Manual de Buenas Prácticas para la Inteligencia Artificial en Medicina

Editado y elaborado por:

OMC ORGANIZACIÓN MÉDICA COLEGIAL DE ESPAÑA | CONSEJO GENERAL DE COLEGIOS OFICIALES DE MÉDICOS

Coordinación Editorial

Dra. María Isabel Moya García,
Vicepresidenta de la OMC y Doctora en Medicina.

Dr. José Antonio Trujillo Ruiz,
Vicepresidente del Colegio de Médicos de Málaga
y Director del Commálaga Health Hub.

El **“Manual de Buenas Prácticas para la Inteligencia Artificial en Medicina”** es una iniciativa de la Organización Médica Colegial (OMC) y forma parte de su proyecto estratégico para establecer un marco técnico, ético, y deontológico en la práctica clínica, en el contexto de la incorporación de la salud digital a la medicina. Todo ello guiado por un principio rector innegociable: Medicina inteligente, decisiones humanas.

Editado y elaborado por: Organización Médica Colegial (OMC).

Coordinación Editorial:

Dra. María Isabel Moya García,
Vicepresidenta de la OMC y Doctora en Medicina.

Dr. José Antonio Trujillo Ruiz,
Vicepresidente del Colegio de Médicos de Málaga y Director del Commálaga Health Hub.

Dirección de Estrategia y Despliegue Técnico:
Alejandro Trujillo, Responsable de IA en Fresh Health.

Comité de Expertos y Autoría:

Dra. María Isabel Moya García,
Vicepresidenta de la OMC y Doctora en Medicina.

Dr. José Ibeas,
Director del Programa para la Promoción y Desarrollo de la IA en el Sistema de Salud de Cataluña, Presidente de GEMAV y Director del Grupo de Nefrología Clínica, Intervencionista y Computacional del Parc Taulí.

Dr. Julio Mayol,
Catedrático de Cirugía y Director Científico del Instituto de Investigación Sanitaria San Carlos - IdISSC.

Dr. José Antonio Trujillo,
Vicepresidente del Colegio de Médicos de Málaga y Experto en IA aplicada al sector médico.

Dra. Cristina Gil Membrado,
Catedrática de Derecho Civil. Universidad de las Islas Baleares. Investigadora principal del Proyecto El Derecho ante la Salud Digital, Personalizada y Robótica (SALUDPYR). Ministerio de Ciencia, Innovación y Universidades.

Dr. José María Domínguez Roldán
Presidente de la Comisión de Ética y Deontología Médica de la Organización Médica Colegial. Profesor de la Facultad de Ciencias de la Salud y de la Vida. Universidad CEU Fernando III, CEU Universities.

OMC ORGANIZACIÓN MÉDICA COLEGIAL DE ESPAÑA | CONSEJO GENERAL DE COLEGIOS OFICIALES DE MÉDICOS

Proyecto PID2022-136964NB-I00 El Derecho ante la Salud Digital, Personalizada y Robótica (SALUDPYR) financiado por MICIU/ AEI/10.13039/501100011033/ y por FEDER, UE

Sumario

Prefacio
IA en medicina: diez garantías para médicos y pacientes

1. Diez principios para el médico 11
2. Diez garantías para el paciente 14

Capítulo I. La realidad asistencial. La IA en la práctica clínica del día a día

1. Palabras clave 17
2. Marco teórico 17
2.1. La transición hacia una medicina basada en datos 17
2.2. De la investigación a la implementación clínica 18
2.3. IA como inteligencia aumentada: el papel del profesional sanitario 18
2.4. Estrategia europea para la inteligencia artificial en salud 21
2.5. Iniciativas institucionales en España 22
2.6. IA como producto sanitario: aspectos regulatorios esenciales 23
2.6.1 Clasificación del riesgo 26
2.7. Validación clínica y adaptación al entorno asistencial 29
2.8. Ficha técnica de una herramienta de IA clínica 30
2.9. Actualizaciones y mantenimiento de los sistemas de IA 32
2.9.1 Actualizaciones de software tradicionales 32
2.9.2 Modelos adaptativos 32
2.10. Interoperabilidad con la historia clínica electrónica 33
3. Aplicación práctica 34
3.1. Sistemas de soporte a la decisión clínica 34
3.2. Interpretación asistida de pruebas diagnósticas 35
3.3. Estratificación de riesgo y predicción clínica 35
3.4. Automatización de documentación clínica 36
3.5. Herramientas clínicas frente a herramientas administrativas 36
3.5.1 Herramientas que son producto sanitario 37
3.5.2 Herramientas de apoyo administrativo o documental 37
3.6. Recomendaciones operativas para la implementación clínica 39
3.6.1 Priorizar herramientas con validación clínica 39
3.6.2 Integración con la historia clínica electrónica 39
3.6.3 Formación básica de los profesionales 39

3.6.4 Supervisión humana permanente 40
3.6.5 Evaluación continua del impacto 40
3.7. Clasificación práctica de herramientas de IA en la asistencia sanitaria 41
3.8. Integración segura en el flujo clínico 42
3.9. Gobernanza clínica de la inteligencia artificial 43
3.10. Checklist para la evaluación de herramientas de IA 44
4. Bibliografía 47

Capítulo II. IA en el ámbito del I+D: el futuro de la investigación

1. Palabras clave y glosario específico 53
1.1. Palabras clave 53
1.2. Glosario mínimo (para lectura clínica) 53
2. Bloque teórico 55
2.1. De la investigación "lineal" a la investigación aumentada 55
2.2. Modelos fundacionales 56
2.3. De la evidencia clásica a la evidencia híbrida: RWD/RWE 56
2.4. Ensayos clínicos adaptativos y "data-driven" 57
2.5. Gemelos digitales: simulación individualizada con exigencia clínica 58
2.6. Sesgo, equidad y validez externa: el cuello de botella 58
2.7. Gobernanza y reproducibilidad algorítmica 59
2.8. Integridad científica en la era generativa 61
2.9. Traslación clínica: del hallazgo al impacto 61
2.10. Límites y condiciones reales de implementación 62
3. Bloque práctico 63
3.1. Mapa operativo: dónde aporta valor la IA en I+D 63
3.2. Checklist de "entregable válido" en un proyecto de IA 64
3.3. Protocolo mínimo en 6 pasos para desplegar IA en I+D 65
3.4. Escenarios prácticos (para anticipar errores típicos) 67
4. Bibliografía 68

Capítulo III. Derechos y obligaciones del médico ante la IA

1. Palabras claves: 71
2. Bloque teórico: 71
2.1. Derechos de cuarta generación 71
2.2. La figura clave del desplegador (deployer) en el *AI act.* 72
2.3. Derechos fundamentales de los médicos frente a la ia. 75
2.4. Objección de conciencia e IA. 80

3. Bloque práctico: 82
3.1. Triaje automatizado con criterios injustificados: 84
3.2. Sistema de ayuda al diagnóstico "caja negra" con falsos negativos: 84
3.3. Modelos no adaptados al contexto local: 85
3.4. Automatización excesiva en salud mental: 85
4. Bibliografía 87

