Especificaciones Revisadas para la Inclusión de Marcadores Biomédicos de Imagen Médica

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
ISC Instituto de Salud Carlos III

IMPaCT
Infraestructura de Medicina de Precisión asociada a la Ciencia y la Tecnología

Programa	IMPacT: Infraestructura de Medicina de Precisión asociada a la Ciencia y la Tecnología			
Nombre Proyecto	IMPaCT-Data: Programa de Ciencia de Datos de IMPaCT			
Expediente	IMP/00019			
Duración	Enero 2021 – Diciembre 2023			
Paquete Trabajo	WP5. Integración de Datos			
Tarea	Tarea 5.2 – Implantación de mecanismos para la identificación automática de información genómica relevante, o contenido de las imágenes médicas que den respuestas anonimizadas al estilo beacon.			
Entregable	E5.5. Especificaciones Revisadas para la Inclusión de Marcadores Biomédicos de Imagen Médica			
Versión	Número de versión del documento			
Fecha Entrega	30/06/2022		***Fecha Aprobado***	17/05/2023
Responsable	FPS			
Nivel Diseminación	X	PU	Público	
		CO-IMP	Confidencial, sólo participantes de los pilares de IMPaCT, incluyendo la comisión de evaluación de IMPaCT.	
		CO-DATA	Confidencial, sólo participantes de IMPaCT-Data, incluyendo la comisión de evaluación de IMPaCT.	

Autores		
Organización	***Nombre***	***Rol***
FPS	Miguel Ángel Armengol de la Hoz	Coordinador/Autor
UMIB-FISABIO	Maria de la Iglesia Vayá	Coordinador/Autor
UMIB-FISABIO	Silvia Nadal	Revisor
INiBICA	Marcial García Rojo	Revisor
INiBICA	Irene Bernal Florindo	Autor
La Fe	Carina Soler	Autor
La Fe	Leonor Cerdá Alberich	Autor
Osasunbidea	Javier Gorricho Mendívil	Autor
IACS	Carlos Tellería	Autor
CLINIC- BCN	Antonio Lopez Rueda	Autor / revisor
Sant Joan de Deu	Josep Munuera	Autor / revisor
CLINIC - BCN	Guillem Bracons Cucó	Autor/Revisor
UMA	Juan Antonio García Ranea	Autor

Historial de versiones			
Nro.	**Fecha**	**Descripción**	**Autor**
v 0.1	11/10/2022	Borrador del índice	Miguel Ángel Armengol (FPS)
v 0.2	11/11/2022	Segunda versión del Índice	María de la Iglesia Vayá (UMIB-FISABIO)
v 0.3	20/11/2022	Procesados comentarios revisores	María de la Iglesia Vayá (UMIB-FISABIO)
v 0.4	30/12/2022	Desarrollo del contenido de las secciones asignadas	Todos los autores
v 0.5	31/12/2022	Revisión inicial de todo el documento	Miguel Ángel Armengol (FPS)
v 0.6	09/02/2023	Procesar comentarios revisores y revisión del documento	Silvia Nadal (UMIB-FISABIO)
V 0.7	09/02/2023	Completar conclusiones e incorporación RQS	María de la Iglesia Vayá (UMIB-FISABIO)
V 1.0	09/02/2023	Revisión final	María de la Iglesia Vayá (UMIB-FISABIO)
v 1.1	17/05/2023	Cambio visibilidad a público y aprobado	Comité Dirección
v 1.1.1	14/06/2023	Cambio de formato para publicar en la Web de IMPaCT	David Velasco (ISCIII)

Contenido

Contenido....4
Tablas....5
Figuras....6
Resumen Ejecutivo....7
Introducción....8
Audiencia....8
Ámbito....8
Relación con otros Entregables....8
Estructura Entregable....8
1. Marcadores de imagen médica....9
1.1 Introducción....9
1.1.1 Definiciones de biomarcadores de imagen....9
1.1.2 Requisitos para el uso eficaz de biomarcadores de imagen....10
1.2 Tipos de biomarcadores....11
1.2.1 Marcadores de radiología....12
1.2.2 Marcadores de patología digital....13
1.2.2.1 Receptor de estrógeno (RE) y de progesterona (RP)....15
1.2.2.2 Receptor HER-2....15
1.2.2.3 Proteína Ki-67....16
1.2.2.4 Receptor EGFR....16
1.2.2.5 Oncogen KRAS....17
1.2.2.6 Gen BRAF....17
1.2.2.7 Gen ACP....18
2. Desarrollo de biomarcadores de imágenes....18
2.1 Relación de los biomarcadores de imagen con la medicina de precisión....18
2.1.1 Flujo del desarrollo de los biomarcadores....19
2.1.1.1 Desarrollo inicial de un biomarcador de imagen....20
2.1.1.2 Adquisición y análisis de un biomarcador de imagen....20
2.1.1.3 Medición de un biomarcador....22
2.1.1.4 Validación de un biomarcador....23
2.1.1.5 La radiómica como parte del flujo del desarrollo de biomarcadores....24
3. Estándares para marcadores....25
3.1 DICOM (Dicom Structured Reporting, DICOM SR)....25

3.1.1 DICOM SR 27
3.1.2 DICOM SEG 27
3.1.3 DICOM RWVM 27
3.1.4 DICOM RT 28
3.2 CIE 28
3.2.1 CIE-10 29
3.2.2 CIE-11 29
3.2.3 CIE-O 30
3.3 RADLEX 30
3.4 UMLS 31
3.5 Indexación con distintas ontologías 32
3.5.1 La ontología del fenotipo humano (HPO) 32
3.5.2 Mondo 32
3.5.3 Mapeo entre ontologías 33
4. Grupos internacionales de trabajo 33
4.1.1 European Imaging Biomarkers Alliance (EIBALL) 33
4.1.2 Quantitative Imaging Biomarkers Alliance (QIBA) 34
5. Biobancos de imágenes y biomarcadores de imágenes basados en imagen poblacional 34
5.1 Introducción 34
5.2 Solicitud y acceso a los biomarcadores 35
5.3 Casos de éxito 35
6. Casos de uso mínimos a abordar 39
7. Conclusiones 41
Referencias 44
Acrónimos y Abreviaturas 47

Tablas

Tabla 1: Biomarcadores de imagen en radiología en función del proceso biopatológico o suceso que lo domina. 13

Figuras

Figura 1: Mapa de procesos para el establecimiento y validación de un biomarcador de imagen. En la parte izquierda de la figura se representan los ítems médico-asistenciales y en la derecha los metodológicos. Fuente: https://www.elsevier.es/es-revista-radiologia-119-articulo-biomarcadores-imagen-imagen-cuantitativa-bioingenieria-S0033833811001123?newsletter=true 19
Figura 2. Entornos principales de extracción de los datos computacionales 22
Figura 3. Transformación a un objeto IOD. Fuente: https://theses.hal.science/tel-01961254/document 26
Figura 4. Grupos de términos de Radlex 31

Resumen Ejecutivo

El objetivo de este documento es proporcionar una guía para la investigación con datos de imagen médica basada en la extracción de biomarcadores a partir de ellas. Además de describir múltiples tipos de biomarcadores de imagen, también se explican los pasos mediante los cuales estos biomarcadores son desarrollados y usados. Esta guía se completa con información de diversos estándares para la clasificación de los marcadores y con casos de uso reales.

Este documento tiene como objetivo proporcionar una guía detallada para la investigación con datos provenientes del campo de la imagen médica, enfocándose en la extracción de biomarcadores de imágenes. La propuesta abarca desde la descripción de los diferentes tipos de biomarcadores hasta la descripción de los procesos necesarios para su desarrollo y uso en la investigación médica. Además, se incluye información sobre los estándares para la clasificación de estos marcadores, así como ejemplos de su aplicación en la medicina de precisión y orientado a facilitar este recurso a cualquier persona interesada en comprender y aplicar la tecnología de biomarcadores de imagen en la investigación médica.

La información contenida en este entregable ofrece una visión general y una comprensión profunda de la importancia y el potencial de los biomarcadores de imagen en el ámbito de la investigación médica.

En resumen, este documento brinda una guía completa para la investigación con biomarcadores de imagen médica, abarcando desde su descripción hasta su aplicación práctica. Es un recurso que permite dotar de las claves necesarias para comprender y aplicar esta tecnología en la investigación médica y en la mejora de la atención médica y la medicina de precisión.

Introducción

Audiencia

Este entregable está dirigido a aquellas personas cuyo trabajo de investigación se centre en la construcción de biomarcadores en datos de imagen.

Ámbito

Este documento pretende ser una referencia para la creación y uso de biomarcadores de imagen y Biobancos.

Relación con otros Entregables

Este entregable forma parte de la Tarea 5.2 —Implantación de mecanismos para la identificación automática de información genómica relevante, o contenido de las imágenes médicas que den respuestas anonimizadas al estilo beacon— en el Paquete de Trabajo 5 —Integración de Datos—. Este entregable guarda relación con el entregable 5.1 —Técnicas de Integración de Datos Biomédicos—.

Estructura Entregable

El entregable tiene la siguiente estructura general: (1) Marcadores de imagen médica, (2) Desarrollo de biomarcadores de imágenes, (3) Estándares para marcadores, (4) Grupos internacionales de trabajo, (5) Biobancos de imágenes y biomarcadores de imágenes basados en imagen poblacional y (6) Casos de uso mínimos a abordar.

En la primera sección se describen los biomarcadores de imagen médica y los tipos que existen; a continuación se describen los pasos para la creación de biomarcadores de imagen; posteriormente se enumeran y explican distintos estándares y ontologías para marcadores; en el cuarto apartado se presentan los grupos internacionales de trabajo EIBALL y QIBA; en el quinto apartado se habla del uso de biobancos de imagen y se listan casos de éxito; finalmente, se listan los casos de uso mínimos a abordar.

1. Marcadores de imagen médica

1.1 Introducción

En los últimos años, el desarrollo de nuevas técnicas de imagen médica ha permitido generar imágenes digitales de alta calidad que contienen una amplia diversidad de información. De manera concordante en el tiempo, la enorme expansión de las capacidades de computación ha mejorado todavía más las imágenes utilizadas para diagnosticar y guiar el tratamiento de numerosas enfermedades.

Actualmente, la interacción multidisciplinaria entre las imágenes digitales y su procesado computacional está definiendo nuevas técnicas de adquisición y aplicando diversos modelos matemáticos que simulan la realidad biológica y fisiológica, y que permiten obtener medidas precisas y reproducibles de la situación de los pacientes. Todo esto está permitiendo desarrollar biomarcadores de imagen que aportan una información cuantitativa que no puede detectarse ni medirse visualmente en las imágenes originales, abriendo el campo de la imagen médica tradicional a otras disciplinas como la ingeniería y la física.

De hecho, los biomarcadores de imagen presentan una doble ventaja. Por un lado, representan variables numéricas que caracterizan y cuantifican diferentes parámetros, extraídos de las imágenes médicas, relevantes para una enfermedad específica. Por otro lado, se pueden generar imágenes paramétricas para representar gráficamente los valores de cada uno de los biomarcadores o parámetros que se calculan a partir del postproceso de las imágenes originales adquiridas. Esto permite analizar la distribución espacial del biomarcador en la muestra observada mediante su representación visual.

El enorme potencial de los biomarcadores de imagen permite que este campo de investigación sea uno de los más activos, permitiendo visualizar y medir procesos fisiológicos y biológicos a partir de modelados tridimensionales de una región de interés en un paciente concreto.

1.1.1 Definiciones de biomarcadores de imagen

Los **biomarcadores de imagen** se definen como todas aquellas características extraídas de las imágenes adquiridas de un sujeto, que pueden medirse y evaluarse de forma objetiva, y que se comportan como indicadores de procesos biológicos normales, cambios patológicos o respuestas a tratamientos [1]. En los últimos años se ha demostrado que los biomarcadores de imagen, desde simples mediciones de tamaño o forma hasta la aplicación de complejos modelos computacionales, ofrecen una información complementaria muy útil al diagnóstico tradicional para establecer la presencia de una alteración o lesión, medir su situación biológica, definir su historia natural y progreso, estratificar las anomalías en fenotipos, y evaluar los efectos de un tratamiento.

Concretamente, se estudian los tejidos biológicos a partir de sus imágenes médicas digitales con el objetivo de describir, con exactitud y veracidad, aquellos fenómenos considerados

como relevantes para el paciente. Para ello, se necesita conocimiento para analizar, mediante procesos computacionales, los cambios biológicos tisulares asociados a la enfermedad desde sus imágenes médicas y así poder detectar, clasificar y representar estas anomalías. Los datos extraídos por análisis computacional *in-silico* se emplean para construir modelos descriptivos y predictivos basándose en las correlaciones con la realidad biopatológica y los resultados clínicos finales. Es decir, estos modelos basados en señales o imágenes son representaciones simbólicas de un proceso real, definidos como una serie de relaciones o ecuaciones entre las distintas variables que intervienen en el mismo. Pero, antes de considerarlos como biomarcadores, los datos de imagen computacional extraídos como descriptores deben validarse. Únicamente cuando hayan demostrado ser eficaces y efectivos, los biomarcadores de imagen pueden utilizarse como objetivos sustitutos para definir una enfermedad, su situación y su respuesta a un tratamiento.

En efecto, para desarrollar un biomarcador de imagen es necesario realizar una serie de pasos destinados a validar su relación con la realidad estudiada y controlar su validez, tanto clínica como técnica. Este proceso, que se describe en mayor detalle en el punto 2.1.1., incluye la definición previa de unas pruebas de concepto y de mecanismo; la adquisición estandarizada y optimizada de imágenes anatómicas, funcionales y moleculares; el análisis de los datos mediante modelos computacionales; la visualización adecuada de los resultados; la obtención de medidas estadísticas apropiadas; y la realización de pruebas de principio, eficacia y efectividad [2].

1.1.2 Requisitos para el uso eficaz de biomarcadores de imagen

El proceso necesario para incorporar un biomarcador de imagen, tanto en la práctica clínica como en un ensayo clínico, es complejo y debe responder a criterios de coherencia conceptual, reproducibilidad técnica, sensibilidad y especificidad. La definición del problema, la adquisición de las imágenes fuente, la metodología de análisis y la extracción del biomarcador a partir de modelos computacionales y herramientas de medida, son aspectos imprescindibles para utilizarlos adecuadamente en el estudio de una enfermedad. Básicamente, el desarrollo de un biomarcador implica validar su relación con la realidad objetivada (estructural, fisiológica, biológica o molecular) y controlar su validez técnica.

