		Páxina	12/18	
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==			
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).			

Presupuesto y cronograma *(máx. 500 palabras)*

Con respecto al importe solicitado para este proyecto, se consideran los siguientes gastos:

- Gastos relativos a cursos de formación: 2000 €
- Gastos para la realización de estancias de formación: 4000 €
- Gastos de asistencia a congresos: 7000 €
- Gastos reservados para cubrir el coste de publicaciones (aceptación de artículo en revistas de alto impacto, etc.): 3000 €

Por lo tanto, se solicita un total de 16000€ (del máximo de 75000€ establecido para cada proyecto).

Por otra parte, se prevé un gasto de aproximadamente 60000€ destinado a contratación de personal:

Año	Salario	Pagas extras	Indem. Fin contrato	Total retribuciones	Seguridad Social	Coste del contrato
2024	25.140,00€	4.190,04€	0,00 €	29.330,04€	9.561,60€	38.891,64€
2025	12.570,00€	2.095,02€	1.466,22€	16.131,24€	4.780,80€	20.912,04 €
			TOTAL	45.461,28€	14.342,40€	59.983,68 €

La siguiente tabla muestra el desglose de todos los gastos considerados asociados a las tareas:

Tarea	Personal contratado	Formación	Participación en congresos
01.T1	X		
02.T1,02.T2	X	X	
02.T3, 02.T4	X	X	X
03.T1	X	X	X
03.T2	X		X
04.T1	X		
04.T2	X	X	X
05.T1	X		X
06.T1	X		
06.T2	X		
06.T3	X	X	
07.T1	X		X
07.T2	X		
08.T1	X		X

La planificación temporal para el desarrollo de las tareas descritas anteriormente se detalla en la siguiente tabla:

Tarea	1-3 meses	4-6 meses	7-9 meses	10-12 meses	12-15 meses	16-18 meses
01.T1	X					
02.T1, 02.T2	X	X				
02.T3		X	X			
02.T4			X	X		
03.T1	X	X	X			
03.T2			X	X		
04.T1		X	X			
04.T2			X	X		
05.T1				X	X	X
06.T1		X	X	X	X	X
06.T2					X	X
06.T3						X
07.T1			X	X	X	
07.T2				X	X	
08.T1			X	X	X	X

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Páxina	13/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Tabla ética

Embriones/fetos humanos

- ¿Su investigación involucra células madre embrionarias humanas (hESC)? En caso afirmativo:
- ¿Se derivarán directamente de embriones dentro de este proyecto?
 - ¿Son líneas celulares previamente establecidas?
 - ¿Están registradas las líneas celulares en el registro de la UE para líneas de células madre embrionarias humanas?
- ¿Su investigación implica el uso de embriones humanos? En caso afirmativo:
- ¿Están registradas las líneas celulares en el registro de la UE para líneas de células madre embrionarias humanas?
- ¿Su investigación implica el uso de células/tejidos fetales humanos?

Humanos

- ¿Su investigación involucra a participantes humanos? En caso afirmativo:
- ¿Son voluntarios para investigaciones en ciencias sociales o humanas?
 - ¿Son personas incapaces de dar su consentimiento informado?
 - ¿Son individuos o grupos potencialmente vulnerables?
 - ¿Son niños/menores?
 - ¿Son pacientes para estudios médicos?
 - ¿Son voluntarios sanos para estudios médicos?
- ¿Su actividad implica intervenciones (físicas que también incluyen tecnología de imágenes, tratamientos conductuales, seguimiento y localización, etc.) en los participantes del estudio? En caso afirmativo:
- ¿Implica técnicas invasivas? (por ejemplo, recolección de células o tejidos humanos, intervenciones quirúrgicas o médicas, estudios invasivos en el cerebro, TMS, etc.)
 - ¿Implica la recogida de muestras biológicas?
- ¿Su actividad implica la realización de un estudio clínico tal como lo define el Reglamento de Ensayos Clínicos 536/2014? (uso de productos farmacéuticos, biológicos, radiofármacos o medicamentos de terapia avanzada). En caso afirmativo:
- ¿Es un ensayo clínico?

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora	
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14	
Observacións		Páxina	14/18	
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==			
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).			

¿Es un ensayo clínico de baja intervención?

Células/Tejidos humanos

¿Su investigación involucra células o tejidos humanos (que no sean embriones/fetos humanos, es decir, la sección 1)? En caso afirmativo:

- ¿Están disponibles comercialmente?
- ¿Se obtienen dentro de este proyecto?
- ¿Se obtienen dentro de otro proyecto, laboratorio o institución?
- ¿Se obtienen de un biobanco?