Capítulo IV. Entre Datos y Diagnósticos: IA Médica con Garantías

1. Palabras claves: 91
2. Bloque teórico: 91
2.1. ¿Por qué hablar de Protección de Datos Personales e IA? 91
2.2. ¿Cómo funciona la IA médica? 93
2.3. Los datos: el combustible de la IA 93
2.4. RGPD vs Ley de IA ¿Qué regula cada norma? 94
2.5. ¿Cuándo se pueden utilizar datos personales de salud? 95
2.6. Decisiones automatizadas: límite crítico. 95
2.7. La supervisión humana del profesional sanitario es imprescindible. 97
2.8. Sesgos en IA: un riesgo clínico real. 99
2.9. Derechos del paciente sobre datos personales utilizados por la IA..... 100
2.10 Uno para todos y todos para uno: el Espacio Europeo de Datos de Salud (EEDS/EHDS). 103
2.11. A modo de conclusión. 105
3. Bloque práctico: 107
3.1. ¿Qué debe saber el médico cuando utiliza la IA? 107
3.2. Decálogo para el médico que utiliza la IA. 107
4. Bibliografía: 108
5. Normativa utilizada 109

Capítulo V. Deontología y ética médica en tiempos de inteligencia artificial

1. Palabras claves 111
2. Bloque teórico 111
2.1. Planteamiento: cuando la técnica altera la moral del acto médico 111
2.2. El marco deontológico español: continuidad y actualización 112
3. Parte Práctica 128
3.1. Hacia una ética operativa 128
3.2. Prospectiva: escenarios futuros y vigilancia ética 129
4. Bibliografía 130

Prefacio

La medicina se enfrenta hoy a una de las encrucijadas más determinantes y esperanzadoras de su historia contemporánea: la irrupción de la inteligencia artificial en nuestro ecosistema sanitario. Esta revolución, que trasciende la mera digitalización para reconfigurar el diagnóstico, la investigación y la propia toma de decisiones, interpela directamente el núcleo de nuestra vocación, donde convergen el rigor científico, el cuidado del paciente y la exigencia ética.

Quienes tenemos la responsabilidad de velar por la excelencia en la práctica clínica asumimos el compromiso ineludible de acompañar a la profesión en esta transición. Partimos de una convicción inquebrantable: la tecnología es una herramienta extraordinaria para expandir nuestras capacidades, pero jamás podrá sustituir el juicio clínico, la empatía y la responsabilidad indelegable del profesional. Nuestro deber es asegurar que el uso de estos sistemas se realice bajo la más estricta prudencia, respeto deontológico y siempre al servicio de las personas.

Con este propósito nace la presente guía. Más que un manual teórico, se erige como una brújula conceptual y práctica diseñada para integrar la inteligencia artificial en la medicina de forma segura, transparente y conforme a los nuevos marcos normativos. Aspiramos a infundir confianza en que la vanguardia tecnológica y el humanismo no son fuerzas opuestas, sino aliadas indispensables para garantizar una atención sanitaria equitativa y de mayor calidad.

Este documento se concibe como un esfuerzo transversal que aborda el fenómeno algorítmico desde todas sus aristas: desde la realidad asistencial en las consultas, pasando por el nuevo horizonte de la investigación biomédica (I+D), hasta llegar al delicado terreno de la protección de datos y el reconocimiento de los nuevos derechos de cuarta generación. No nos limitamos a cartografiar los riesgos, como los sesgos, la opacidad o el peligro de deshumanización, sino que trazamos rutas seguras y ofrecemos criterios operativos para una adopción responsable.

Es, ante todo, una herramienta viva para el día a día. Una guía que asegura que pilares irrenunciables como la primacía del paciente, la independencia clínica y el secreto profesional mantengan su plena vigencia en un entorno dominado por los datos.

Invitamos a todos los médicos, investigadores y gestores sanitarios a hacer suyo este texto. Solo a través del conocimiento, la formación y la aplicación reflexiva en la práctica diaria lograremos que la innovación cumpla su verdadera promesa, evitando que la técnica altere la moral del acto médico. También a los pacientes y ciudadanos que deben comprender la integración de la innovación en los principios básicos de la medicina y qué nuevos roles vamos a desempeñar en las decisiones sobre su salud, su vida y su enfermedad.

"Nuestra convicción es firme: la inteligencia artificial debe ser el instrumento; el cuidado, el fin."

Medicina inteligente, decisiones humanas.

Dr. Tomás Cobo Castro
Presidente de la OMC

IA en medicina: diez garantías para médicos y pacientes

Dra. María Isabel Moya García
Vicepresidenta de la OMC y Doctora en Medicina

Los diez principios que siguen resumen lo esencial que todo médico debe tener presente antes de utilizar sistemas de inteligencia artificial en su práctica profesional, y ofrecen al mismo tiempo un marco comprensible de garantías para los pacientes.

1. Diez principios para el médico

1. La inteligencia artificial no sustituye al médico

La inteligencia artificial puede ampliar la capacidad diagnóstica, analítica y organizativa del profesional, pero nunca reemplazar su juicio clínico, su responsabilidad deontológica ni su deber de cuidado. La decisión final sobre el paciente debe seguir siendo humana, prudente y clínicamente contextualizada.

2. Toda IA clínica exige supervisión humana significativa

La supervisión médica no puede reducirse a una mera formalidad. El profesional debe conservar la capacidad real de comprender, revisar, contrastar, matizar y, en su caso, rechazar las recomendaciones del sistema antes de que estas produzcan efectos sobre la atención sanitaria, evitando todo proceso de automatización ciega.

3. Ninguna herramienta debe utilizarse sin validación suficiente (check-list)

No debe incorporarse a la práctica clínica ningún sistema de IA que no haya demostrado adecuadamente su utilidad, su seguridad y su pertinencia en el entorno asistencial concreto en el que vaya a ser empleado. La validación local y la evaluación de su impacto en resultados en salud constituyen una exigencia de prudencia y buena práctica.

4. No hay IA segura sin datos de calidad y adecuadamente protegidos

La fiabilidad de cualquier sistema depende de la calidad, representatividad e integridad de los datos que lo nutren. A ello se añade la obligación de preservar la privacidad, limitar accesos, minimizar los datos utilizados y reforzar las medidas de seguridad para evitar usos indebidos o resultados injustos.

5. La equidad debe vigilarse de forma activa

La IA puede reproducir o amplificar desigualdades ya existentes. Por ello, su rendimiento debe evaluarse en distintos subgrupos de pacientes y corregir cualquier sesgo que comprometa la justicia, la calidad asistencial o la seguridad clínica. La innovación solo merece ese nombre cuando mejora la atención sin excluir ni perjudicar a los más vulnerables.