Por tanto, los requisitos que deben cumplir los biomarcadores de imagen para que puedan ser usados de forma eficaz pueden resumirse de la siguiente manera [3]:

Un biomarcador de imagen debe:

- adquirirse desde equipos tecnológicamente estables, para evitar fuentes de variabilidad en la adquisición;
- obtenerse por métodos incruentos y seguros para el paciente, a partir de las imágenes médicas;
- estar disponible de forma generalizada, para garantizar su difusión y utilización universal;

- ser reproducible, es decir, tener la capacidad de replicar el valor obtenido, de forma que se obtengan siempre los mismos resultados en centros con equipamientos diferentes;
- estar estandarizado en cuanto a la adquisición de la imagen (parámetros técnicos), preparación y proceso (filtrados, mejora de resolución de señal-ruido y de resolución espacial), análisis y modelado de la señal (ajuste de modelos computacionales y extracción de variables);
- estar validado con respecto a su precisión y eficacia;
- ser clínicamente útil, permitiendo una mejora clínica medible;
- tener una alta sensibilidad, para mejorar la relación entre la variación del biomarcador y la del efecto que se pretende medir;
- tener un alto porcentaje de verdaderos positivos, es decir, poder clasificar correctamente como anormal un hallazgo realmente alterado;
- tener una alta especificidad, que es la probabilidad de que un sujeto verdaderamente sano obtenga un resultado negativo en la prueba;
- estar, además, muy vinculado al proceso biológico o la alteración patológica con la que se relaciona;
- tener un coste lo más bajo posible, tanto económico (barato) como de dedicación temporal (rápido).

1.2 Tipos de biomarcadores

En líneas generales, los biomarcadores se pueden clasificar según su objetivo final en [4]:

- **Biomarcadores de diagnóstico/fenotipado**: aquellos que detectan o confirman la presencia de un proceso patoló-gico, una enfermedad concreta o un subtipo de la enfermedad. Estas métricas son marcadores subrogados de agresividad, incluyendo tiempo de duplicación de volumen, invasión local y ex-tensión a distancia. En este grupo están los bio-marcadores de susceptibilidad, que son aquellos cuya presencia identifica el potencial para desa-rrollar una enfermedad.
- **Biomarcadores predictivos de mutaciones o variantes en el genoma**: son los que permiten predecir la presencia de una variante o mutación clínicamente relevante en el genoma, a partir de imágenes microscópicas con tinciones convencionales o con técnicas inmunohistoquímicas (IHC), o a partir de otras imágenes clínicas (endoscopias, radiología, etc.).
- **Biomarcadores predictivos de tratamiento**: características que señalan a los indi-viduos con más probabilidades de experimentar un efecto favorable (o desfavorable) con el tra-tamiento. Se consideran en esta categoría a los biomarcadores de respuesta, cuando muestran el efecto biológico generado por el tratamiento; y de seguridad, cuando indican la probabilidad, pre-sencia o grado de toxicidad como efecto adver-so al tratamiento. Estas métricas son marcadores subrogados relacionados con el tratamiento, que in-cluyen aspectos como ratios coste/beneficio (coste material), riesgo/beneficio (toxicidad evi-table) y oportunidad/beneficio (tiempo a efecto).

- **Biomarcadores pronósticos**: valores que determinan la probabi-lidad o el tiempo a un evento clínico, a una re-currencia o a la progresión de la enfermedad, y la supervivencia del paciente. Estas métricas son marcadores subrogados relacionados con la res-puesta clínica objetiva.

Además, dentro de cada especialidad médica y según los procesos biológicos que se quieren estudiar, se desarrollan diferentes tipos de biomarcadores. A continuación, se van a exponer ejemplos concretos en el campo de la radiología y la anatomía patológica.

1.2.1 Marcadores de radiología

Atendiendo a las imágenes médicas radiológicas y de medicina nuclear, la radiómica es uno de los campos de investigación más empleados en los últimos años. Se basa en la extracción, el almacenamiento y el análisis de alto rendimiento de una gran cantidad de características cuantitativas extraídas de las imágenes médicas para descubrir nuevos biomarcadores de imagen que proporcionan información cuantitativa (biopsias virtuales) para el diagnóstico temprano de enfermedades, fenotipado de enfermedades, clasificación de enfermedades, terapias dirigidas y evaluación de la respuesta de la enfermedad al tratamiento [5].

Estas características cuantitativas pueden clasificarse, tal y como se detalla en el *entregable E4.5. Comparación de Técnicas de Gestión de Información de Imagen Médica*, en diferentes niveles en función de su complejidad:

- Las características morfológicas describen propiedades semánticas, como la forma de la región de interés trazada y sus propiedades geométricas, como el volumen, el diámetro máximo a lo largo de diferentes direcciones ortogonales, la superficie máxima, la compacidad del tumor y la esfericidad.
- Las características estadísticas de primer orden describen la distribución de los valores de las intensidades de los vóxeles individuales, sin preocuparse por las relaciones espaciales. Estas propiedades están basadas en distribuciones de histograma y proporcionan información sobre la media, la mediana, el máximo y el mínimo de las intensidades de los vóxeles en la imagen, así como de su asimetría, curtosis, uniformidad y aleatoriedad (entropía). Dependiendo de la modalidad de imagen empleada, estas características adquieren un significado concreto que permiten aumentar su explicabilidad.
- Las características estadísticas de segundo orden se obtienen calculando las interrelaciones estadísticas entre vóxeles vecinos. Proporcionan una medida de la disposición espacial de las intensidades de vóxel y, por lo tanto, de la heterogeneidad dentro de la región de interés. Tales características pueden derivarse, por ejemplo, de la matriz de co-ocurrencia de nivel de gris (GLCM), que cuantifica la incidencia de vóxeles con las mismas intensidades a una distancia predeterminada a lo largo de una dirección fija, o de la matriz de longitud de ejecución de nivel de gris (GLRLM), que cuantifica vóxeles consecutivos con la misma intensidad a lo largo de direcciones fijas.
- Las características de orden superior se refieren a aquellas características estadísticas que no se obtienen directamente de las imágenes originales, sino que se

extraen después de aplicar diferentes filtros o transformaciones matemáticas. Se obtienen, por ejemplo, con el objetivo de identificar patrones repetitivos, suprimir ruido o resaltar detalles. Estos incluyen métodos como el análisis fractal, el funcional de Minkowski, la transformada de tipo "wavelet" y la transformada laplaciana aplicada a imágenes procesadas por un filtro gaussiano, a partir de los que se pueden extraer regiones con patrones de textura cada vez más complicados.

Aunque la radiómica es una de las disciplinas más extendidas en el área de biomarcadores e imagen médica y, en particular, en oncología debido al apoyo de la Red de Imágenes Cuantitativas (QIN) del Instituto Nacional del Cáncer (NCI) y otras iniciativas del NCI, se pueden encontrar una gran variedad de biomarcadores de imagen que pueden aplicarse a multitud de condiciones y áreas de la medicina. En la Tabla 1, se muestran a modo de resumen algunos tipos de biomarcadores más utilizados.

Tabla 1: Biomarcadores de imagen en radiología en función del proceso biopatológico o suceso que lo domina.

Proceso biopatológico dominante	Biomarcadores de imagen	Modalidad de imagen
Tumores	Radiómica / Radiómica profunda	RM, TC, PET, SPECT, RX, US, mapas paramétricos
	Parámetros de difusión (ADC, IVIM, Kurtosis)	RM (DWI)
	Parámetros farmacocinéticos (Ktrans, ve, vp)	RM (DCE)
Inflamación	Área / Volumen de edema	RM (STIR)
Depósitos	Concentración de hierro (R2*)	RM (T2* multieco)
	Concentración de grasa (PDFF)	RM (T2* multieco con información de fase)
	Mapas cuantitativos de susceptibilidad (QSM)	RM (T2* multieco con información de fase)
	Volumen extracelular (ECV) en tejido (fibrosis, isquemia...)	RM (T1 pre- y post- contraste)
Morfología	Segmentación de órganos/lesiones	RM, TC, PET
	Espesor cortical	RM (T1)
	Volumetría	RM (T1)
	Esqueletonización	RM, TC, RX
	Tensor de difusión	RM (DTI)
	Parámetros trabeculares (BV/TV, TbTh, TbSp, Tn, D2D, D3D, QTS)	RM, TC
Funcional/Metabólica	Conectividad funcional	RM funcional
	SUV, MTV, TLG	PET

1.2.2 Marcadores de patología digital

La medicina personalizada persigue poder tratar a cada paciente de una forma clínicamente individual. En los últimos años, las terapias dirigidas contra muchos tumores malignos han obtenido altas tasas de respuestas clínicamente significativas, y han permitido a los oncólogos desarrollar estrategias individuales para cada paciente. La clave de estos logros

se encuentra en la identificación correcta de la molécula diana tumoral contra la que hay que dirigirse (la molécula que hay que “atacar”).

En tumores sólidos, una lista cada vez mayor de moléculas dianas se evalúan rutinariamente mediante la técnica inmunohistoquímica (IHC) en biopsias y en material tumoral extraído de cirugías, aportando información clave para la toma de decisiones clínicas e información pronóstica relevante. Los biomarcadores establecidos actualmente ayudan en la detección temprana del cáncer y en seleccionar a los pacientes adecuados para determinados tratamientos. En patología, los biomarcadores se pueden detectar fundamentalmente mediante técnicas IHC, métodos proteómicos cuantitativos y métodos de biología molecular, que incluyen técnicas de PCR y de secuenciación del material genético.

En general, los biomarcadores de imagen en patología digital se aplican en tres grandes grupos de imagen microscópica en campo claro: las tinciones convencionales (hematoxilina eosina [HE] en biopsias y Papanicolaou en citología), la inmunohistoquímica (IHC) y las tinciones especiales (histoquímica) y las imágenes de patología molecular, generalmente de hibridación in situ (ISH). En microscopía de fluorescencia, podemos tener biomacadores en inmunofluorescencia in situ (especialmente interesantes en piel y en riñón) y en hibridación in situ fluorescente (FISH).

La IHC representa una tecnología muy importante para el desarrollo de nuevos biomarcadores en la estratificación de pacientes y en la medicina personalizada. Sin embargo, esta técnica presenta cierta subjetividad y es propensa a errores, ya que varios estudios en diferentes biomarcadores obtuvieron poca reproducibilidad entre patólogos. Aplicar la patología digital mediante el análisis de imagen permite usar los datos cuantitativos de la imagen digital para extraer datos numéricos sobre densidad y distribución de la expresión de los biomarcadores en la muestra de tejido. La inteligencia artificial proporciona objetividad y repetibilidad, permitiendo identificar cambios sutiles en la expresión de los biomarcadores que a simple vista podrían pasar desapercibidos.

Actualmente, hay varias herramientas de análisis de imagen disponibles comercialmente. Estas herramientas permiten medir cuantitativamente los biomarcadores IHC, extrayendo los datos cuantitativos de su expresión. Muchos estudios han usado ya con éxito el análisis de imagen para medir la expresión IHC en un amplio rango de biomarcadores. La mayoría de los sistemas comerciales ofrecen plataformas de desarrollo de análisis de imagen que permiten construir algoritmos para cualquier biomarcador candidato, donde su desarrollo y validación están en manos del usuario. Sin embargo, están surgiendo algoritmos con objetivos específicos que han sido probados y validados comercialmente para un propósito concreto, y que bajo condiciones controladas funcionan sin ninguna intervención del usuario. Por ejemplo, varias empresas (Leica, Roche) tienen autorización de la “Food and Drug Administration” (FDA) para el uso de algoritmos específicos en la detección del receptor de estrógeno (RE), receptor de progesterona (RP), HER-2 y Ki-67. Estos algoritmos generalmente están ligados a instrumentos específicos de escaneado, y deben usarse exactamente como lo indica el fabricante. Actualmente se usan de forma rutinaria en el diagnóstico.

Varios de los biomarcadores usados en patología digital se detallan a continuación.

1.2.2.1 Receptor de estrógeno (RE) y de progesterona (RP)

Uno de los biomarcadores más relevantes es el receptor de estrógeno (RE), expresado en el 70-75% de los carcinomas de mamas infiltrantes. Los tumores que expresan estas moléculas dependen en gran medida de la disponibilidad de estrógeno para su supervivencia. La expresión de RE es un importante biomarcador de carácter predictivo, permitiendo seleccionar qué pacientes se beneficiarán de un tratamiento hormonal. El nivel de expresión de RE se determina mediante IHC en el tejido de la biopsia y el punto de corte para considerar el estado positivo es que el porcentaje de células tumorales positivas para este receptor sea mayor o igual al 1%.

Los tumores positivos para RE suelen mostrar expresión de otros genes, entre ellos el receptor de progesterona (RP). Esta molécula pertenece a la familia del receptor de hormonas nucleares de factores de transcripción dependiente de ligandos. La presencia de progesterona y estrógeno producen la activación de RP, transcribiendo genes implicados en la supervivencia y proliferación celular. Esta molécula se expresa en aproximadamente la mitad de los cánceres de mama ER positivos y en muy rara ocasión en los ER negativos. El estado de RP se considera un biomarcador pronóstico; su expresión en los tumores ER positivos está asociada con un mejor pronóstico y menor riesgo de recaída comparado con aquellos que son PR negativos. Al igual que ocurre con el RE, se considera que el tumor es RP positivo cuando el porcentaje de células tumorales que lo expresan es mayor o igual al 1%.

La IHC está aceptada como método estándar de evaluación para determinar los estados de RE y PR. El “College of American Pathologist” (CAP) y “American Society of Clinical Oncology” (ASCO) desarrollaron y publicaron una guía para mejorar la precisión de los ensayos IHC para determinar RE y RP en cáncer de mama. En esta guía se detallan los procesos de validación de estos ensayos, ya que la precisión y evaluación cuantitativa de los resultados es crucial. Sin embargo, varios factores pueden afectar a los resultados de los estados de RE y RP en cáncer de mama, incluyendo la fijación del tejido, la elección del anticuerpo usado, el proceso de tinción (manual o automatizado) y la elección de los umbrales para considerar los resultados positivos. Para mejorar la calidad de los ensayos a escala global es necesario que todos los laboratorios implementen los ensayos estandarizados y validados recomendados por CAP/ASCO para la evaluación de RE y RP.

La microscopía automatizada y el procesamiento computacional han proporcionado mayor precisión en la cuantificación y estandarización en IHC. Diferentes grupos han publicado datos de la evaluación automática de estos receptores en el cáncer de mama, describiendo una correlación excelente entre el análisis manual y automático. Varios sistemas de análisis automático de imagen ofrecen la posibilidad de una puntuación rápida y objetiva de los núcleos teñidos, pero sólo algunos tienen la aprobación de la FDA.