Datos personales

¿Su investigación implica la recopilación y/o el procesamiento de datos personales? En caso afirmativo:

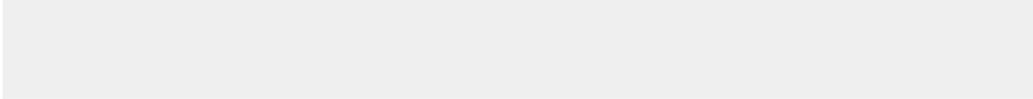
¿Implica la recolección y/o procesamiento de datos personales sensibles? En caso afirmativo, (p. ej., salud, estilo de vida sexual, origen étnico, opinión política, convicciones religiosas o filosóficas)

¿Se trata de un tratamiento de datos genéticos, biométricos o de salud?

¿Implica la elaboración de perfiles, el seguimiento sistemático de personas o el procesamiento a gran escala de categorías especiales de datos o métodos intrusivos de procesamiento de datos?

¿Su investigación implica un mayor procesamiento de datos personales recopilados previamente? (incluido el uso de conjuntos de datos o fuentes preexistentes, la fusión de conjuntos de datos existentes)?

¿Está previsto exportar datos personales (transferencia de datos) desde la UE a países no pertenecientes a la UE? (Especifique el tipo de datos personales y países involucrados)



¿Está previsto importar datos personales (transferencia de datos) de países no pertenecientes a la UE a la UE o de un país no perteneciente a la UE a otro país no perteneciente a la UE? (Especifique el tipo de datos personales y países involucrados)

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Página	15/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



¿Su actividad implica el tratamiento de datos personales relacionados con condenas o delitos penales?

Animales

¿Su investigación involucra animales?
 En caso afirmativo, indique las especies involucradas:

- ¿Son vertebrados?
- ¿Son primates no humanos?
- ¿Son modificados genéticamente?
- ¿Son animales de granja clonados?
- ¿Son especies en peligro de extinción?

Terceros países

¿Su investigación involucra a países no pertenecientes a la UE? En caso afirmativo, especifique los países involucrados:

En caso de que participen países no pertenecientes a la UE, ¿las actividades realizadas en estos países plantean posibles problemas éticos? Especificar los países:

¿Planea utilizar recursos locales? por ejemplo, animal; y/o muestras de tejidos humanos, material genético, animales vivos, restos humanos, materiales de valor histórico, muestras de fauna o flora amenazadas, etc.

¿Está previsto importar algún material (que no sean datos) de países no pertenecientes a la UE a la UE o de un país no perteneciente a la UE a otro país no perteneciente a la UE?

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Página	16/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Para la importación de datos, complete también la sección 4.

Para las importaciones relativas a células o tejidos humanos, rellene también la sección 3.

En caso de que participen países no pertenecientes a la UE, ¿las actividades realizadas en estos países plantean posibles problemas éticos? Especificar los países:

¿Está previsto exportar algún material (que no sean datos) de la UE a países extracomunitarios? En caso afirmativo, especifique el material y los países involucrados:

Para la exportación de datos, complete también la sección 4.

Si su investigación involucra países de ingresos bajos y/o medianos bajos, ¿se prevén medidas de distribución de beneficios?

¿La situación del país podría poner en riesgo a la persona que participa en la investigación?

Medio ambiente, salud y seguridad

¿Su investigación involucra el uso de sustancias o procesos (o tecnologías) que pueden causar daño al medio ambiente, a los animales o plantas (durante la implementación de la actividad o posterior al uso de los resultados, como un posible impacto)?

Para investigaciones que involucren experimentos con animales, complete también la sección 5.

¿Esta actividad trata sobre fauna y/o flora en peligro de extinción/áreas protegidas?

¿Esta actividad implica el uso de sustancias o procesos (o tecnologías) que pueden causar daño a los seres humanos, incluidos los que realizan la actividad (durante la implementación de la actividad o posterior al uso de los resultados, o el despliegue de la tecnología como un posible impacto)?

Para actividades que involucran participantes humanos, ver la sección 2.

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14
Observacións		Página	17/18
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==		
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).		



Inteligencia artificial

¿Esta actividad implica el desarrollo, despliegue y/o uso de sistemas basados en Inteligencia Artificial?

 Doble uso

¿Su investigación tiene potencial para aplicaciones militares?

 Mal uso

¿Su investigación tiene el potencial para el abuso malévolo/criminal/terrorista?

 Otras cuestiones éticas

¿Existen otras cuestiones éticas que deban tenerse en cuenta? En caso afirmativo, especifique:

Si ha identificado algún problema ético anterior, describa con más detalle cómo pretende abordarlo dentro de su proyecto:

Firma del/a representante de la entidad

Código Seguro De Verificación	XQw/AQhrGJWiaD8vd6nJCA==	Estado	Data e hora	
Asinado Por	Ana López Cheda	Asinado	08/11/2023 10:31:14	
Observacións		Página	18/18	
Url De Verificación	https://sede.udc.gal/services/validation/XQw/AQhrGJWiaD8vd6nJCA==			
Normativa	Este informe ten o carácter de copia electrónica auténtica con validez e eficacia administrativa de ORIGINAL (art. 27 Lei 39/2015).			