6. La ética y la deontología médica deben guiar la implementación de la inteligencia artificial

La ética y la deontología no llegan al final para corregir la tecnología ni para impedir la innovación, sino al principio para orientarla. La IA no puede reemplazar el núcleo moral de la medicina; la deontología no es un freno externo, sino una disciplina interna de la profesión y la brújula que permite integrar la IA con prudencia, responsabilidad y sentido humanista.

7. El paciente tiene derecho a una información clara y comprensible

Cuando la inteligencia artificial participe de forma relevante en el proceso asistencial, el paciente debe poder conocer su utilización, entender qué papel desempeña y cuáles son sus límites. Esa información debe ser clínicamente útil, comprensible y compatible con una relación médico-paciente basada en la confianza y el respeto.

8. El médico tiene derecho a garantías profesionales ante la IA

La implantación de la IA no solo impone deberes al profesional; también exige reconocerle garantías: formación adecuada, condiciones seguras de uso, acceso a información suficiente sobre la herramienta, apoyo institucional y capacidad real de discrepancia cuando el sistema no sea fiable, no resulte adecuado al caso o entre en conflicto con la buena práctica clínica.

9. La inteligencia artificial debe acelerar la investigación sin comprometer la integridad científica

La IA está transformando la forma en que el médico investiga, analiza la evidencia y obtiene conocimiento útil para la práctica clínica, desde el descubrimiento biomédico hasta el diseño de ensayos clínicos y el análisis de datos del mundo real. Su utilización exige trazabilidad, supervisión metodológica, control de sesgos, vigilancia de la deriva de los modelos y prevención de prácticas que comprometan la validez científica. La buena investigación médica del futuro requerirá, además de excelencia científica, alfabetización en datos y capacidad crítica.

10. La trazabilidad, la gobernanza y la formación continua son imprescindibles

La seguridad en el uso de la IA no depende exclusivamente del médico. Requiere estructuras de gobernanza, protocolos de validación y monitorización, trazabilidad a lo largo de todo el ciclo de vida de los sistemas, registro de incidencias, criterios de retirada, soporte técnico y formación continuada. Utilizar IA con responsabilidad forma ya parte del profesionalismo médico contemporáneo

La incorporación de la inteligencia artificial a la medicina debe reforzar la confianza del paciente y no debilitarla. Por ello, junto a los principios que orientan al profesional, conviene expresar de forma clara las garantías mínimas que toda persona debe poder esperar cuando estas tecnologías participan en su atención.

2. Diez garantías para el paciente

1. La inteligencia artificial es una ayuda para su médico, no un sustituto

La atención sanitaria sigue descansando en la valoración del profesional, que es quien interpreta su situación, integra la información clínica y toma la decisión final.

2. Ninguna decisión importante sobre su salud debe quedar solo en manos de una máquina

Las decisiones relevantes deben contar siempre con supervisión médica y con una valoración individualizada de su caso.

3. Las herramientas de IA deben haber sido evaluadas antes de utilizarse

No basta con que una tecnología sea novedosa; debe haber demostrado que funciona adecuadamente y que puede emplearse con seguridad en el entorno clínico real.

4. Su médico debe poder revisar y cuestionar lo que indique la IA

Si una recomendación automatizada no encaja con su situación clínica, no debe aplicarse automáticamente.

5. Sus datos de salud deben estar protegidos. No compartirlos no es la solución.

La utilización de estas herramientas exige medidas estrictas de privacidad, seguridad y control de acceso a la información que se han desarrollado en un marco jurídico actualizado para toda Europa.

6. La IA debe funcionar de manera justa

No debe ofrecer peores resultados por razones de edad, sexo, enfermedad, discapacidad u otras circunstancias que puedan generar desigualdad.

7. Tiene derecho a saber cuándo y para qué se usa IA en tu atención

La transparencia es parte del buen cuidado. Debe poder conocer, de manera comprensible, qué función cumple esa herramienta en su proceso asistencial.

8. Tiene derecho a pedir explicaciones

La tecnología no debe dificultar la comunicación con el profesional, sino reforzarla. Poder preguntar y recibir respuestas claras forma parte de una atención respetuosa y segura.

9. Los centros sanitarios deben supervisar estas herramientas de forma continua

La responsabilidad no recae solo en el médico. Los hospitales y organizaciones sanitarias deben vigilar su funcionamiento, registrar incidencias y retirarlas si no son útiles o seguras.

10. La buena medicina sigue necesitando humanidad

La inteligencia artificial puede ayudar mucho, pero nunca debe sustituir la escucha, la empatía, la prudencia ni la relación de confianza entre su médico y usted.

La inteligencia artificial puede y debe contribuir a una mejor medicina. Pero su legitimidad dependerá siempre de que se utilice con prudencia, con transparencia, con supervisión humana real y con pleno respeto a la dignidad del paciente, a la independencia clínica del médico y a los valores de la profesión.

La realidad asistencial. La IA en la práctica clínica del día a día

Dr. José Ibeas
Director del Programa para la Promoción y Desarrollo de la IA en el Sistema de Salud de Cataluña, Presidente de GEMAV y Director del Grupo de Nefrología Clínica, Intervencionista y Computacional del Parc Taulí

1. Palabras clave

Inteligencia artificial, Inteligencia aumentada, Sistema de soporte a la decisión clínica, Medicina basada en datos, Producto sanitario software, Historia clínica electrónica, Diagnóstico asistido por IA, Gestión de datos sanitarios, Seguridad del paciente, Supervisión humana de la IA, Interoperabilidad sanitaria.

2. Marco teórico

2.1. La transición hacia una medicina basada en datos

La medicina contemporánea genera una cantidad creciente de información: historias clínicas electrónicas, pruebas de laboratorio, imágenes médicas, monitorización continua, datos genómicos y datos generados por dispositivos digitales de salud.

El volumen y complejidad de estos datos superan en muchos casos la capacidad de análisis individual del profesional sanitario. En este contexto, la inteligencia artificial permite identificar patrones clínicos complejos, correlaciones y señales predictivas que pueden ser difíciles de detectar mediante análisis tradicionales, facilitando el desarrollo de modelos de medicina personalizada basados en la estratificación del riesgo, la predicción de la evolución clínica y la adaptación de las intervenciones terapéuticas a las características individuales de cada paciente.

Las tecnologías utilizadas en este ámbito incluyen principalmente:

- Aprendizaje automático (machine learning)
- Aprendizaje profundo (deep learning)
- Procesamiento del lenguaje natural (NLP)

Estas tecnologías permiten construir sistemas capaces de aprender a partir de datos clínicos históricos y generar predicciones o recomendaciones clínicas de forma individualizada, adaptadas a las características y al contexto clínico de cada paciente.

2.2. De la investigación a la implementación clínica

Durante la última década, la mayoría de los desarrollos en IA médica se han producido en el ámbito de la investigación. Sin embargo, el verdadero reto consiste en trasladar estos sistemas al entorno clínico real e integrarlos de forma segura y eficaz, donde deben convivir con flujos asistenciales complejos, sistemas de información heterogéneos y contextos clínicos variables.