1.2.2.2 Receptor HER-2

Este receptor tiene un dominio de unión al ligando extracelular, un dominio transmembrana y uno intracelular tirosin quinasa. Junto con los receptores hormonales descritos anteriormente, es probablemente el biomarcador predictivo y de pronóstico más importante en el cáncer de

mama. Está implicado en varias rutas de señalización celular, regulando la proliferación y la supervivencia.

En el cáncer de mama, la actividad de este gen aumenta por amplificación del mismo, estimulando el crecimiento tumoral. Esta sobreexpresión se produce en aproximadamente el 20% de los cánceres de mama y se asocia con una biología más agresiva de la enfermedad y un mal pronóstico; mayor riesgo de recaída y mortalidad. Con la llegada de las terapias dirigidas, el pronóstico de estos pacientes ha mejorado, ya que se han desarrollado fármacos contra el receptor de HER-2. El estado de HER-2 predice la respuesta a determinados agentes quimioterapéuticos.

El estudio de este biomarcador pronóstico se puede hacer mediante técnicas de IHC o ISH (evalúa el número de copias del gen). Cuando se usa la IHC, la expresión de HER-2 se clasifica de 0 a +3. Aquellas que se encuentran entre 0 y +1 se consideran negativas (aunque, recientemente, se ha observado que los tumores con baja expresión de HER2 responden de forma distinta a ciertos tratamientos), y las que se encuentran entre +2 y +3 se consideran positivas. Si se usa la técnica ISH, se considera positivo cuando el número de copias del gen HER2 es superior a 4.

1.2.2.3 Proteína Ki-67

Esta molécula es un antígeno nuclear asociado a la proliferación celular, es un biomarcador pronóstico. El estado de Ki-67 se informa como el porcentaje de células invasivas que presentan función nuclear positiva. En cáncer de mama, si la evaluación de este parámetro resulta en un valor menor al 10%, estamos ante un pronóstico favorable. Si la evaluación de este parámetro resulta en un valor superior al 20%, estamos ante un pronóstico desfavorable. Si la evaluación se encuentra entre el 10% y el 20%, nos encontramos en una categoría intermedia. La fracción del Ki-67 está relacionada con el grado del tumor en la mayoría de los cánceres, siendo el mayor porcentaje en el grado 3 de carcinomas invasivos. En el carcinoma de mama, los tumores RE y PR negativos tienden a tener también un mayor porcentaje de expresión del Ki-67. Aún existe bastante debate sobre cómo evaluar correctamente este biomarcador inmunohistoquímico, no solo en cáncer de mama sino en múltiples tipos de neoplasias.

1.2.2.4 Receptor EGFR

Los receptores del factor de crecimiento epidérmico se encuentran entre los biomarcadores de cáncer más estudiados debido a su actividad oncogénica en muchos tipos de tumores, por lo que su expresión está correlacionada con una progresión de la enfermedad y una supervivencia pobre. Hay 2 tipos de agentes que tienen acción contra este receptor: anticuerpos monoclonales anti-EGFR y pequeñas moléculas inhibidoras del dominio tirosin quinasa del receptor EGFR.

El gen EGFR tiene mucha importancia en el cáncer de pulmón de células no pequeñas (NSCLC). La identificación de los subtipos moleculares de esta enfermedad y la selección de los pacientes para el tratamiento adecuado es crucial en la terapia personalizada del NSCLC.

La guía clínica para esta enfermedad ha incorporado el test genético para evaluar el gen EGFR, y así identificar pacientes que se beneficiarían de una terapia oral dirigida en lugar de la administración de la quimioterapia. Los adenocarcinomas de pulmón presentan frecuentemente mutaciones en este gen y un mayor número de copias. En Estados Unidos, aproximadamente el 15% de los pacientes con adenocarcinoma de pulmón presentaban mutaciones que activaban el gen EGFR. La mayoría de esas mutaciones se encuentran en los exones 19 y 21 del gen. En aproximadamente el 10% de los pacientes occidentales y en el 50% de los pacientes asiáticos con adenocarcinoma presentan mutaciones en EGFR.

El carcinoma de células grandes rara vez tiene mutaciones activadoras del gen EGFR, siendo más frecuente en el subtipo papilar. La evaluación del gen EGFR está recomendada para los subtipos adenocarcinoma, carcinoma de células grandes y NSCLC.

Aunque este biomarcador es genético, es decir, originalmente no es un biomarcador de imagen, se han desarrollado algoritmos de IA que utilizando imágenes de hematoxilina-eosina (HE) son capaces de predecir si el tumor presenta alguna mutación clínicamente relevante en el gen EGFR.

1.2.2.5 Oncogen KRAS

El oncogen KRAS es crítico en el desarrollo de varios cánceres. Sus mutaciones son muy frecuentes en muchos tipos de cánceres, incluidos el de páncreas (65%), adenoma colorectal (30%), cáncer colorectal (del 30% al 50%), cáncer de pulmón (20%) y en cáncer de ovario (15%). El estado de mutación del gen KRAS tiene mucho valor en la ayuda de toma de decisiones terapéuticas. En el cáncer colorectal, mutaciones de este gen en zonas críticas, como los codones 12 y 13, son predictores negativos de la respuesta a anticuerpos anti-EGFR. El estado de este gen se determina por varios métodos, incluyendo reacciones de biología molecular como la PCR y las técnicas de secuenciación de ácidos nucleicos en las muestras del tumor.

Algunos algoritmos de IA, aplicados a imágenes de HE, son capaces de predecir la presencia de mutaciones en el gen KRAS tanto en cáncer colorectal como en otras neoplasias.

1.2.2.6 Gen BRAF

Este gen codifica para una proteína activadora de la ruta de señalización MAP quinasa/ERK, y está codificada en el cromosoma 7q34. Se han identificado más de 30 mutaciones en este gen relacionadas con cánceres en humanos. Las frecuencias de estas mutaciones varían desde un 80% en melanomas hasta un 0-18% en otros tumores, como en el del cáncer de pulmón (1-3%) y cáncer colorectal (5%). Las mutaciones más comunes son V600E y V600K. El desarrollo de terapias dirigidas contra melanomas va desde el descubrimiento de la molécula diana al desarrollo del fármaco. PLX4032 es relevante, suprimiendo la ruta oncogénica activada al inhibir la cascada de quinasa ERK.

1.2.2.7 Gen ACP

Este gen se encuentra mutado en la mayoría de los cánceres colorectales. La proteína codificada resultante juega un papel fundamental en muchos procesos celulares que determinan si una célula se desarrolla a tumoral. Aproximadamente entre un 3-5% de los cánceres de colon están relacionados con síndromes de cánceres de colon hereditarios de alto riesgo. "*Familial Adenomatous Polyposis*" (FAP) es un síndrome autosómico-dominante que afecta al >1% de todos los cánceres colorectales. Más del 90% de las familias afectadas por este síndrome tienen mutaciones en este gen. El cribado genético para FAP implica normalmente la secuenciación del ADN, normalmente precedida por otros métodos complementarios que identifican correctamente hasta el 95% de las mutaciones.

En algunos casos es necesario conocer el estado de una serie de proteínas o de genes en su conjunto, como en el caso de la inestabilidad de microsatélites o en los déficits de reparación de ADN.

2. Desarrollo de biomarcadores de imágenes

2.1 Relación de los biomarcadores de imagen con la medicina de precisión

La imagen médica ha contribuido notablemente al de-sarrollo de la medicina personalizada. En un principio, las diversas modalidades de imagen han contribuido en la detección, caracterización y estadiaje de numerosas lesio-nes y enfermedades. Con su utilización masiva, gran par-te de los pacientes se benefician de esta precisión a la hora de definir si padecen una enfermedad, la localización de las lesiones y la estimación de la mejor aproximación te-rapéutica. Sin embargo, y pese a la gran resolución espa-cial obtenida en las imágenes tomográficas (tales como RM -*Resonancia Magnética*-, TC -*Tomografía Computarizada*--- y PET -*Tomografía por Emisión de Positrones*-), la información que estas pruebas gene-ran no es completa y presenta algunas lagunas para poder tomar decisiones individualizadas en base a la evidencia actual [6]. El desarrollo de biomarcadores de imagen médica computacional pretende dar respuesta a estas necesidades.

En el caso de **radiología**, los biomarcadores de imagen son parámetros cuantitativos subrogados, que se extraen computacionalmente de las imágenes médicas radiológicas. Los datos, obtenidos píxel-a-píxel, pueden representarse como imágenes paramétricas donde la distribución y magnitud del cambio permite el análisis de la localización y heterogeneidad espacial. Además, y dado que no se destruye la muestra cuando se adquieren las imágenes, estos cambios pueden analizarse en estudios longitudinales, pudiendo considerarse como resueltas también en el tiempo.

En **anatomía patológica**, la medicina de precisión se beneficia de la existencia de los biomarcadores de imagen en cuatro aplicaciones específicas: a) la objetividad en la cuantificación de biomarcadores; b) mejora en la precisión diagnóstica y la correcta

clasificación de tipos histológicos, y, por tanto, la concordancia intra- e interobservador; c) el soporte a la decisión clínica, disminuyendo los tiempos de respuesta, especialmente en diagnósticos complejos; d) permitiendo descubrir nuevas características morfológicas, a veces no perceptibles por el ser humano, asociadas a diagnóstico, pronóstico o respuesta terapéutica.

2.1.1 Flujo del desarrollo de los biomarcadores

A continuación, se describen los pasos que conforman el flujo de desarrollo de los biomarcadores de imagen (Figura 1). Desde las imágenes adquiridas, conociendo los parámetros que se quieren calcular y los conceptos que se quieren inferir, y tras preparar las imágenes a un marco de referencia común espacial (corregistro, escalado) y de señal (filtrado de ruido y normalización de la señal), se extraen de los datos computaciones como descriptores simples o en asociación (firmas tisulares). Por último, y solo si son validados, pueden considerarse como biomarcadores de imagen.

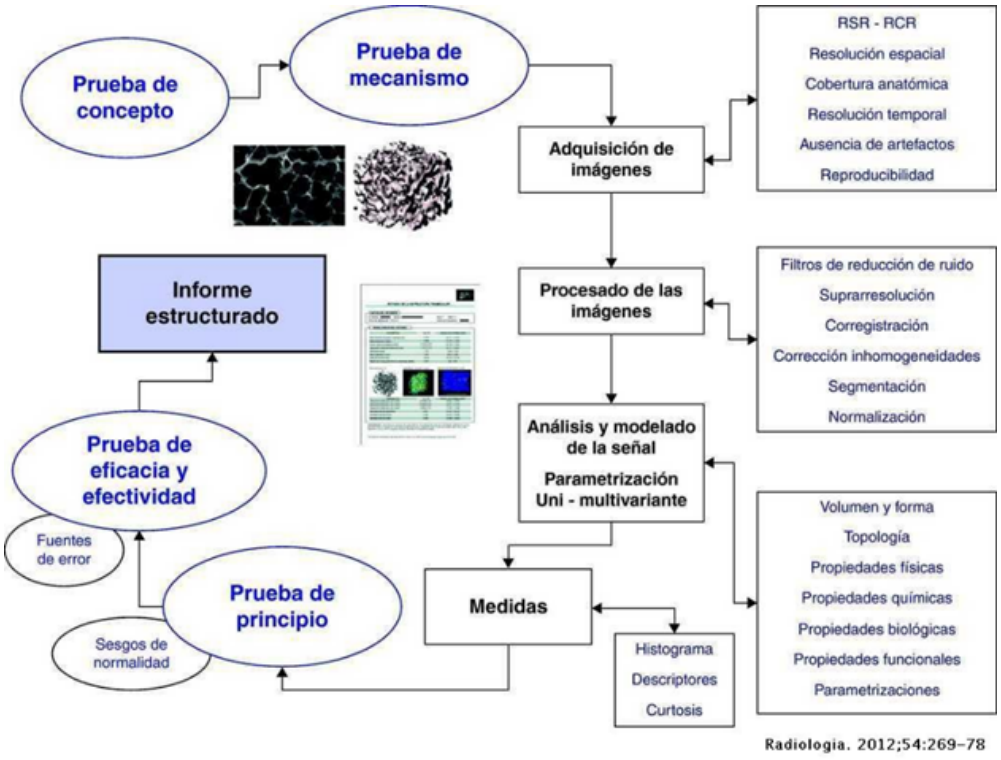


Figura 1: Mapa de procesos para el establecimiento y validación de un biomarcador de imagen. En la parte izquierda de la figura se representan los ítems médico-asistenciales y en la derecha los metodológicos. Fuente: https://www.elsevier.es/es-revista-radiologia-119-articulo-biomarcadores-imagen-imagen-cuantitativa-bioingenieria-S0033833811001123?newsletter=true

2.1.1.1 Desarrollo inicial de un biomarcador de imagen

El primer paso que se debe realizar es establecer una **prueba de concepto**, es decir, definir los conceptos por los que se quiere medir un aspecto determinado de una enfermedad. Esta prueba de concepto debe verificar que un proceso biológico determinado, entendido como una cadena de causas y efectos, se puede estudiar con las pruebas disponibles de diagnóstico por imagen y computación.

Tras definir el objetivo, se debe conocer la relación existente entre el biomarcador, el parámetro que se analiza, y la enfermedad objeto de estudio. En este sentido, la **prueba de mecanismo** establece que hay que demostrar la interrelación entre el biomarcador y el concepto, centrándose en comprender la acción (en magnitud y dirección) que una determinada enfermedad o un tratamiento tienen sobre el biomarcador. Por ejemplo, esta prueba permite establecer si la relación entre la enfermedad y un determinado biomarcador es directa o inversamente proporcional a la gravedad, entre otras.

2.1.1.2 Adquisición y análisis de un biomarcador de imagen

La definición del método y la técnica de **adquisición** de las imágenes es uno de los pasos más importantes ya que, sin unas imágenes adecuadas, es imposible extraer biomarcadores útiles. Independientemente de la técnica elegida (radiografía, ecografía, TC, RM, PET o SPECT –Single Photon Emission Computed Tomography-), se deben cuidar varios aspectos. La calidad de la imagen, medida como la relación señal-ruido (RSR), y la relación contraste-ruido (RCR), deben optimizarse para detectar el parámetro que se desea estudiar. Este paso es crítico para poder analizar de forma adecuada la señal del objeto de interés y sus cambios, facilitando la implementación de las herramientas de computación más adecuadas.