Para que estas herramientas puedan incorporarse de forma efectiva a la práctica clínica diaria, es necesario que cumplan una serie de requisitos fundamentales:

- Validación clínica rigurosa
- Integración con los sistemas de información sanitaria
- Interpretabilidad suficiente para el profesional
- Fiabilidad y seguridad
- Impacto real en resultados clínicos o eficiencia asistencial

En este sentido, la IA debe entenderse como una herramienta de apoyo a la decisión, no como un sustituto del proceso clínico. La implementación clínica de la IA requiere, por tanto, un enfoque multidisciplinar que combine conocimientos clínicos, tecnológicos, regulatorios y organizativos.

2.3. Inteligencia artificial como inteligencia aumentada: el papel del profesional sanitario

La integración de la IA redefine parcialmente el rol del médico. En lugar de reemplazar su función, estas herramientas permiten liberar tiempo clínico, mejorar la precisión diagnóstica y apoyar decisiones complejas.

El profesional sanitario sigue siendo responsable de:

- Interpretar los resultados generados por los sistemas
- Contextualizarlos en la situación clínica del paciente
- Tomar la decisión final

La IA no sustituye el razonamiento clínico, sino que lo complementa proporcionando herramientas que:

- Facilitan el análisis de información compleja
- Identifican patrones difíciles de detectar
- Ayudan a priorizar riesgos clínicos
- Reducen tareas administrativas

En este contexto, el uso de la inteligencia artificial en medicina se enmarca en el concepto de *inteligencia aumentada*, entendido como un modelo en el que la tecnología amplía la capacidad de análisis y apoyo a la decisión del profesional sanitario, sin sustituir su juicio clínico. Los sistemas de IA pueden analizar grandes volúmenes de información, identificar patrones complejos o generar predicciones clínicas, pero su interpretación y aplicación en el contexto de cada paciente siguen dependiendo del médico.

En este modelo, el profesional sanitario permanece en el centro del proceso asistencial y mantiene en todo momento la responsabilidad final de la decisión clínica. Este enfoque se describe habitualmente como "*human in the loop*", en el que los sistemas de inteligencia artificial actúan como herramientas de apoyo dentro del circuito de decisión clínica y requieren supervisión y validación por parte del profesional.

2.4. Estrategia europea para la inteligencia artificial en salud

La Comisión Europea considera que la inteligencia artificial puede contribuir de forma significativa a transformar el futuro de la medicina al facilitar el análisis de grandes volúmenes de datos clínicos, mejorar la precisión diagnóstica, apoyar la toma de decisiones clínicas y avanzar hacia modelos de medicina personalizada. Asimismo, estas tecnologías pueden ayudar a optimizar la organización de los sistemas sanitarios mediante modelos predictivos que anticipen necesidades asistenciales y permitan una mejor planificación de recursos. En la actualidad, más de 180 algoritmos de inteligencia artificial aplicados al ámbito sanitario cuentan ya con marcado CE en Europa, obtenido a través de organismos notificados conforme al Reglamento Europeo de Productos Sanitarios, lo que refleja el crecimiento progresivo de estas tecnologías en la práctica clínica.

No obstante, la integración de la inteligencia artificial en la asistencia sanitaria plantea también diversos retos. Entre los más relevantes destacan la calidad y representatividad de los datos utilizados para entrenar los algoritmos, la interoperabilidad entre sistemas de información sanitarios, la transparencia y explicabilidad de los modelos, la validación clínica en entornos reales, la protección de datos personales y la necesidad de mantener siempre la supervisión humana en la toma de decisiones médicas.

Para abordar estos desafíos, la Unión Europea ha desarrollado un marco legislativo específico que regula el desarrollo y la utilización de la inteligencia artificial en el ámbito sanitario. Este marco incluye principalmente el Reglamento Europeo de Productos Sanitarios (MDR) y el Reglamento de Diagnóstico In Vitro (IVDR), que regulan los dispositivos médicos basados en software, así como el Reglamento Europeo de Inteligencia Artificial (AI Act), que establece un enfoque basado en el riesgo para garantizar que los sistemas de IA sean seguros, transparentes y fiables. Estas normas se complementan con la legislación europea de protección de datos (RGPD) y con iniciativas como el Espacio Europeo de Datos de Salud (European Health Data Space), orientadas a facilitar el uso seguro de los datos sanitarios para investigación, innovación y mejora de la asistencia.

2.5. Iniciativas institucionales en España

Además del marco europeo, en España se están desarrollando iniciativas específicas para impulsar el uso responsable de la inteligencia artificial en el sistema sanitario. En septiembre de 2025, el Ministerio de Sanidad anunció la puesta en marcha de la Estrategia de Inteligencia Artificial del Sistema Nacional de Salud (eIASNS), concebida como una hoja de ruta para promover el desarrollo y la adopción segura y coordinada de estas tecnologías en el SNS. Esta estrategia tiene como objetivo mejorar la calidad asistencial, optimizar la eficiencia del sistema y favorecer la investigación y la innovación en salud, al tiempo que establece mecanismos de gobernanza, evaluación y supervisión que garanticen un uso seguro, ético y equitativo de la inteligencia artificial. Entre sus líneas de actuación destacan el desarrollo de capacidades en IA entre los profesionales sanitarios, la identificación de casos de uso clínico relevantes, la mejora de la interoperabilidad de los sistemas de información y la integración progresiva de estas tecnologías en los procesos asistenciales. Asimismo, la estrategia identifica retos clave para su implementación, como la disponibilidad de datos de calidad, infraestructuras tecnológicas adecuadas, validación clínica de los algoritmos e integración efectiva en los flujos de trabajo asistenciales.

En paralelo, la Agencia Española de Supervisión de Inteligencia Artificial (AESIA) ha publicado diversas guías orientadas a promover el desarrollo y uso responsable de la inteligencia artificial en distintos sectores, incluido el sanitario que proporcionan orientaciones prácticas para el desarrollo, evaluación y uso de sistemas de IA. Estas guías promueven un enfoque basado en el riesgo y en la supervisión humana, e incluyen recomendaciones sobre gobernanza de los sistemas de IA, calidad y representatividad de los datos, transparencia y explicabilidad de los algoritmos, gestión de sesgos, evaluación del impacto y monitorización continua tras su despliegue. Asimismo, subrayan la importancia de garantizar la protección de los derechos fundamentales, la seguridad de los sistemas y la confianza de los profesionales y pacientes en el uso de estas tecnologías.

Las guías también recomiendan establecer procesos de evaluación previa, monitorización continua y auditoría de los sistemas de IA, así como fomentar la formación de los profesionales y la creación de estructuras de gobernanza que supervisen su uso dentro de las organizaciones.

2.6. IA como producto sanitario: aspectos regulatorios esenciales

La implementación de inteligencia artificial en salud se desarrolla dentro de un marco regulatorio cada vez más definido.

- Las dos agencias regulatorias más influyentes son:
- La Food and Drug Administration (FDA) en Estados Unidos
- El sistema regulatorio europeo basado en MDR, IVDR y AI Act

Una parte importante de las aplicaciones de inteligencia artificial utilizadas en medicina son consideradas productos sanitarios software, conocidos como Software as a Medical Device (SaMD).

Ambos sistemas comparten principios regulatorios fundamentales:

- Seguridad del paciente
- Evidencia clínica
- Transparencia del algoritmo
- Gestión de riesgos
- Seguimiento postcomercialización

La FDA mantiene una base de datos pública de dispositivos médicos con inteligencia artificial autorizados. Actualmente existen más de 1400 dispositivos aprobados, concentrándose la mayoría en el ámbito de la radiología.

Las aplicaciones más frecuentes incluyen:

- Detección automática de lesiones en imagen médica
- Diagnóstico asistido de cáncer
- Análisis automatizado de ECG
- Detección de retinopatía diabética

Por su parte, en Europa el marco regulatorio incluye:

- Reglamento Europeo de Productos Sanitarios (MDR)
- Reglamento de Diagnóstico In Vitro (IVDR)
- Reglamento Europeo de Inteligencia Artificial (AI Act)

En Europa, estos productos están regulados por el Reglamento Europeo de Productos Sanitarios (MDR 2017/745).

Un software se considera producto sanitario cuando su función es:

- Diagnosticar enfermedades
- Apoyar decisiones diagnósticas
- Recomendar tratamientos
- Monitorizar pacientes
- Predecir evolución clínica

En estos casos, el software debe disponer del marcado CE, lo que certifica que cumple los requisitos de seguridad, calidad y rendimiento establecidos por la normativa europea.

El marcado CE implica que el producto ha sido evaluado por un organismo notificado y que el fabricante ha demostrado:

- Seguridad del sistema
- Calidad del proceso de desarrollo
- Evidencia clínica del rendimiento
- Gestión adecuada de riesgos

El proceso de obtención del marcado CE incluye:

- Calificación y clasificación como producto sanitario
- Preparación de documentación técnica
- Implantación de un sistema de gestión de calidad
- Evaluación de conformidad
- Registro en bases regulatorias europeas

La documentación técnica debe incluir:

- Descripción del producto
- Finalidad prevista
- Análisis beneficio-riesgo
- Evaluación clínica
- Verificación y validación
- Plan de control de cambios
- Seguimiento postcomercialización

El uso de herramientas de IA en decisiones clínicas sin la certificación correspondiente puede suponer un riesgo para la seguridad del paciente y una responsabilidad legal para las instituciones sanitarias.

2.6.1 Clasificación del riesgo

El Reglamento Europeo de Productos Sanitarios (MDR) establece diferentes clases de riesgo para los productos sanitarios software, que dependen del impacto potencial del sistema sobre la salud del paciente y del grado en que influye en las decisiones diagnósticas o terapéuticas.

Clase I (bajo riesgo)

Software que no influye directamente en decisiones clínicas.

Ejemplos:

- Organización de datos clínicos
- Herramientas administrativas
- Sistemas de gestión documental

Clase IIa (riesgo moderado)

Software que apoya decisiones clínicas pero cuya influencia es limitada.

Ejemplos:

- Sistemas de estratificación de riesgo
- Herramientas de apoyo al diagnóstico
- Análisis automatizado de ECG

Clase IIb (alto riesgo)

Software que influye de forma significativa en decisiones diagnósticas o terapéuticas y cuyo uso puede tener un impacto relevante en la salud del paciente.

Ejemplos:

- Sistemas avanzados de soporte diagnóstico
- Algoritmos de detección precoz de patologías graves
- Herramientas de apoyo a decisiones terapéuticas complejas

Clase III (muy alto riesgo)

Software que puede realizar funciones diagnósticas críticas o cuya utilización tiene un impacto directo en decisiones clínicas de alto riesgo para el paciente.

Ejemplos:

- Sistemas de diagnóstico automatizado en imagen médica
- Algoritmos de detección automática de cáncer
- Sistemas que generan recomendaciones terapéuticas directas para enfermedades graves

Clase	Nivel de riesgo	Ejemplo
I	Bajo	gestión de datos
IIa	Moderado	predicción clínica
IIb	Alto	soporte diagnóstico
III	Muy alto	diagnóstico automatizado

En general, cuanto mayor es la influencia del software en la toma de decisiones clínicas y el potencial impacto sobre la salud del paciente, mayor es la clase de riesgo asignada y más exigentes son los requisitos regulatorios para su evaluación y certificación.

2.7. Validación clínica y adaptación al entorno asistencial

Disponer de marcado CE es un requisito necesario, pero no siempre suficiente para garantizar que una herramienta funcione adecuadamente en todos los entornos clínicos.

Los modelos de IA pueden comportarse de forma diferente cuando se aplican en poblaciones distintas a aquellas con las que fueron entrenados.

Por este motivo es recomendable que las herramientas utilizadas en un centro sanitario:

- Hayan sido evaluadas en entornos clínicos similares
- Dispongan de estudios de validación externa
- Sean evaluadas localmente durante su implementación, mediante la realización de proyectos piloto o evaluaciones en el entorno asistencial del propio centro.

Esta validación puede incluir:

- Auditorías de rendimiento
- Comparación con práctica clínica habitual
- Análisis de impacto en resultados clínicos

La adaptación al contexto local es especialmente importante cuando existen diferencias en:

- Características demográficas de la población
- Protocolos clínicos
- Sistemas de información sanitaria
- Disponibilidad de recursos diagnósticos

La adaptación al contexto local es esencial para garantizar el valor clínico de estas tecnologías.

2.8. Ficha técnica de una herramienta de IA clínica

Toda herramienta de IA utilizada en práctica clínica debería disponer de una ficha técnica transparente y accesible para los profesionales sanitarios.

Esta ficha técnica permite comprender cómo funciona la herramienta y cuáles son sus limitaciones.

La información mínima que debería incluir es:

1. Finalidad clínica

Descripción clara del propósito del sistema.

Ejemplo:

"Detección automática de retinopatía diabética en imágenes de fondo de ojo."

2. Tipo de algoritmo

Descripción general del tipo de modelo utilizado.

Ejemplo:

Red neuronal convolucional entrenada con imágenes retinianas.

3. Datos de entrenamiento

Descripción de los datos utilizados para entrenar el modelo:

- Tamaño del dataset
- Origen de los datos
- Características de la población

4. Rendimiento del sistema

Indicadores de rendimiento clínico:

- Sensibilidad
- Especificidad
- AUC
- Tasas de falsos positivos y negativos

5. Limitaciones conocidas

Situaciones en las que el modelo puede tener menor rendimiento.

Ejemplo:

- Imágenes de baja calidad
- Poblaciones no representadas en el entrenamiento

6. Clasificación como producto sanitario

- Clase de riesgo
- Número de marcado CE

7. Requisitos técnicos

- Integración con sistemas clínicos
- Requisitos de hardware
- Conectividad

8. Procedimiento de actualización

Información sobre cómo se actualiza el sistema y cómo se gestionan los cambios en el modelo.