Respecto a la cobertura anatómica, debe estudiarse el órgano con la suficiente resolución espacial, definida ésta por el tamaño del vóxel. Tanto la resolución como la cobertura deben ajustarse al problema de estudio y al tamaño esperado de la alteración. Para que las imágenes sean reproducibles deben estar exentas de artefactos que degraden o corrompan la señal y mantener en el tiempo su calidad (RSR, RCR) sin modificar los parámetros de adquisición. En los estudios funcionales y dinámicos es además importante asegurar el mejor compromiso entre los requerimientos de resolución espacial y resolución temporal (frecuencia de adquisición de las imágenes de un mismo volumen), para poder observar detalladamente los procesos fisiológicos.

Para garantizar la estabilidad y el buen funcionamiento de los equipos de adquisición de imágenes es necesario realizar **controles de calidad** periódicos (aconsejable cada 2 meses) mediante maniquís donde se verifique que las variaciones de RSR, RCR, frecuencia de sintonización, relación espacial y grado de distorsión son estables y están dentro de las especificaciones del equipo proporcionadas por el fabricante.

Antes de analizar y modelar las señales, las imágenes deben **pre-procesarse** garantizando que los datos adquiridos sean óptimos para el análisis. Para ello, se debe incrementar la RSR tanto como sea posible eliminando el ruido disperso mediante los filtros adecuados [7]; homogeneizar la imagen para eliminar señales y brillos artefactuales [8] y maximizar la

resolución espacial sin afectar la señal mediante algoritmos de super-resolución [9], garantizando una coherencia espacial exacta entre todos los puntos del órgano durante todo el estudio mediante un corregistro [10] que garantice un reposicionamiento de los vóxeles para que un área de un órgano esté representada por los mismos puntos en todas las imágenes adquiridas en distintas series.

En numerosos procesos es recomendable **normalizar** la señal y el volumen del órgano estudiado a un marco de referencia que permita su comparación con otros datos poblacionales diferentes. Este proceso de normalización suele realizarse mediante algoritmos de transformación espacial, que establecen las operaciones geométricas necesarias para relacionar la estructura de interés a una plantilla de referencia que permita un análisis conjunto [11].

La **segmentación** permite clasificar los tejidos para facilitar su análisis y visualización posterior [12]. Para ello pueden emplearse varios métodos, basados generalmente en la señal del tejido a segmentar tras optimizar la adquisición, mejorar el contraste y aplicar los filtros de reducción de ruido. Para una segmentación adecuada es aconsejable alcanzar en la imagen final la mayor resolución espacial posible, minimizando el problema del volumen parcial.

El **análisis y modelado de la señal** permite extraer de las imágenes médicas digitales la información de los biomarcadores mediante los procesos computacionales adecuados. Aunque los métodos existentes son muy numerosos y sería imposible ni siquiera definirlos en este trabajo, pueden agruparse en función de los parámetros que se analizan. Así, hay métodos para calcular el volumen y la forma de un tejido (como en el carcinoma de pulmón), su topología (como en el hueso trabecular), y algunas de sus propiedades físicas (como los tiempos de relajación transversal y longitudinal), químicas (como la espectroscopia), biológicas (como la angiogénesis y la densidad celular) y funcionales (como el consumo local de oxígeno).

Una vez obtenidas estas variables para cada uno de los vóxeles de interés, puede representarse la distribución espacial del biomarcador mediante **imágenes paramétricas**. En estas imágenes los píxeles representan los valores de un determinado parámetro (morfológico o funcional) obtenido mediante el ajuste numérico de un modelo y representado habitualmente en una escala de colores diferente a la escala de niveles de gris de las imágenes fuente originales (aunque en algunos casos, como en los mapas del coeficiente aparente de difusión, la imagen paramétrica se representa también en escala de grises). Estas imágenes paramétricas también pueden representarse en escala de color como solo los vóxeles anómalos superpuestos a la imagen original, con los órganos y estructuras representados en escala de grises.

Un paso más para disminuir la redundancia de los datos adquiridos, es decir, acotar las imágenes que deben visualizarse tras optimizar la información útil, se alcanza con la utilización de **imágenes paramétricas multivariantes**. A diferencia de las imágenes paramétricas, en estas imágenes el color de cada vóxel viene determinado por el resultado de una función estadística multivariante, que es a su vez una combinación de varios parámetros o biomarcadores. Con esta aproximación se pretende valorar conjuntamente los datos para mejorar el análisis estadístico unidimensional y ayudar al radiólogo a tomar las

decisiones más adecuadas. Algunos de los métodos estadísticos multivariantes más utilizados son la regresión lineal, el análisis de función discriminante y el de componentes independientes o principales.

2.1.1.3 Medición de un biomarcador

De las imágenes paramétricas, tanto convencionales como multivariantes, se pueden obtener mediciones bien de todo el tejido u órgano estudiado, bien solo de aquellas áreas que se consideren más representativas o anormales. En este punto existen dos entornos principales de extracción de los datos computacionales (Figura 2).

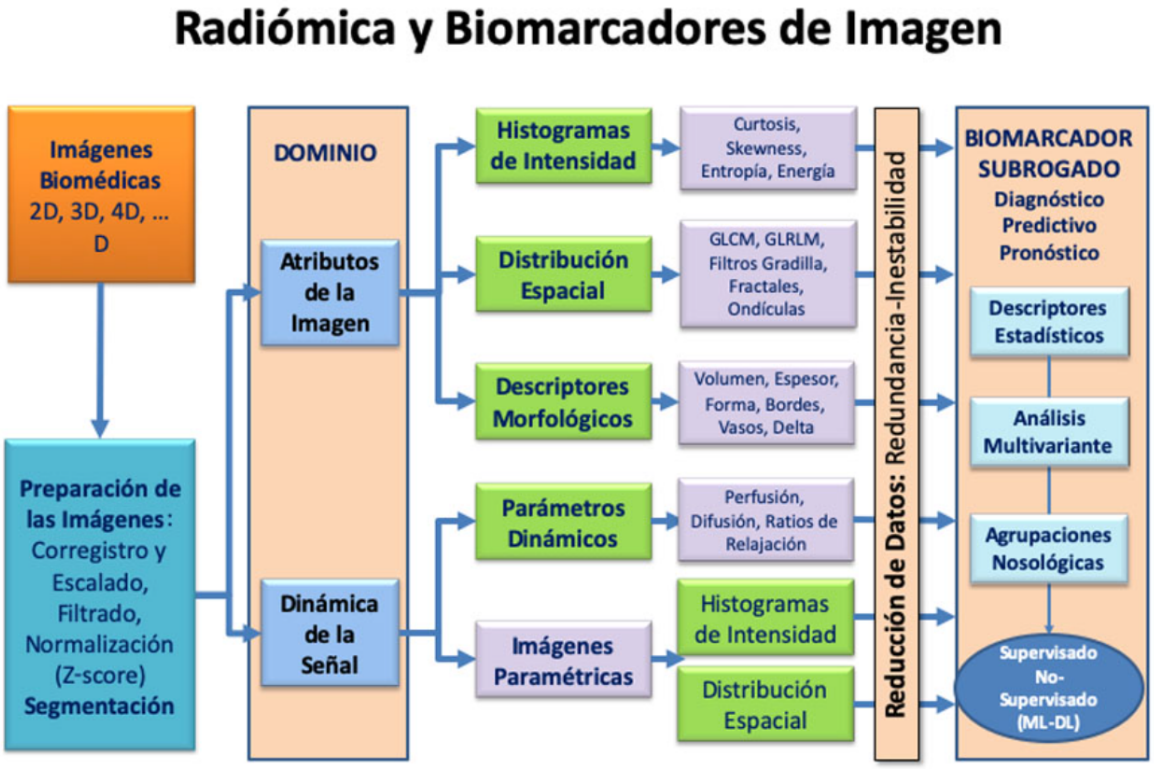


Figura 2. Entornos principales de extracción de los datos computacionales

Por un lado, está el **do-minio de los atributos de la imagen**, desde el que se computan texturas a partir del histograma de las in-tensidades (para extraer métricas como la curtosis, asimetría -skewness-, entropía, y energía), la distri-bución espacial de la señal (para obtener datos como las matrices GLCM -Gray Level Co-Occurrence Ma-trix- y GLRLM -Gray Level Run-Length Matrix-, fil-tros gradilla, fractales, y ondículas), y los descrip-tores morfológicos (entre los que podemos incluir aquellos de volumen, espesor, forma, bordes, vasos, y sus cambios temporales).

Por el otro lado, se consi-deran los parámetros obtenidos del **análisis de series dinámicas** donde existen modificaciones de la señal, como por ejemplo los modelos de perfusión (heurís-tico, farmacocinético), difusión (monoexponencial, multiexponencial), ratios de relajación (valores de T1 y T2 en RM) y el análisis espectral de absor-ción en TC. Todos estos atributos y parámetros, antes de correlacionarse como biomarcadores subrogados, deben

reducirse para evitar redundancias e inestabi-lidades que impidan un análisis correcto y preciso de estas relaciones.

Para establecer la correlación de los datos compu-tacionales con estos valores (diagnóstico de feno-tipo, predictivo de tratamiento o pronóstico de re-sultado clínico) pueden emplearse **descrip-tores estadísticos** clásicos, análisis multivariante y agrupaciones nosológicas. También pueden re-lacionarse mediante modelos de agrupamiento no supervisado empleando Inteligencia Artificial, con métodos basados en agrupamientos (clusterición) jerárquicos. Estos métodos pueden dar el resulta-do para todo el tumor o para sus hábitats o com-ponentes principales. Estos hábitats expresan re-giones de comportamiento similar obtenidos como firmas nosológicas multiparamétricas, bajo la pre-misa de que para cada evento biopatológico hay una firma nosológica que puede extraerse.

2.1.1.4 Validación de un biomarcador

Cuando ya se dispone del concepto, las imágenes adecuadas, el análisis y los modelados pertinentes y la metodología de medición, es necesario realizar una experiencia piloto que valide el proceso en un grupo de sujetos, con y sin enfermedad, con un tamaño muestral pequeño. La prueba de principio, o **prueba piloto**, se centra en demostrar la validez de las pruebas de concepto y de mecanismo, ambas teóricas, en un grupo inicial reducido, antes de emprender un desarrollo clínico completo. En ella es muy importante además definir si existen variaciones en las medidas asociadas a la edad, al sexo u otros factores que pueden ser de confusión.

Tras el éxito en la prueba piloto se debe trabajar ya con series de un tamaño muestral adecuado y bien definido para extraer conclusiones de peso estadístico. En ellas se define la **prueba de eficacia** como aquella que permite valorar la capacidad de un biomarcador de medir adecuadamente (de modo reproducible, fiable y preciso) el concepto. La eficacia es un estudio demostrativo que define la potencia de la tecnología sanitaria en condiciones de control ideales. La **prueba de efectividad** analiza, en condiciones habituales de uso, la bondad de un biomarcador para medir la variable clínica relevante a la enfermedad (endpoint). Es importante recordar que los objetivos sustitutos o subrogados, es decir, los biomarcadores de imagen pueden reemplazar al objetivo clínico final solo si se han demostrado eficaces y efectivos.

Finalmente, para poder considerar los datos de imagen computacional extraídos como biomarcadores, debe completarse el proceso de validación a nivel técnico y clínico.

Por un lado, está la **validación técnica** que ana-liza la variabilidad en su proceso de obtención. Para ello se debe controlar la posible redundan-cia de datos muy correlacionados y eliminar ade-más los que no sean reproducibles por inestabi-lidad en la extracción del atributo. Se debe con-trolar también su precisión y estabilidad con el coeficiente de correlación intraclase, aplicado a la repetición de experimentos sin cambios (repe-titividad para los mismos equipos con estudios tipo test-retest) y a la repetición con cambios (reproducibilidad en otros equipos con datos in-dependientes). Otro aspecto fundamental en esta validación técnica es el estudio de la exactitud de las medidas mediante el análisis de su asociación con la verdad, considerada como la realidad de los dispositivos tipo fantomas o

de la referencia patológica. Esta validación en la exactitud per-mite además establecer calibraciones para mejo-rar la relación de un método concreto de obtener el biomarcador con la verdad.

Por otro lado, la **validación clínica** analiza la relevancia de la asociación en el entorno asistencial, es decir, su eficiencia práctica. Así, la fiabilidad se establece con la relación existente entre el biomarcador y los resultados clínicos definidos como objetivos en la prueba de concepto, y la de impacto de-termina si se producen cambios en la práctica clínica asociados a la presencia de estas nuevas aproximaciones.

2.1.1.5 La radiómica como parte del flujo del desarrollo de biomarcadores

La finalidad de la Radiómica es extraer características clínicamente relevantes presentes en imágenes radiológicas. Para ello, se apoya en datos que a menudo están "escondidos" a la percepción del ojo humano. Las conclusiones obtenidas a veces no reflejan un significado biológico real, lo que requiere de una validación externa.

Los últimos avances en este campo están facilitando la validación biológica de las características radiómicas mediante la correlación con datos genéticos e histológicos. Se espera que en el futuro se estandarice la correlación biológica de estas características, lo que aumentará la confianza en el método radiómico.

Hay muchos estudios que muestran la relación entre las características radiómicas y biológicas, correlacionando las características encontradas con el grado histológico del tumor o con la expresión génica. Sin embargo, hay que destacar aquellos estudios que buscan correlaciones independientes para validar la naturaleza biológica de las características radiómicas obtenidas. El primer grupo de estudios, están basados en la validación a través del **análisis semántico**, donde se evalúan características semánticas de la imagen, como la irregularidad de los bordes y la necrosis central, y se buscan correlaciones con la naturaleza de la lesión. La segunda forma es la validación a través de la **Radiogenómica**, donde se establece una relación entre las características radiómicas de la lesión y sus datos genéticos, sin embargo, las relaciones pueden ser ambiguas debido a la complejidad del código genético y su vínculo con el fenotipo. Otro grupo versa en la validación a través de la **Inmunohistoquímica** (IHQ) es otra herramienta que describe los fenotipos de las células tumorales y su microambiente. La correlación anatomopatológica es otro enfoque que relaciona los resultados radiómicos del tumor con los hallazgos patológicos. Además la **Patómica** es un campo en expansión que se enfoca en estudiar los patrones en los datos patológicos y trata de mejorar la predicción de resultados. Finalmente, el ***Habitat imaging*** es un enfoque diferente que trata de identificar distintas áreas tumorales y subpoblaciones celulares [13].

Hasta la fecha, muchos estudios del ámbito de la radiómica no realizan la validación de sus resultados más allá de compararlos con un grupo control. Esta tendencia contribuye a una menor reproducibilidad de los resultados y, por ende, menor impacto de los mismos en la práctica clínica. Cuando se atribuya una mayor importancia a la comprensión biológica de las características radiómicas, se podrá desarrollar un sistema de validación estandarizado. No cabe duda de que ya existen muchas formas diferentes de validar y/o aportar contexto

biológico a las firmas radiómicas. Para seguir avanzando, proponemos que los estudios de radiómica intenten siempre presentar un análisis de validación o una hipótesis para la explicación biológica de los resultados. Esto permitirá la discusión de las características biológicas que hay detrás de los resultados, reforzando esta materia.

Una recomendación versa en la utilización del RQS (Radiomic Quality Score) [14] para evaluar la calidad de los datos de radiómica. El RQS es un puntaje numérico que refleja la calidad de los datos obtenidos a través de la radiómica y se utiliza para determinar si es adecuado para un análisis específico. Este índice se basa en diversos criterios, como la calidad de la imagen, la resolución, la consistencia y la representatividad de los datos. Un índice de puntuación más alto indica una calidad superior de los datos, mientras que un puntaje bajo puede indicar la necesidad de mejorar o adquirir nuevos datos.

3. Estándares para marcadores

3.1 DICOM (Dicom Structured Reporting, DICOM SR)

Digital Imaging and Communications in Medicine (DICOM) es el estándar para la comunicación y gestión de la información de imágenes médicas y los datos relacionados. DICOM se utiliza sobre todo para almacenar y transmitir imágenes médicas, lo que permite la integración de dispositivos de imagen médica como escáneres, servidores, estaciones de trabajo, impresoras, hardware de red y sistemas de comunicación y archivo de imágenes (PACS) de múltiples fabricantes. Ha sido ampliamente adoptado por los hospitales y se está abriendo paso en aplicaciones más pequeñas, como las consultas de dentistas y médicos.

Los archivos DICOM pueden intercambiarse entre dos entidades capaces de recibir datos de imágenes y pacientes en formato DICOM. Los distintos dispositivos vienen acompañados de declaraciones de conformidad DICOM que indican las clases DICOM que admiten. La norma incluye una definición de formato de archivo y un protocolo de comunicaciones de red que utiliza TCP/IP para la comunicación entre sistemas.

La Asociación Nacional de Fabricantes Eléctricos (NEMA) posee los derechos de autor de la norma publicada, que fue desarrollada por el Comité de Normas DICOM, cuyos miembros son también en parte miembros de la NEMA. También se conoce como norma NEMA PS3, y como norma ISO 12052:2017 "Health informatics - Digital imaging and communication in medicine (DICOM) including workflow and data management".

DICOM (Digital Imaging and Communications in Medicine) no es solo un formato de archivo para imágenes médicas, es un estándar que busca garantizar todas las necesidades de un sistema PACs (Pictures Archiving and Communication System) como almacenamiento, transmisión, comunicación e impresión [15]. DICOM está universalmente aceptado y proporciona los medios para apoyar intercambio no solo de imágenes adquiridas, sino también de datos clínicos y diversos tipos de resultados con el objetivo de permitir su intercambio y reutilización. Actualmente se encuentra en su versión 3.0, aunque la versión más ampliamente extendida es la versión 2.0.

De manera general, los ficheros DICOM están formados por dos partes diferenciadas:

- Una Cabecera con multitud de campos estandarizados que especifican tanto datos administrativos (datos de paciente o datos del hospital donde se realizó la adquisición) como datos sobre la imagen.
- Cuerpo de la Imagen, que puede estar comprimido con distintos estándares.

En un contexto más técnico, el formato DICOM cuenta con objetos IOD (Information Object Definition), formados por la imagen y su información asociada (son una representación lógica de objetos del mundo real) y DIMSE (DICOM Message Service Element), operaciones que pueden realizarse sobre un objeto. IOD y DICOM forman SOP (Service Object), la unidad funcional de DICOM.

Un IOD se compone de IEs (Entidades de información) (Hay IE de paciente, de estudio, de serie, de equipo, de imagen…) que a su vez se componen de uno o varios módulos, los cuales contienen varios atributos. Un atributo se define con nombre, etiqueta, tipo y descripción (Figura 3) [16].

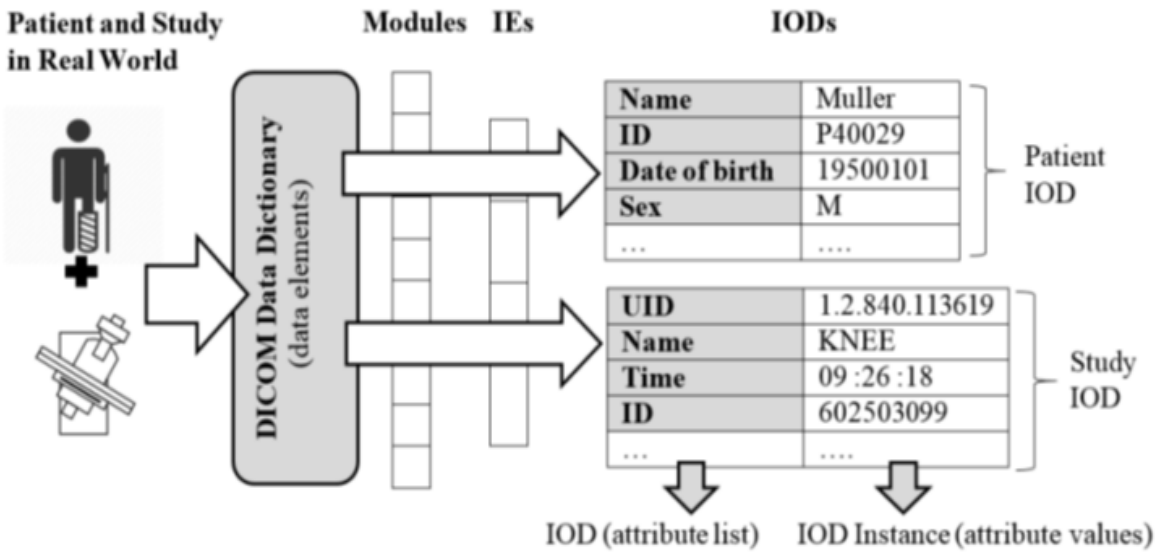


Figura 3. Transformación a un objeto IOD. Fuente: https://theses.hal.science/tel-01961254/document

A la hora de trabajar con imágenes y procesarlas para realizar cualquier procedimiento posterior, se añaden capas de información que nos permiten segmentar ciertas estructuras (órganos o lesiones), etiquetarlas e incluso añadir biomarcadores o medidas. El estándar DICOM proporciona una gran variedad de objetos que se pueden usar para almacenar la información derivada del procesamiento de las imágenes. A continuación, se explican de forma resumida algunos de estos objetos, entre los que destaca DICOM SR, aunque se puede consultar más información en el *entregable E4.4. Normas Internacionales de Anotación de Información de Imagen Médica.*

3.1.1 DICOM SR

Para innovar clínicamente hay que comunicar los resultados de los biomarcadores utilizados de forma intuitiva. Un sistema adecuado es el informe estructurado DICOM (Dicom Structured Reporting, DICOM SR), que es un documento que contiene texto con enlaces a otros datos, como imágenes y coordenadas espaciales o temporales. Esta arquitectura está diseñada para la codificación de documentos y el intercambio de información pertinente mediante el uso de una estructura jerarquizada. DICOM SR introduce las definiciones de objetos de información DICOM y los servicios utilizados para el almacenamiento y transmisión de estos informes estructurados [17].

DICOM SR no solo define las estructuras de datos (paciente, episodio, imágenes, biomarcadores e informes parciales), sino que también permite su almacenamiento, consulta, recuperación, análisis estadístico y transferencia. Además, DICOM SR tiene la capacidad de vincular el documento clínico de los biomarcadores al episodio del paciente en el PACS. El informe estructurado debe contener una información completa y precisa con evaluación de los posibles sesgos y una generalización de los resultados. La norma para la presentación de informes de diagnóstico (tipo STARD -Standards for Reporting of Diagnostic Accuracy-) contiene una lista de los principales elementos para elaborar estudios de eficacia diagnóstica y el diagrama de flujo necesario para garantizar la presencia de toda la información pertinente [18].

Entre las directrices del documento, cabe resaltar las siguientes:

- Describir el proceso de recogida de datos.
- Describir la población o técnica utilizada como referencia.
- Describir la metodología estadística y proporcionar información sobre la reproducibilidad del método.
- Incluir una distribución de la gravedad de la enfermedad y la situación del paciente respecto al grupo de referencia.
- Describir la conclusión del análisis basada en la aplicabilidad clínica final para ayudar a resolver los objetivos clínicos relevantes.

3.1.2 DICOM SEG

DICOM SEG (DICOM Segmentation). Además de poder almacenar las segmentaciones en los objetos DICOM SR, también se pueden almacenar en un objeto DICOM SEG, que define la información gráfica ya sea como objetos rasterizados o de malla de superficie. Representan una clasificación de píxeles que puede ser binaria o fraccionada.

3.1.3 DICOM RWVM

DICOM RWVM (DICOM Real World Value Mapping). Este objeto proporciona un mecanismo para describir el cálculo que se utilizó (y puede reutilizarse) para crear "valores del mundo real" (como el SUV) a partir de los valores de datos de píxeles almacenados. Un RWVM puede estar incrustado dentro de otro objeto DICOM (como una imagen adquirida o derivada)

o puede ser codificado como un objeto independiente, que a su vez puede ser referenciado desde otros objetos, como los SR, o reconocido como relevante a partir de lo común de los identificadores del paciente y del estudio. Estos pueden incluir parámetros de difusión, valores normalizados de algunas modalidades, volumen sanguíneo cerebral regional (RCBV) o flujo (RCBF), etc.

3.1.4 DICOM RT

DICOM RT (DICOM Radiation Therapy) fue creado para la integración de información de radioterapia para un registro electrónico de pacientes, facilitando la interoperabilidad de diferentes sistemas de radioterapia, haciendo posible el intercambio de información de diferentes sistemas. Estos objetos incluyen información sobre el plan de radioterapia, el registro del tratamiento, dosis, etc.

3.2 CIE

La Clasificación Internacional de Enfermedades (CIE) es una herramienta de diagnóstico utilizada en todo el mundo para la epidemiología, la gestión sanitaria y los fines clínicos. La CIE es mantenida por la Organización Mundial de la Salud (OMS), que es la autoridad que dirige y coordina la salud dentro del sistema de las Naciones Unidas. La CIE se diseñó originalmente como un sistema de clasificación de la atención sanitaria, proporcionando un sistema de códigos de diagnóstico para clasificar las enfermedades, incluyendo clasificaciones matizadas de una amplia variedad de signos, síntomas, hallazgos anormales, quejas, circunstancias sociales y causas externas de lesión o enfermedad. Este sistema está diseñado para asignar las condiciones de salud a las categorías genéricas correspondientes junto con las variaciones específicas, asignando para estas un código designado de hasta seis caracteres. Así, las categorías principales están diseñadas para incluir un conjunto de enfermedades similares.

La CIE es publicada por la OMS y se utiliza en todo el mundo para las estadísticas de morbilidad y mortalidad, los sistemas de reembolso y el apoyo a la toma de decisiones automatizadas en la atención sanitaria. Este sistema está diseñado para promover la comparabilidad internacional en la recogida, procesamiento, clasificación y presentación de estas estadísticas. Al igual que el análogo Manual Diagnóstico y Estadístico de los Trastornos Mentales (que se limita a los trastornos psiquiátricos y es casi exclusivo de Estados Unidos), la CIE es un proyecto de gran envergadura para clasificar estadísticamente todos los trastornos de la salud y proporcionar ayuda al diagnóstico. La CIE es un sistema básico de diagnóstico clasificatorio basado en estadísticas para cuestiones relacionadas con la atención sanitaria de la Familia de Clasificaciones Internacionales de la OMS (WHO-FIC).

La CIE se revisa periódicamente y actualmente se encuentra en su 11ª revisión. La CIE-11, como se conoce, por tanto, fue aceptada por la Asamblea Mundial de la Salud (AMS) de la OMS el 25 de mayo de 2019 y entró oficialmente en vigor el 1 de enero de 2022. Sin embargo, es probable que la adopción generalizada tarde algún tiempo: Estados Unidos no adoptó la

CIE-10 hasta 2015, y en 2018 algunos países seguían utilizando la CIE-9 e incluso la CIE-8. La versión para preparar su aprobación en la AMS se publicó el 18 de junio de 2018.

La CIE forma parte de una "familia" de clasificaciones internacionales (WHOFIC) que se complementan entre sí, incluyendo también la Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud (CIF) que se centra en los dominios de funcionamiento (discapacidad) asociados a las condiciones de salud, tanto desde el punto de vista médico como social, y la Clasificación Internacional de Intervenciones Sanitarias (CIIS) que clasifica toda la gama de intervenciones médicas, de enfermería, de funcionamiento y de salud pública.

El título de la CIE es formalmente la Clasificación Estadística Internacional de Enfermedades y Problemas de Salud Relacionados, aunque el título original, Clasificación Internacional de Enfermedades, sigue siendo informalmente el nombre por el que se suele conocer.

3.2.1 CIE-10

Es la clasificación actual. La CIE-10 es la décima revisión de la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (CIE), una lista de clasificación médica de la Organización Mundial de la Salud (OMS). Contiene códigos para enfermedades, signos y síntomas, hallazgos anormales, quejas, circunstancias sociales y causas externas de lesiones o enfermedades. La CIE-10 comenzó a elaborarse en 1983, fue aprobada por la 43ª Asamblea Mundial de la Salud en 1990 y fue utilizada por primera vez por los Estados miembros en 1994.

Aunque la OMS gestiona y publica la versión básica de la CIE, varios Estados miembros la han modificado para adaptarla mejor a sus necesidades. En la clasificación base, el conjunto de códigos permite más de 14.000 códigos diferentes y permite el seguimiento de muchos diagnósticos nuevos en comparación con la CIE-9 precedente. Mediante el uso de subclasificaciones opcionales, la CIE-10 permite la especificidad en cuanto a la causa, la manifestación, la localización, la gravedad y el tipo de lesión o enfermedad. Las versiones adaptadas pueden diferir en varios aspectos, y algunas ediciones nacionales han ampliado aún más el conjunto de códigos; algunas llegan a añadir códigos de procedimiento. La CIE-10-CM, por ejemplo, tiene más de 70.000 códigos.