2.9. Actualizaciones y mantenimiento de los sistemas de IA

A diferencia de otros productos sanitarios tradicionales, muchos sistemas de IA evolucionan mediante actualizaciones periódicas.

Estas actualizaciones pueden incluir:

- Mejoras del algoritmo
- Corrección de errores
- Adaptación a nuevos datos clínicos

Es importante que las instituciones sanitarias conozcan cómo se gestionan estas actualizaciones.

Existen dos modelos principales:

2.9.1 Actualizaciones de software tradicionales

El algoritmo permanece fijo y se actualiza mediante nuevas versiones del software.

Cada actualización puede requerir validación adicional.

2.9.2 Modelos adaptativos

Algunos sistemas pueden actualizarse mediante aprendizaje continuo.

Este enfoque plantea retos regulatorios importantes, ya que los cambios en el algoritmo pueden modificar su comportamiento clínico.

Por este motivo, las actualizaciones deben incluir:

- Control de versiones
- Documentación de cambios
- Evaluación del impacto clínico

Las instituciones sanitarias deben establecer procedimientos para supervisar las actualizaciones de los sistemas de IA utilizados en práctica clínica.

2.10. Interoperabilidad con la historia clínica electrónica

Uno de los factores críticos para el éxito de la IA en entornos asistenciales es su capacidad de integrarse con los sistemas de información clínica existentes.

La interoperabilidad permite que los sistemas de IA:

- Accedan a los datos clínicos relevantes
- Devuelvan resultados directamente en la historia clínica
- Se integren en el flujo de trabajo del profesional

Sin esta integración, las herramientas pueden generar:

- Duplicación de tareas
- Pérdida de tiempo clínico
- Errores de transcripción

Los estándares de interoperabilidad más utilizados incluyen:

- HL7
- FHIR
- DICOM en imagen médica

La integración adecuada permite que las recomendaciones generadas por la IA aparezcan directamente en el contexto clínico donde el profesional toma decisiones. Por el contrario, la falta de interoperabilidad puede generar duplicación de tareas, aumentar el riesgo de errores y dificultar la adopción de estas herramientas por parte de los profesionales sanitarios.

3. Aplicación práctica

La incorporación de la inteligencia artificial en la práctica clínica se está produciendo de forma progresiva a través de diferentes herramientas de apoyo al profesional sanitario. Estas aplicaciones abarcan desde sistemas de soporte a la decisión clínica hasta herramientas de análisis de datos o automatización de tareas administrativas. En este apartado se describen algunos de los principales ámbitos en los que la inteligencia artificial está comenzando a integrarse en la asistencia sanitaria diaria, así como ejemplos de uso que ilustran su aplicación práctica en el cuidado del paciente.

3.1. Sistemas de soporte a la decisión clínica

Los sistemas de soporte a la decisión clínica (CDSS) representan una de las aplicaciones más consolidadas de la IA en medicina.

Estos sistemas analizan datos clínicos del paciente en tiempo real y ofrecen recomendaciones o alertas al profesional sanitario.

Aplicaciones frecuentes incluyen:

- Alertas de interacciones farmacológicas
- Recomendaciones terapéuticas basadas en guías clínicas
- Detección precoz de deterioro clínico
- Apoyo al diagnóstico diferencial

Ejemplo clínico:

En un paciente hospitalizado con infección, un sistema de IA puede identificar patrones tempranos compatibles con sepsis, alertando al equipo clínico antes de que se manifiesten signos clínicos evidentes.

3.2. Interpretación asistida de pruebas diagnósticas

La IA ha demostrado especial eficacia en el análisis de datos estructurados y de imágenes médicas.

Ámbitos con mayor desarrollo:

- Radiología
- Dermatología
- Oftalmología
- Anatomía patológica
- Cardiología (ECG)

Estos sistemas pueden detectar patrones sutiles que pueden pasar desapercibidos en una lectura convencional.

Ejemplo clínico:

Un sistema de IA en radiología puede analizar una tomografía torácica y señalar regiones sospechosas de nódulos pulmonares, facilitando su revisión por el radiólogo.

El objetivo no es sustituir la interpretación médica, sino aumentar la sensibilidad diagnóstica y reducir errores de omisión.

3.3. Estratificación de riesgo y predicción clínica

Los modelos predictivos permiten estimar la probabilidad de determinados eventos clínicos:

- Reingreso hospitalario
- Deterioro clínico
- Complicaciones postquirúrgicas

- Progresión de enfermedad crónica

Estos sistemas pueden ayudar a priorizar intervenciones clínicas y optimizar recursos asistenciales.

Ejemplo clínico:

Un modelo predictivo puede identificar pacientes con insuficiencia cardiaca con alto riesgo de reingreso en los próximos 30 días, permitiendo intensificar el seguimiento.

3.4. Automatización de documentación clínica

Una de las áreas con mayor potencial de impacto inmediato es la reducción de carga administrativa.

Aplicaciones emergentes incluyen:

- Transcripción automática de consultas
- Generación automática de informes clínicos
- Codificación clínica automatizada
- Resumen de historias clínicas

Estas herramientas permiten liberar tiempo clínico que puede dedicarse a la atención directa del paciente.

3.5. Herramientas clínicas frente a herramientas administrativas

Es importante distinguir entre dos grandes categorías de herramientas basadas en inteligencia artificial: aquellas que influyen en decisiones clínicas y aquellas destinadas a tareas administrativas o de apoyo documental. Esta diferenciación resulta fundamental desde el punto de vista clínico y regulatorio, ya que las primeras suelen considerarse productos sanitarios y están sujetas a requisitos regulatorios específicos, mientras que las segundas se orientan principalmente a mejorar la eficiencia de los procesos asistenciales y administrativos.

3.5.1 Herramientas que son producto sanitario

Son aquellas que influyen directamente en decisiones clínicas.

Ejemplos:

- Detección de cáncer en imágenes
- Predicción de deterioro clínico
- Sistemas de recomendación terapéutica

3.5.2 Herramientas de apoyo administrativo o documental

No influyen directamente en decisiones clínicas.

Ejemplos:

- Transcripción automática de consultas
- Generación de informes
- Resumen de historias clínicas
- Organización documental

Estas herramientas pueden mejorar la eficiencia del trabajo clínico, pero no se consideran necesariamente productos sanitarios.

Sin embargo, su uso debe seguir garantizando:

- Confidencialidad de los datos
- Seguridad de la información
- Supervisión humana

3.6. Recomendaciones operativas para la implementación clínica

Para una integración segura y eficaz de la IA en entornos asistenciales, se recomienda:

3.6.1 Priorizar herramientas con validación clínica

Los sistemas deben haber demostrado utilidad clínica en estudios robustos o en implementaciones reales. En el caso de herramientas que influyen en decisiones diagnósticas o terapéuticas, es necesario verificar que se trata de productos sanitarios con marcado CE, conforme al Reglamento Europeo de Productos Sanitarios (MDR) o al Reglamento de Diagnóstico In Vitro (IVDR).