3.2.2 CIE-11

Es una nueva clasificación en desarrollo. La CIE-11 es la undécima revisión de la Clasificación Internacional de Enfermedades (CIE). Sustituye a la CIE-10 como norma mundial para registrar la información sanitaria y las causas de muerte. La Organización Mundial de la Salud (OMS) desarrolla y actualiza anualmente la CIE. El desarrollo de la CIE-11 comenzó en 2007 y se prolongó durante más de una década de trabajo, con la participación de más de 300 especialistas de 55 países divididos en 30 grupos de trabajo, y otras 10.000 propuestas de personas de todo el mundo. Tras una versión alfa en mayo de 2011 y un borrador beta en mayo de 2012, el 18 de junio de 2018 se publicó una versión estable de la CIE-11, que fue

aprobada oficialmente por todos los miembros de la OMS durante la 72ª Asamblea Mundial de la Salud, el 25 de mayo de 2019.

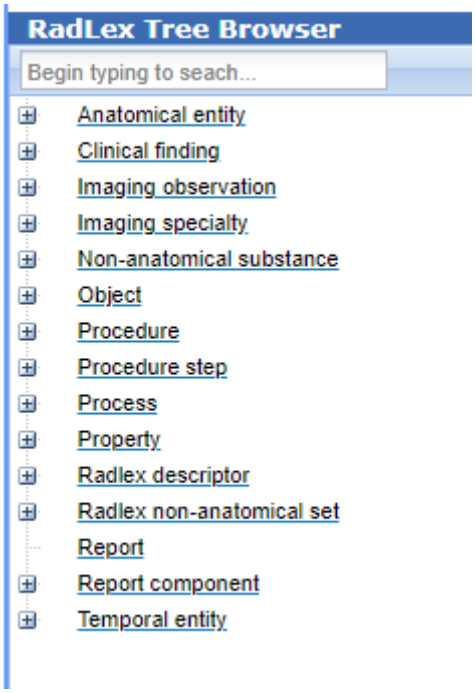
3.2.3 CIE-O

La Clasificación Internacional de Enfermedades para la Oncología (CIE-O) es una extensión específica de la Clasificación Estadística Internacional de Enfermedades y Problemas de Salud Relacionados para las enfermedades tumorales. Esta clasificación es ampliamente utilizada por los registros de cáncer.

Actualmente se encuentra en su tercera revisión (CIE-O-3). La CIE-10 incluye una lista de códigos morfológicos. Estos proceden de la segunda edición de la CIE-O (CIE-O-2) que era válida en el momento de su publicación.

3.3 RADLEX

En los últimos años en el campo de la radiología ha surgido la necesidad de crear un vocabulario común estándar de términos radiológicos específicos, para así poder compartir los datos entre diferentes entidades y además poder crear reportes radiológicos normalizados [19]. Para cumplir esta función la Radiological Society of North America (RSNA) ha creado RadLex (Radiology Lexicon), un conjunto de términos estandarizados que representan los diferentes conceptos utilizados en el campo de la radiología, agrupados en diferentes campos semánticos. Estos se pueden explorar en su página web[1].



1 https://radlex.org

Figura 4Grupos de términos de Radlex

Radlex ofrece varias ventajas, entre ellas:

- Unificación de términos: Proporciona un vocabulario común para nombrar y describir los elementos del sistema radiológico, permitiendo una mayor precisión y claridad en la comunicación.
- Facilita la búsqueda y recuperación de información mediante el uso de términos específicos y relaciones semánticas estructuradas.
- Permite la interoperabilidad entre diferentes sistemas de información radiológica, facilitando la integración de datos y el desarrollo de aplicaciones médicas.

En lo que hace referencia a las ontologías, una ontología es un modelo conceptual que representa un conjunto de conceptos y relaciones semánticas en un dominio específico. Esto permite una mejor comprensión y organización del conocimiento en ese dominio. La ontología de RadLex es una herramienta importante para estandarizar y organizar los términos y conceptos utilizados en el campo de la radioterapia. Esto permite a los profesionales de la salud comprender y comunicarse de manera consistente sobre el tema, lo que facilita la colaboración y el intercambio de información.

Además, la ontología de RadLex puede vincularse con sistemas comunes de datos, como OMOP (Observational Medical Outcomes Partnership)[2], para mejorar la recopilación, análisis y utilización de datos de salud en investigación clínica y epidemiológica. Esto puede ayudar a identificar patrones y tendencias en el uso y efectividad de la radioterapia, lo que puede mejorar la atención médica y la toma de decisiones en el campo.

En un futuro se espera que Radlex contribuya a una mejor identificación de enfermedades y tratamientos más efectivos. También se espera que facilite la integración de sistemas médicos y la recopilación de datos en bases de datos médicas.

Sin embargo, todavía queda mucho trabajo por hacer en relación a Radlex, ya que aún hay que introducir la definición de los términos de RadLex en lengua española, para así promover su uso y estandarización entre los profesionales de la salud para garantizar su utilidad y efectividad en la práctica médica.

3.4 UMLS

UMLS o Unified Medical Language System es un conjunto de archivos y herramientas software que recopila múltiples vocabularios y estándares biomédicos y de la salud para posibilitar la interoperabilidad entre sistemas. El objetivo de UMLS es identificar los conceptos biomédicos mediante un código único (CUI) y relacionarlos en forma de red.

La base de UMLS es el Metatesauro UMLS, una herramienta que incluye códigos y términos de varios vocabularios, entre ellos CPT®, ICD-10-CM, LOINC®, MeSH®, RxNorm y SNOMED CT®. El Metatesauro relaciona los términos y códigos de los más de 200 distintos

2 https://www.ohdsi.org/data-standardization/

vocabularios fuente que incluye adaptándolos a un formato común de datos. Estos términos y códigos se agrupan en conceptos, los cuales se categorizan según distintos tipos semánticos. Además, en este formato común de datos se incorporan relaciones y atributos que proporcionan los vocabularios.

Este Metatesauro se crea a partir de otras dos herramientas proporcionadas por UMLS: la red semántica (Semantic Network) y el léxico especializado y herramientas léxicas (SPECIALIST Lexicon and Lexical Tools). La red semántica proporciona las categorías semánticas y sus relaciones, mientras que el léxico especializado y herramientas léxicas son un conjunto de herramientas de lenguaje natural.

3.5 Indexación con distintas ontologías

Además del CIE, existen otras ontologías que se pueden utilizar para describir y clasificar a los pacientes en términos de enfermedad, cada una con sus propias ventajas. Estas incluyen ontologías destinadas a describir fenotipos patológicos, que pueden considerarse como componentes de la enfermedad, y ontologías destinadas a describir enfermedades en sí mismas, aunque hay mucha superposición entre las dos. Por ejemplo, algunos fenotipos como la anemia hipocrómica pueden considerarse tanto un fenotipo componente de una enfermedad como una enfermedad en sí misma. Estas ontologías se pueden utilizar para ampliar y complementar los códigos CIE y, a menudo, se utilizan como entrada para metodologías externas para el diagnóstico genético. Como ejemplos de ontologías importantes tenemos “The Human Phenotype Ontology” (HPO) [20] y Mondo [21].

3.5.1 La ontología del fenotipo humano (HPO)

HPO [20] proporciona un conjunto organizado jerárquicamente de definiciones estandarizadas para anomalías fenotípicas relacionadas con enfermedades. Está dirigido a enfermedades hereditarias, por lo que su aplicabilidad a la imagen médica depende del contexto. Sin embargo, en casos apropiados, por ejemplo, deformidades congénitas y otras anomalías físicas, pueden ser muy poderosas, ya que permiten una descripción precisa del problema. Además, tienen la ventaja añadida de ser la entrada a una gran gama de software basado en fenotipos para diagnóstico genómico y diferencial. Por ejemplo, Phenomiser se puede usar para sugerir enfermedades para un paciente con un conjunto determinado de términos HPO asignados [22]; Exomizer se puede utilizar para priorizar las variantes que causan enfermedades en pacientes para los que están disponibles tanto la anotación del término HPO como los datos de variantes genómicas [23].

3.5.2 Mondo

Mondo [21] es una ontología que busca unificar las definiciones de enfermedades en todos los recursos de anotación de enfermedades. Se puede considerar como un complemento de UMLS de muchas maneras, con la particularidad de que proporciona un mapeo estricto, uno a uno, entre términos cuando se mapea a otros recursos de anotación, mientras que UMLS es menos restrictivo. Mondo se desarrolla en estrecha colaboración con HPO, y presenta

referencias cruzadas tanto con términos HPO como CIE, además de otros muchos otros recursos de anotación.

3.5.3 Mapeo entre ontologías

El mapeo entre los códigos CIE y los términos HPO se puede lograr de varias maneras. Si bien HPO OBO incluye campos de referencia cruzada que se pueden usar para mapear directamente entre los términos HPO y los códigos CIE, estos son limitados en número. Para evitar este problema, Metathesaurus, una red semántica incluida en UMLS (descrito anteriormente), podría usarse para vincular conceptos similares entre diferentes ontologías a través de UMLS. También se puede usar potencialmente Mondo, mapeando CIE a entidades Mondo y luego a través de estos a términos HPO. Además, hay estudios recientes que utilizan diferentes técnicas para mapear entre términos HPO y códigos CIE. Estos incluyen enfoques basados en minería de texto [24], el uso de códigos CIE obtenidos a través de UMLS de OMIM y otros recursos de anotación de enfermedades, y la búsqueda de menciones conjuntas con términos HPO dentro de los resúmenes de Pubmed [24]. Otro enfoque es a través de "phecodes", obtenidos agregando códigos CIE similares en rasgos, y usándolos para mapear entre HPO e CIE [25].

4. Grupos internacionales de trabajo

Diferentes sociedades científicas radiológicas internacionales han puesto en marcha iniciativas con el fin de agrupar, analizar y compartir estrategias, definiciones y listados de biomarcadores de imagen médica. De forma coordinada, las sociedades que mayores progresos han realizado son la europea y norteamericana de radiología. Desde hace varios años, ambas sociedades crearon grupos de trabajo que posteriormente confluyeron en alianzas para el estudio e implementación de los biomarcadores de imagen. Cabe señalar por otro lado que la sociedad japonesa de radiología (JRS) puso también en marcha en 2015 una iniciativa similar a la norteamericana, que tituló J-QIBA[3].

4.1.1 European Imaging Biomarkers Alliance (EIBALL)

A nivel de la Sociedad Europea de Radiología (ESR), el subcomité de biomarcadores de imágenes de ESR, el Grupo de trabajo de ESR sobre medicina personalizada y el Grupo de trabajo de ESR-EORTC se fusionaron en una nueva iniciativa llamada EIBALL.

Esta alianza tiene como principales objetivos:

- Establecer qué biomarcadores funcionales son funcionales actualmente. En este sentido, en su web existe un inventario de los biomarcadores, y cuáles son las

[3] http://www.radiology.jp/j-qiba/english/index.html

medidas en cada uno[4]. Este listado se agrupa según órganos diana y a su vez, en función de las técnicas de imagen.

- Permitir el uso clínico de biomarcadores mediante el establecimiento de estándares para la adquisición de datos e imagen. Esta iniciativa se lleva a cabo conjuntamente con QIBA.
- Finalmente, divulgar y promocionar la educación sobre el uso apropiado uso de imagen funcional biomarcadores así como su interpretación.

4.1.2 Quantitative Imaging Biomarkers Alliance (QIBA)

Por otro lado, la Radiological Society of North America (RSNA) en 2007 puso en marcha una alianza similar a la de la sociedad europea. Esta iniciativa tiene como objetivos específicos:

- Colaboración entre centros e instituciones para identificar necesidades, barreras y soluciones para crear resultados de imágenes cuantitativos consistentes, confiables, válidos y alcanzables de forma general en todos los centros que usen imagen médica.
- Acelerar el desarrollo y la adopción de estándares de hardware y software para lograr resultados cuantitativos precisos y reproducibles a partir de métodos de imágenes.

En este sentido, QIBA se ha centrado especialmente en la metodología y definición de los conceptos, así como en la métrica en biomarcadores. De forma similar a EIBALL, disponen de un repositorio de varios biomarcadores que clasifican según el grado de evidencia así como de implementación en la práctica diaria[5].

5. Biobancos de imágenes y biomarcadores de imágenes basados en imagen poblacional

5.1 Introducción

Esta especialidad pretende establecer estrategias y herramientas destinadas a transformar datos e información en conocimiento y ciencia por lo que jugará un papel esencial en el desarrollo de Data-Lakes poblacionales, basados en Imagen Médica. Las imágenes médicas son, con mucho, la tecnología más avanzada y de mayor crecimiento de datos en la investigación biológica y biomédica moderna en el ámbito de IA. El término "Imagen médica Poblacional" cubre una amplia gama de tecnologías y aplicaciones tanto clínicas, epidemiológicas, así como científicas. Su fundamento está basado en la cuantificación de los valores de píxel registrados por un sistema de imágenes o modalidad. Estas mediciones cuantitativas se corresponden con la constitución, estructura o dinámica de la biología u

[4] https://www.myesr.org/research/biomarkers-inventory
[5] https://qibawiki.rsna.org/index.php/Profiles

organismo que se está estudiando, y serán el inicio de la exploración poblacional en diversas patologías como por ejemplo el estudio del Alzheimer. En cuanto a los Data-Lakes, estos se conciben como repositorios que se utilizan para almacenar cantidades masivas de datos (información estructurada y no estructurada), generalmente para análisis futuros, revisiones metodológicas, metaanálisis, procesamiento de big data y aprendizaje automático. Un Data-Lake permite aplicar múltiples enfoques con los datos, sin embargo, hay que aplicar implementaciones metodológicas con pautas bien definidas (por ejemplo, Medical Imaging Data Structure MIDS) y que cumpla con los objetivos FAIR. Las palabras clave (o grupos de palabras y sus abreviaturas) que sirven para caracterizarla a los efectos de considerar específica la bibliografía, experiencia o los cursos son las siguientes (apareciendo dos o más): Imagen Médica Poblacional (IMP), Sistemas de Gestión de Imagen Médica Digital, Gestión de Contenidos de Salud, Data-Lakes, Data warehouse, Administración de datos, Almacenes de datos, Big data analytics, Datos estructurados y no estructurados, Salud electrónica, E-Health, Healthcare Content Management, Anonimización imagen clínica, Data curation, Medical Population Imaging Data Structure (MIDS-BIDS), DICOM, HIPPA, Digital Anomymation y anonimización.

5.2 Solicitud y acceso a los biomarcadores

El proceso de solicitud de datos de biomarcadores de imagen biomédica poblacional se llevaría a cabo en varios pasos en el marco del protocolo de gobernanza. Primero, el solicitante debería completar una solicitud de acceso en la que se describa el propósito del estudio, se detalle cómo se garantizará el cumplimiento de las normas de privacidad y confidencialidad de los datos, y se incluya cualquier información adicional relevante. En segundo lugar, un comité de revisión de acceso revisaría la solicitud y determinaría si se cumplen los criterios para el acceso a los biomarcadores de imagen biomédica. Estos criterios pueden incluir, entre otros, el valor científico del estudio, la capacidad del solicitante para manejar adecuadamente los datos y garantizar su confidencialidad, y la calidad de la propuesta en general.

Si la solicitud es aprobada, se establecerían las condiciones para el acceso a los datos, incluyendo asesoramiento que el solicitante deba completar antes de acceder a los datos. Además, se establecería un plazo para la realización del estudio y se definirían las medidas de seguimiento para garantizar el cumplimiento de las condiciones establecidas. Finalmente, se proporcionaría acceso a los biomarcadores de imagen biomédica a través de un mecanismo seguro, como un portal web o un servidor seguro, y se llevaría un registro de los datos accedidos. Este proceso de solicitud de datos seguiría el protocolo de gobernanza de otros repositorios de datos clínicos existentes.

5.3 Casos de éxito

A continuación, se proporciona una breve descripción de algunos repositorios de imagen médica disponibles de manera abierta:

- **The Cancer Imaging Archive (TCIA)**: es un repositorio de imagen médica que contiene conjuntos de datos de imágenes relacionadas con el cáncer, incluyendo tomografías computarizadas, resonancias magnéticas y radiografías.
- **National Institutes of Health (NIH) 3D Print Exchange**: es un repositorio que permite a los usuarios descargar y compartir archivos de impresión 3D de estructuras anatómicas relevantes para la investigación biomédica y la educación.
- **The National Library of Medicine (NLM) Visible Human Project**: es un proyecto que proporciona imágenes digitales de secciones transversales de un cuerpo humano sano, tanto en hombre como en mujer, con una resolución de 1 mm.
- **Open-i**: es un repositorio que proporciona acceso a imágenes médicas de artículos científicos publicados en la Biblioteca Nacional de Medicina de los Estados Unidos (NLM).
- **Radiopaedia.org**: es una base de datos de imágenes médicas que incluye radiografías, tomografías computarizadas, resonancias magnéticas y otros tipos de imágenes médicas, junto con descripciones y análisis por parte de radiólogos y otros profesionales médicos.
- **Banco de Imagen Médica de la Comunidad Valenciana (BIMCV)**[6]: Se trata de un recurso concebido como un repositorio de conocimiento orientado a favorecer los avances tecnológicos en imagen médica y proporcionar servicios de cobertura tecnológica para el soporte a proyectos I+D. Actualmente, BIMCV tiene ámbito geográfico e incluye todos los centros públicos de la Comunidad Valenciana.
- Base de datos NIH de 100,000 imágenes de radiografías de tórax[7]: Datos clínicos asociados, anotaciones y diagnósticos.
- **The Cancer Imaging Archive (TCIA)**[8]
- **Base de datos de imágenes ultrasónicas rastreadas de alta calidad de fantasmas**[9]: "una base de datos de secuencias rastreadas de imágenes de EE. UU. de fantasmas médicos adquiridas con una metodología que garantiza el control espacial y de fuerza de la sonda de EE. brazo y un sistema de seguimiento óptico". (Fantasmas abdominales y de bebés)
- **Base de datos MedPix de NLM**[10]: Una base de datos de imágenes médicas en línea gratuita con más de 59 000 imágenes indexadas y seleccionadas, de más de 12 000 pacientes
- **GrepMed**[11]: "Encuentre algoritmos, ayudas para la toma de decisiones, listas de verificación, pautas, diferenciales, ecografía en el punto de atención (POCUS), clips de exámenes físicos y más"

[6] https://bimcv.cipf.es/bimcv-projects/
[7] https://nihcc.app.box.com/v/ChestXray-NIHCC
[8] https://www.cancerimagingarchive.net/
[9] https://www.vicomtech.org/demos/us_tracked_dataset/UsTrackedDataset.htm
[10] https://medpix.nlm.nih.gov/home
[11] https://medpix.nlm.nih.gov/home

- **OASIS**[12]: Datos de IRM transversales en adultos mayores jóvenes, de mediana edad, sin demencia y con demencia. Datos de IRM longitudinales en adultos mayores con demencia y sin demencia
- **ADNI**[13]: La Iniciativa de Neuroimagen de la Enfermedad de Alzheimer de ADNI (ADNI) une a los investigadores con datos de estudio como trabajan para definir la progresión de la enfermedad de Alzheimer. Los investigadores de ADNI recopilan, validan y utilizan datos como imágenes de MRI y PET, genética, pruebas cognitivas, LCR y biomarcadores sanguíneos como predictores de la enfermedad.
- **FITBIR**[14]: El sistema informático Federal Interagency Traumatic Brain Injury Research (FITBIR): MRI, PET, contraste y otros datos sobre una variedad de condiciones de TBI
- **STARE**[15]: Análisis estructurado de la retina: esta investigación se refiere a un sistema para diagnosticar automáticamente enfermedades del ojo humano.
- **Cancer Digital Slide Archive**[16]: Imágenes de preparaciones histológicas completas de muestras de todos los tipos de neoplasias de The Cancer Genome Atlas (TCGA), con acceso a tos genómicos y de anatomía patológica (https://portal.gdc.cancer.gov/projects) de 86.513 casos pertenecientes a 74 proyectos.
- **Repositorio de imágenes médicas de SICAS**[17]: TC post mortem de 50 sujetos TC, microCT, segmentación y modelos de cóclea Copias de datos de desafío seleccionados (p. ej., BRATS2015)
- **Colección DTI del Instituto Médico Johns Hopkins**[18]: Atlas DTI: adultos, niños, ...
- **Centro de recursos de datos e imágenes médicas (MIDRC)**[19]: Un esfuerzo conjunto entre el Colegio Americano de Radiología (ACR), la Sociedad Radiológica de América del Norte (RSNA), y la Asociación Estadounidense de Físicos en Medicina (AAPM)
- **Duke Center for In Vivo Microscopy**[20]: MRI, CT, ... de animales pequeños
- **Data.Kitware.Com**[21]: Imágenes y conjuntos de datos de una amplia variedad de dominios de computación científica (incluidas las imágenes médicas)
- **100 Resonancias magnéticas cerebrales sanas**: 18-90 años. Colección equilibrada por edad y género MRA, T1, T2 y algunos vasos intracraneales DTI extraídos de pacientes seleccionados.
- **MIDAS @ Kitware.Com**[22]: Recopilación principalmente de datos históricos del desarrollo de 3D Slicer, el grupo NAMIC y NLM.

[12] http://www.oasis-brains.org/
[13] https://adni.loni.usc.edu/
[14] https://fitbir.nih.gov/
[15] http://cecas.clemson.edu/~ahoover/stare/
[16] http://cancer.digitalslidearchive.net/
[17] https://www.smir.ch/
[18] https://cmrm.med.jhmi.edu/
[19] https://www.midrc.org/
[20] http://www.civm.duhs.duke.edu/SharedData/DataSupplements.htm
[21] https://data.kitware.com/
[22] https://www.insight-journal.org/midas/

- **Datos de NAMIC**[23]: Alianza nacional para la computación de imágenes médicas
- **NLM Datos**[24]: diseñados para la segmentación de imágenes médicas y la evaluación de algoritmos de registro
- **Base de datos de IA en medicina de Stanford Combinación**[25]: de radiografías, tomografías computarizadas y resonancias magnéticas de tórax, manos, etc
- **Archivo de imágenes en línea de visión por computadora**: Gran lista de múltiples bases de datos en visión por computadora y biomédica Imaging
- **Cornell Visualization and Image Analysis (VIA) group**[26]: Proporciona una lista de las bases de datos disponibles, muchas de las cuales también se enumeran aquí.
- **OmniMedicalSearch.com**[27]: Bases de datos y bibliotecas de imágenes médicas
- **Base de datos digital para mamografías de detección (DDSM)**[28]: Gran colección con hallazgos normales y anormales y datos reales.
- **Imágenes digitales de la retina para la extracción de vasos (DRIVE)**[29]: Imágenes digitales y segmentaciones expertas de los vasos de la retina.
- **Base de datos de la Sociedad Japonesa de Tecnología Radiológica (JSRT) Imágenes**[30]: digitales de rayos X de tórax con ubicaciones de nódulos pulmonares, datos reales y controles.
- **Base de datos de segmentación en radiografías de tórax (SCR)**[31]: Imágenes digitales de rayos X de tórax con segmentaciones de campos pulmonares, corazón y clavículas.
- **Base de datos pulmonar pública para abordar la respuesta a los medicamentos**[32]: Imágenes de TC de tórax bien documentadas.
- **Base de datos de la Sociedad de Análisis de Imágenes Mamográficas (mini-MIAS)**[33]: Imágenes mamográficas y marcado.
- **Base de datos estándar de retinopatía diabética (DIARETDB1)**[34]**:** Imágenes: digitales de la retina para detectar y cuantificar la retinopatía diabética.
- **SpineWeb**[35]: SpineWeb es una plataforma de colaboración en línea para todos los interesados en la investigación sobre imágenes de la columna vertebral y análisis de imágenes.
- **OAinitiative**[36]: de caderas, rodillas y otros sitios afectados por osteoartritis

[23] https://www.insight-journal.org/midas/community/view/17
[24] https://www.insight-journal.org/midas/community/view/15
[25] https://aimi.stanford.edu/research/public-datasets
[26] https://homepages.inf.ed.ac.uk/rbf/CVonline/Imagedbase.htm#biomed
[27] https://omnimedicalsearch.com/
[28] http://marathon.csee.usf.edu/Mammography/Database.html
[29] https://www.isi.uu.nl/Research/Databases/DRIVE/
[30] https://www.jsrt.or.jp/web_data/english03.php
[31] https://www.isi.uu.nl/Research/Databases/SCR/
[32] http://www.via.cornell.edu/crpf.html
[33] http://peipa.essex.ac.uk/info/mias.html
[34] http://www2.it.lut.fi/project/imageret/diaretdb1/
[35] http://spineweb.digitalimaginggroup.ca/
[36] https://nda.nih.gov/oai/

- **Repositorio pediátrico**[37]: RM Cerebro
- **Facebase**[38]: Mediciones de superficie craneofacial en 3D
- **Embodi3D**[39]: Colección de archivos destinados a la impresión 3D, pero incluye escaneos médicos volumétricos (es decir, CT y MRI en formato NRRD) para una variedad de estructuras anatómicas (huesos, músculos, buques).
- **Open Microscopy**[40]: Colección de preparaciones digitales de patología con imágenes en formato propietario de mútiples fabricantes y ejemplos de imágenes en DICOM WSI.
- **Open Slide**[41]: OpenSlide es una librería en C que proporciona una interfaz sencilla y única para mostrar imágenes de preparaciones completas (WSI). Dispone de un número reducido de ejemplos de ocho fabricantes.
- **DICOM Whole Slide Imaging (WSI)**[42]: Además de incluir una descripción detallada del suplemento 145 de DICOM (WSI), contiene enlaces de interés a repositorios y servidores de DICOM WSI

6. Casos de uso mínimos a abordar

Los biomarcadores de imagen pueden ser utilizados para diversos propósitos y aplicaciones, incluyendo el diagnóstico y seguimiento de enfermedades y alteraciones patológicas, la estratificación de anomalías en fenotipos, la evaluación de los efectos de un tratamiento, la evaluación del impacto de enfermedades crónicas, el monitoreo del tratamiento y la investigación médica. La utilización de biomarcadores de imagen puede proporcionar información complementaria útil al diagnóstico tradicional y puede ser de gran valor en la evaluación y el tratamiento de una amplia variedad de enfermedades y condiciones médicas.

A continuación, se detallan los casos de uso mínimos que la plataforma de imagen biomédica para explotación mediante uso secundario debería abordar:

- Diagnóstico y seguimiento de enfermedades y alteraciones patológicas: Los biomarcadores de imagen pueden utilizarse para establecer la presencia de una alteración o lesión, medir su situación biológica y definir su historia natural y progreso.
- Estratificación de anomalías en fenotipos: Los biomarcadores de imagen también pueden utilizarse para estratificar las anomalías en diferentes fenotipos, lo que puede ayudar a identificar las características específicas de una enfermedad o alteración patológica.
- Evaluación de los efectos de un tratamiento: Los biomarcadores de imagen también pueden utilizarse para evaluar los efectos de un tratamiento, como medicamentos o procedimientos quirúrgicos, en una enfermedad o alteración patológica. Esto puede

[37] https://nda.nih.gov/edit_collection.html?id=1151nihpd/info/index.html
[38] https://www.facebase.org/
[39] https://www.embodi3d.com/democratiz3D/
40 https://downloads.openmicroscopy.org/images/
[41] https://openslide.org/
[42] https://dicom.nema.org/dicom/dicomwsi/

incluir el uso de biomarcadores para evaluar la respuesta a un tratamiento, o para determinar si un tratamiento está siendo efectivo en modificar la historia natural de la enfermedad.

- Evaluación del impacto de enfermedades crónicas: Los biomarcadores de imagen también pueden utilizarse para evaluar el impacto de enfermedades crónicas, como la diabetes o la enfermedad pulmonar obstructiva crónica (EPOC), en diferentes órganos y sistemas del cuerpo. Esto puede incluir el uso de biomarcadores para evaluar la progresión de la enfermedad y para determinar el impacto de la enfermedad en la calidad de vida del paciente.
- Monitoreo del tratamiento: Los biomarcadores de imagen también pueden utilizarse para monitorear el progreso del tratamiento de una enfermedad o condición médica, como el cáncer o la enfermedad cardiovascular. Por ejemplo, pueden utilizarse biomarcadores para evaluar la respuesta a un tratamiento oncológico o para monitorizar la progresión de una enfermedad cardiaca.
- Investigación médica: Los biomarcadores de imagen también pueden ser utilizados en el contexto de la investigación médica, para explorar y comprender mejor las causas de enfermedades y condiciones médicas, así como para desarrollar nuevas terapias y tratamientos. Esto puede incluir el uso de biomarcadores para investigar las alteraciones en la estructura o la función de órganos o tejidos en diferentes enfermedades, o para evaluar el efecto de nuevas terapias en diferentes modelos animales o en humanos.