En determinados casos, los centros sanitarios pueden desarrollar soluciones "*in house*", es decir, software desarrollado y utilizado internamente dentro de una organización sanitaria. Estas soluciones pueden quedar exentas de ciertos procedimientos de certificación, siempre que cumplan las condiciones establecidas en la normativa europea, como su uso exclusivo dentro de la institución, la justificación de la necesidad clínica y la aplicación de procedimientos adecuados de calidad, seguridad y evaluación clínica.

3.6.2 Integración con la historia clínica electrónica

Las herramientas deben integrarse de forma fluida en los sistemas existentes para evitar duplicación de tareas y asegurar la interoperabilidad con los sistemas clínicos, permitiendo que la información generada por la IA se incorpore directamente al flujo de trabajo asistencial y a la historia clínica electrónica.

3.6.3 Formación básica de los profesionales

Los profesionales sanitarios deben disponer de una formación básica que les permita comprender las capacidades, limitaciones y condiciones de uso de los sistemas de inteligencia artificial. Esto incluye conocer el tipo de datos en los que

se basan los algoritmos, el significado de las recomendaciones o predicciones generadas y los posibles sesgos o limitaciones del sistema. Esta formación es fundamental para interpretar adecuadamente los resultados proporcionados por la IA, integrarlos en el razonamiento clínico y mantener una supervisión crítica de las herramientas utilizadas en la práctica asistencial.

3.6.4 Supervisión humana permanente

Las decisiones clínicas deben mantenerse bajo la responsabilidad del profesional sanitario. Los sistemas de inteligencia artificial deben utilizarse como herramientas de apoyo a la decisión y no como sustitutos del juicio clínico. Por ello, es fundamental mantener un modelo de supervisión humana continua ("*human in the loop*"), en el que el profesional evalúe críticamente las recomendaciones generadas por el sistema, las interprete en el contexto clínico del paciente y determine su aplicabilidad en cada situación concreta. Este enfoque permite aprovechar el potencial analítico de la IA manteniendo la seguridad del paciente, la responsabilidad profesional y la calidad del proceso de toma de decisiones clínicas.

3.6.5 Evaluación continua del impacto

Tras la implementación de una herramienta de inteligencia artificial en el entorno asistencial, es necesario establecer mecanismos de "monitorización y evaluación continua" que permitan analizar su impacto real en la práctica clínica. Esto incluye evaluar resultados clínicos, indicadores de seguridad del paciente, eficiencia de los procesos asistenciales y grado de aceptación por parte de los profesionales. Asimismo, es importante detectar posibles desviaciones en el rendimiento del sistema, identificar efectos no previstos y revisar periódicamente su utilidad clínica, de forma que puedan realizarse ajustes, actualizaciones o incluso reconsiderar su uso si no aporta valor en el contexto asistencial.

3.7. Clasificación práctica de herramientas de IA en la asistencia sanitaria

Para facilitar su comprensión en la práctica clínica, resulta útil clasificar las herramientas de IA según su función y nivel de riesgo.

Tipo de herramienta	Ejemplo	¿Producto sanitario?	Nivel de riesgo aproximado
Automatización documental	Transcripción de consultas	Generalmente no	Bajo
Generación de informes	Resumen automático de historias clínicas	Generalmente no	Bajo
Organización de datos clínicos	Clasificación automática de documentos	No	Bajo
Soporte a la decisión clínica	Alertas de sepsis	Sí	Moderado
Predicción clínica	Riesgo de reingreso	Sí	Moderado
Análisis automatizado de ECG	Detección de arritmias	Sí	Moderado
Análisis de imagen médica	Detección de cáncer en radiología	Sí	Alto
Diagnóstico automatizado	Retinopatía diabética	Sí	Alto

3.8. Integración segura en el flujo clínico

Para que la IA tenga impacto real en la práctica clínica diaria debe integrarse adecuadamente en los flujos asistenciales.

Antes de implementar una herramienta de IA en un centro sanitario se recomienda evaluar:

Evaluación clínica

- Evidencia científica
- Utilidad clínica

Evaluación regulatoria

- Marcado CE
- Clasificación de riesgo

Evaluación técnica

- Interoperabilidad
- Requisitos de infraestructura

Evaluación operativa

- Integración en el flujo clínico
- Formación de profesionales

Evaluación continua

- Monitorización de su rendimiento clínico.

Los sistemas que requieren aplicaciones externas o procesos manuales adicionales suelen tener baja adopción.

Las características de una buena integración incluyen:

- Acceso automático a los datos clínicos relevantes
- Visualización de resultados dentro de la historia clínica
- Mínima interrupción del flujo de trabajo clínico
- Recomendaciones claras y comprensibles

Además, es importante evitar la denominada "fatiga de alertas", un fenómeno bien descrito en los sistemas de soporte a la decisión clínica. Cuando los sistemas generan demasiadas alertas, los profesionales tienden a ignorarlas. Por este motivo, el diseño de los sistemas de IA debe priorizar alertas clínicamente relevantes y bien contextualizadas.

3.9. Gobernanza clínica de la inteligencia artificial

La adopción segura de la inteligencia artificial en instituciones sanitarias requiere establecer estructuras de gobernanza adecuadas que permitan evaluar, supervisar y monitorizar el uso de estas tecnologías. En este contexto, resulta recomendable que los centros sanitarios dispongan de comités o grupos de evaluación de tecnologías basadas en IA, encargados de analizar su utilidad clínica, seguridad y adecuación al entorno asistencial.

Estos órganos deberían contar con la participación de:

- Profesionales clínicos
- Especialistas en informática médica
- Expertos en protección de datos
- Responsables de calidad y seguridad del paciente
- Gestores sanitarios

Estos comités pueden evaluar:

- Nuevas herramientas
- Impacto clínico
- Seguridad
- Cumplimiento regulatorio

La gobernanza institucional permite asegurar que la implementación de IA responde a necesidades asistenciales reales y mantiene estándares adecuados de seguridad y calidad.

3.10. Checklist para la evaluación de herramientas de inteligencia artificial antes de su implementación

Antes de incorporar una herramienta de inteligencia artificial en un entorno clínico, es recomendable que los centros sanitarios realicen una evaluación estructurada de su utilidad, seguridad y adecuación al contexto asistencial. Con este objetivo, se propone a continuación un checklist orientativo que puede servir de apoyo a responsables clínicos, gestores sanitarios y equipos de evaluación tecnológica en el proceso de análisis previo a su implantación.

Evaluación clínica

- ☐ ¿Existe una necesidad clínica real que justifique la herramienta?
- ☐ ¿Existe evidencia científica que respalde su uso?
- ☐ ¿Se han publicado estudios de validación clínica?

Evaluación regulatoria

- ☐ ¿La herramienta está clasificada como producto sanitario?
- ☐ ¿Dispone de marcado CE si corresponde?
- ☐ ¿Se conoce su clasificación de riesgo (I, IIa, IIb o III)?