7. Conclusiones

La imagen médica ha sido fundamental en el avance de la medicina personalizada y se utiliza para detectar, caracterizar y estudiar enfermedades. Sin embargo, a pesar de la gran resolución de las imágenes tomográficas, aún existen lagunas en la información que generan. Por lo tanto, el desarrollo de biomarcadores de imagen médica computacional es crucial para cubrir estas necesidades. Los biomarcadores son parámetros cuantitativos extraídos computacionalmente de imágenes médicas radiológicas. En anatomía patológica, los biomarcadores de imagen son importantes en cuatro aplicaciones específicas: a) cuantificación objetiva de biomarcadores; b) mejora en la precisión diagnóstica; c) soporte a la decisión clínica y d) descubrimiento de nuevas características morfológicas. Para establecer y validar un biomarcador, se deben seguir varios pasos que incluyen adquirir imágenes, conocer los parámetros que se quieren calcular, preparar las imágenes para un marco de referencia común y extraer descriptores de los datos computacionalmente.

Los biomarcadores de imagen médica son parámetros cuantitativos obtenidos computacionalmente de imágenes médicas radiológicas y son importantes en la medicina personalizada para la cuantificación objetiva, mejora de la precisión diagnóstica, soporte a la decisión clínica y descubrimiento de nuevas características morfológicas. La establecimiento y validación de un biomarcador incluye la adquisición de imágenes, conocimiento de los parámetros a calcular, preparación de imágenes con marco de referencia común y extracción computacional de descriptores de los datos.

En el estudio de imágenes paramétricas (convencionales y multivariantes), se pueden obtener datos de todo el tejido o solo de áreas seleccionadas. Los datos se extraen desde dos entornos principales: atributos de la imagen (texturas, distribución espacial de la señal, descripción morfológica) y parámetros obtenidos del análisis de series dinámicas (modelos de perfusión, difusión, ratios de relajación, análisis espectral de absorción). Antes de relacionarlos como biomarcadores, se deben reducir para evitar redundancias e inestabilidades. Estos datos pueden relacionarse con valores clínicos a través de descripciones estadísticas, análisis multivariantes y modelos de agrupamiento no supervisado con Inteligencia Artificial. Para validar un biomarcador, se deben realizar pruebas piloto en un grupo pequeño de pacientes, con y sin enfermedad, para demostrar la validez de las pruebas de concepto y mecanismo, así como su sensibilidad y especificidad.

En lo que a estándares se refiere, DICOM es un estándar para la comunicación y gestión de información de imágenes médicas y datos relacionados. Se utiliza en hospitales y en aplicaciones médicas más pequeñas y permite la integración de dispositivos de imagen médica y la comunicación entre sistemas de diferentes fabricantes. La norma incluye un formato de archivo y un protocolo de comunicación de red basado en TCP/IP. Los archivos DICOM incluyen una cabecera con campos estandarizados y un cuerpo de la imagen que puede estar comprimido. DICOM utiliza objetos IOD y DIMSE para realizar operaciones sobre imágenes y datos. Los IOD están formados por IEs que a su vez están compuestas por módulos y atributos.

El estándar DICOM ofrece varios objetos para almacenar información derivada del procesamiento de imágenes médicas. Uno de estos objetos es el DICOM SR, que permite la codificación y el intercambio de informes estructurados mediante una estructura jerarquizada. DICOM SR permite almacenar, consultar y transferir información clínica de biomarcadores relacionados con el episodio del paciente. El informe estructurado debe contener información completa y precisa con evaluación de sesgos y generalización de resultados. Otro objeto es el DICOM SEG, que permite almacenar segmentaciones de imágenes como objetos rasterizados o de malla de superficie. Por último, el DICOM RWVM permite describir el cálculo utilizado para crear valores del mundo real a partir de los valores de los píxeles.La Clasificación Internacional de Enfermedades (CIE) es un sistema de clasificación de diagnóstico de la OMS utilizado en todo el mundo para la epidemiología, la gestión sanitaria y los fines clínicos. Proporciona un sistema de códigos de diagnóstico y clasifica las enfermedades y lesiones con un código de hasta seis caracteres. La CIE es actualmente en su 11ª revisión (CIE-11) y entró en vigor en 2022. La CIE forma parte de una "familia" de clasificaciones internacionales complementarias, incluyendo la Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud y la Clasificación Internacional de Intervenciones Sanitarias.

El CIE 10 es fue publicado por la Organización Mundial de la Salud en 1990 y actualmente es ampliamente utilizado en el sector de la salud.

El CIE 11 fue publicada en junio de 2018. La revisión incluye una mayor consideración de los trastornos mentales y de la conducta, así como una mayor integración con los avances en la medicina digital y tecnológica.

La CIE-O es la versión de la clasificación de la OMS para su uso en el sector de la salud oral. Esta clasificación se utiliza para describir y clasificar los problemas y enfermedades relacionados con la salud oral, incluyendo la odontología, la estomatología y la maxilofacial.

El UMLS (Unified Medical Language System) es un conjunto de herramientas y archivos que recopila varios vocabularios y estándares en salud y biomedicina para permitir la interoperabilidad entre sistemas. El objetivo de UMLS es identificar los conceptos biomédicos a través de un código único (CUI) y relacionarlos en una red. La base de UMLS es el Metatesauro UMLS, que incluye términos y códigos de más de 200 vocabularios fuente y los relaciona en un formato común de datos. Además de la CIE, existen otras ontologías que se pueden utilizar para describir y clasificar enfermedades, como HPO y Mondo. HPO proporciona definiciones estandarizadas para anomalías fenotípicas relacionadas con enfermedades y Mondo es una ontología para describir enfermedades y relacionarlas en una red jerárquica.

Existen varias ontologías además del CIE (Clasificación Internacional de Enfermedades) para describir y clasificar a los pacientes en términos de enfermedad. Algunas de estas ontologías incluyen HPO (The Human Phenotype Ontology) que proporciona definiciones estandarizadas para anomalías fenotípicas relacionadas con enfermedades y Mondo que es una ontología que busca unificar las definiciones de enfermedades en todos los recursos de anotación de enfermedades. Hay diferentes formas de hacer un mapeo entre los códigos CIE

y los términos HPO, por ejemplo, a través de UMLS (Unified Medical Language System) o directamente a través de HPO OBO (Ontology for Biomedical Investigations).

Finalmente, los biobancos de imágenes y biomarcadores poblacionales han contribuido significativamente al avance tecnológico gracias a su capacidad para proporcionar una gran cantidad de datos y muestras diversas para su análisis y utilización en investigaciones médicas y de salud. Estos biobancos permiten la recopilación y almacenamiento sistemático de imágenes y biomarcadores, lo que facilita la comparación y el análisis de grandes de muestras heterogéneas, lo que a su vez conduce a un mejor entendimiento de enfermedades y trastornos y a un desarrollo más eficiente de nuevos tratamientos y terapias.

Referencias

1. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001 Mar;69(3):89-95. doi: 10.1067/mcp.2001.113989. PMID: 11240971
2. Martí-Bonmatí, L., & Alberich-Bayarri, A. (2016). Imaging biomarkers: Development and clinical integration. Imaging Biomarkers: Development and Clinical Integration, 1–376. https://doi.org/10.1007/978-3-319-43504-6/COVER
3. Martí Bonmatí, L., Alberich-Bayarri, A., García-Martí, G., Sanz Requena, R., Pérez Castillo, C., Carot Sierra, J. M., & Manjón Herrera, J. v. (2012). Biomarcadores de imagen, imagen cuantitativa y bioingeniería. Radiología, 54(3), 269–278. https://doi.org/10.1016/J.RX.2010.12.013
4. About Biomarkers and Qualification | FDA. (n.d.). Retrieved December 7, 2022, from https://www.fda.gov/drugs/biomarker-qualification-program/about-biomarkers-and-qualification
5. Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images are more than pictures, they are data. Radiology, 278(2), 563–577. https://doi.org/10.1148/RADIOL.2015151169/ASSET/IMAGES/LARGE/RADIOL.2015151169.FIG7.JPEG
6. Martí-Bonmatí, L. (2019). Radiomics and imaging biomarkers in observational clinical studies with retrospective data. ANALES RANM, 136(01), 34–42. https://doi.org/10.32440/AR.2019.136.01.REV07
7. Behrenbruch, C. P., Petroudi, S., Bond, S., Declerck, J. D., Leong, F. J., & Brady, J. M. (2014). Image filtering techniques for medical image post-processing: an overview. Http://Dx.Doi.Org/10.1259/Bjr/17464219,77(SPEC. ISS. 2). https://doi.org/10.1259/BJR/17464219
8. Vovk, U., Pernuš, F., & Likar, B. (2007). A review of methods for correction of intensity inhomogeneity in MRI. IEEE Transactions on Medical Imaging, 26(3), 405–421. https://doi.org/10.1109/TMI.2006.891486
9. Placidi, G., & Sotgiu, A. (2004). A novel algorithm for the reduction of undersampling artefacts in magnetic resonance images. Magnetic Resonance Imaging, 22(9), 1279–1287. https://doi.org/10.1016/J.MRI.2004.09.010
10. Lohman, B., Boric-Lubecke, O., Lubecke, V. M., Ong, P. W., & Sondhi, M. M. (2002). A digital signal processor for Doppler radar sensing of vital signs. IEEE Engineering in Medicine and Biology Magazine,21(5), 161–164. https://doi.org/10.1109/MEMB.2002.1044188
11. Ashburner, J., & Friston, K. J. (1999). Nonlinear Spatial Normalization Using Basis Functions. Hum. Brain Mapping, 7, 254–266. https://doi.org/10.1002/(SICI)1097-0193(1999)7:4<254::AID-HBM4>3.0.CO;2-G
12. Clarke, L. P., Velthuizen, R. P., Camacho, M. A., Heine, J. J., Vaidyanathan, M., Hall, L. O., Thatcher, R. W., & Silbiger, M. L. (1995). MRI segmentation: methods and applications. Magnetic Resonance Imaging, 13(3), 343–368. https://doi.org/10.1016/0730-725X(94)00124-L
13. Tomaszewski, M. R., & Gillies, R. J. (2021). The biological meaning of radiomic features. Radiology, 298(3), 505-516

14. Spadarella, G., Stanzione, A., Akinci D'Antonoli, T., Andreychenko, A., Fanni, S. C., Ugga, L., ... & Cuocolo, R. (2022). Systematic review of the radiomics quality score applications: An EuSoMII Radiomics Auditing Group Initiative. European Radiology, 1-1
15. Vista de Representación conceptual de imágenes médicas digitales: Integración de Contexto y Contenido Visual. (n.d.). Retrieved December 7, 2022, from https://revistas.ucm.es/index.php/RGID/article/view/54719/49922
16. Workload- and Data-based Automated Design for a Hybrid Row-Column Storage Model and Bloom Filter-Based Query Processing for Large-Scale DICOM Data Management. (n.d.). Retrieved December 7, 2022, from https://www.researchgate.net/publication/329828718_Workload-_and_Data-based_Automated_Design_for_a_Hybrid_Row-Column_Storage_Model_and_Bloom_Filter-Based_Query_Processing_for_Large-Scale_DICOM_Data_Management
17. Noumeir, R. (2003). Dicom Structured Report Document Type Definition. IEEE Transactions on Information Technology in Biomedicine,7(4), 318–328. https://doi.org/10.1109/TITB.2003.821334
18. Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., Lijmer, J. G., Moher, D., Rennie, D., & de Vet, H. C. W. (2012). Towards Complete and Accurate Reporting of Studies of Diagnostic Accuracy: The STARD Initiative. Http://Dx.Doi.Org/10.2214/Ajr.181.1.1810051,181(1), 51–56. https://doi.org/10.2214/AJR.181.1.1810051
19. Datta S, Godfrey-Stovall J, Roberts K. RadLex Normalization in Radiology Reports. AMIA Annu Symp Proc. 2021 Jan 25;2020:338-347. PMID: 33936406; PMCID: PMC8075450.
20. Köhler S, Gargano M, Matentzoglu N, Carmody LC, Lewis-Smith D, Vasilevsky NA, Danis D, Balagura G, Baynam G, Brower AM, Callahan TJ, Chute CG, Est JL, Galer PD, Ganesan S, Griese M, Haimel M, Pazmandi J, Hanauer M, Harris NL, Hartnett MJ, Hastreiter M, Hauck F, He Y, Jeske T, Kearney H, Kindle G, Klein C, Knoflach K, Krause R, Lagorce D, McMurry JA, Miller JA, Munoz-Torres MC, Peters RL, Rapp CK, Rath AM, Rind SA, Rosenberg AZ, Segal MM, Seidel MG, Smedley D, Talmy T, Thomas Y, Wiafe SA, Xian J, Yüksel Z, Helbig I, Mungall CJ, Haendel MA, Robinson PN. The Human Phenotype Ontology in 2021. Nucleic Acids Res. 2021 Jan 8;49(D1):D1207-D1217.
21. Mungall CJ, McMurry JA, Köhler S, Balhoff JP, Borromeo C, Brush M, Carbon S, Conlin T, Dunn N, Engelstad M, Foster E, Gourdine JP, Jacobsen JO, Keith D, Laraway B, Lewis SE, NguyenXuan J, Shefchek K, Vasilevsky N, Yuan Z, Washington N, Hochheiser H, Groza T, Smedley D, Robinson PN, Haendel MA. The Monarch Initiative: an integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res. 2017 Jan 4;45(D1):D712-D722.
22. Köhler S, Schulz MH, Krawitz P, Bauer S, Dölken S, Ott CE, Mundlos C, Horn D, Mundlos S, Robinson PN. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. Am J Hum Genet. 2009 Oct;85(4):457-64.
23. Smedley D, Jacobsen JO, Jäger M, Köhler S, Holtgrewe M, Schubach M, Siragusa E, Zemojtel T, Buske OJ, Washington NL, Bone WP, Haendel MA, Robinson PN. Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat Protoc. 2015 Dec;10(12):2004-15.

24. Kafkas Ş, Althubaiti S, Gkoutos GV, Hoehndorf R, Schofield PN. Linking common human diseases to their phenotypes; development of a resource for human phenomics. J Biomed Semantics. 2021 Aug 23;12(1):17.
25. Bastarache L, Hughey JJ, Goldstein JA, Bastraache JA, Das S, Zaki NC, Zeng C, Tang LA, Roden DM, Denny JC. Improving the phenotype risk score as a scalable approach to identifying patients with Mendelian disease. J Am Med Inform Assoc. 2019 Dec 1;26(12):1437-1447

Acrónimos y Abreviaturas

AG	Asamblea General
IMPaCT	Infraestructura de Medicina de Precisión asociada a la Ciencia y la Tecnología
IMPaCT-Data	Programa de ciencia de datos de IMPaCT
CD	Comité Dirección
WP	Paquete de trabajo (siglas en inglés, Work Package)