Evaluación técnica

- ☐ ¿Es interoperable con la historia clínica electrónica del centro?
- ☐ ¿Utiliza estándares sanitarios (HL7, FHIR, DICOM)?
- ☐ ¿Cuál es el impacto en la infraestructura tecnológica?

Evaluación del algoritmo

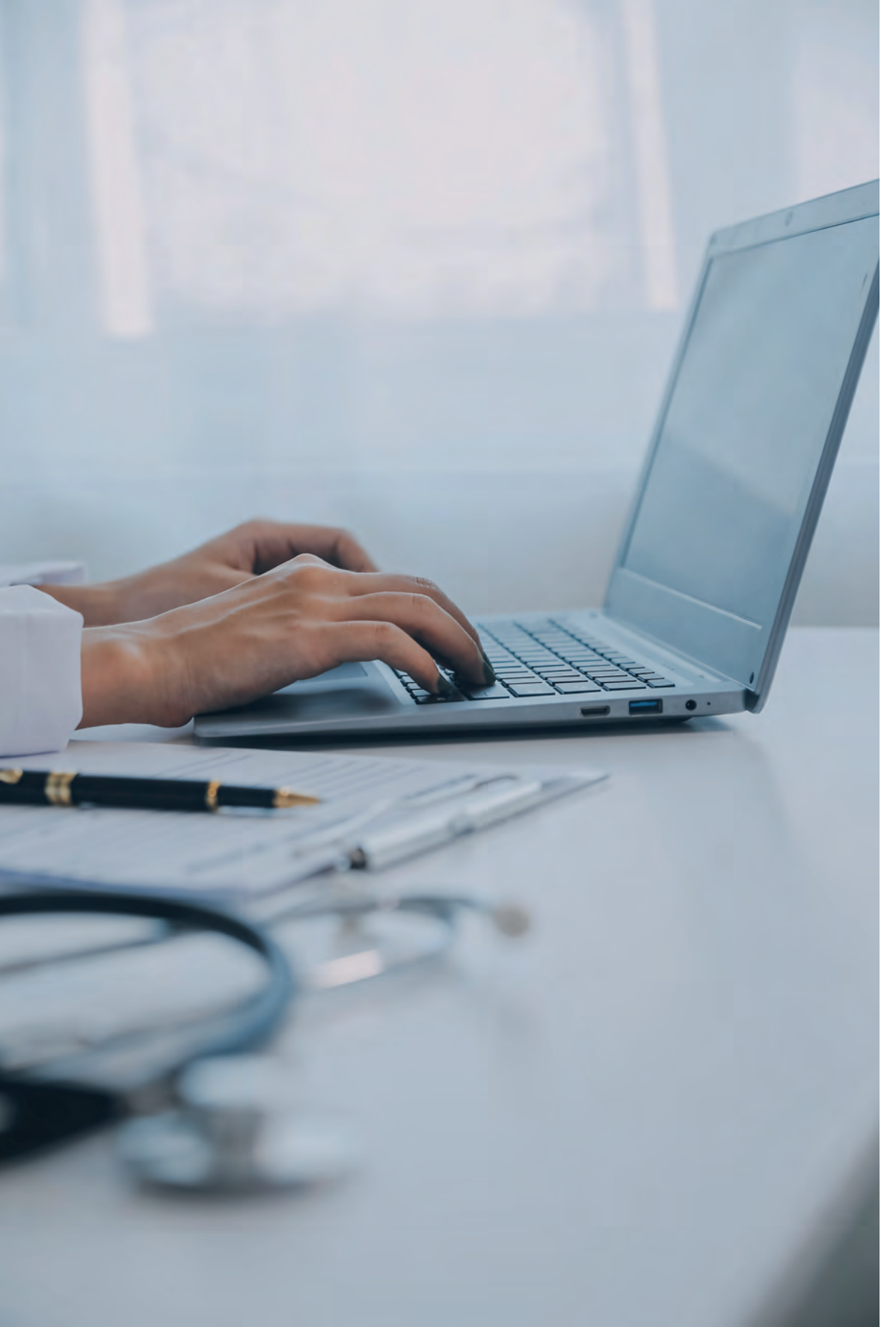
- ☐ ¿Se conoce el tipo de modelo utilizado?
- ☐ ¿Se dispone de información sobre los datos de entrenamiento?
- ☐ ¿Se conocen las limitaciones del sistema?

Evaluación operativa

- ☐ ¿Se integra en el flujo de trabajo clínico habitual?
- ☐ ¿Se ha planificado la formación de los profesionales?
- ☐ ¿Se ha definido un sistema de monitorización del rendimiento?

Seguridad y ética

- ☐ ¿Cómo se gestionan los datos clínicos?
- ☐ ¿Existe supervisión humana en las decisiones clínicas?
- ☐ ¿Se ha evaluado el riesgo de sesgos en el algoritmo?



4. Bibliografía

1. Declaración de la Asociación Médica Mundial sobre la Inteligencia artificial y aumentada en la Atención Médica. Acceso en https://www.wma.net/es/policies-post/declaracion-de-la-amm-sobre-la-inteligencia-artificial-y-aumentada-en-la-atencion-medica/ Consultado 2 de marzo de 2026

2. Inteligencia artificial en la asistencia sanitaria. Comisión Europea. Acceso: https://health.ec.europa.eu/ehealth-digital-health-and-care/artificial-intelligence-healthcare_es#legislaci%C3%B3n-de-la-ue-que-da-forma-a-la-ia-en-la-asistencia-sanitaria. Consultado 2 de marzo de 2026

3. Guias de la Agencia Española de Supervisión de Inteligencia Artificial (AESIA). Acceso: https://aesia.digital.gob.es/es/guias Consultado 2 de marzo de 2026

4. Estrategia de Inteligencia Artificial para el Sistema Nacinal de Salud. EIASNS. Septiembre 2025. Acceso www.sanidad.gob.es/areas/saludDigital/doc/eIASNS_v13.pdf Consultado 2 de marzo de 2026

5. Guía de evaluación de tecnologías de salud digital que incorporan inteligencia artificial (IA). AquAS. Generalitat de Cataluya. Acceso en https://aquas.gencat.cat/web/.content/minisite/aquas/publicacions/2024/guia-evaluacion-tecnologias-salud-digital-IA-aquas2024.pdf Consultado del 2 de marzo de 2026

6. Ibeas J, Suppi R, Eizaguerri S, Oliver I. El cambio de paradigma en el uso primario y secundario de los datos o la transformación estructural del sistema sanitario en la era de la inteligencia artificial. Medicina Clínica. En prensa.

7. Wachter RM, Brynjolfsson E. Will Generative AI Deliver on Its Promise in Health Care? JAMA. 2024;331(1):65-69.

8. Warraich HJ, Tazbaz T, Califf RM. FDA Perspective on the Regulation of Artificial Intelligence in Health Care and Biomedicine. JAMA. 2025 Jan 21;333(3):241-247.

9. FDA. Artificial intelligence & medical products: how CBER, CDER, CDRH, and OCP are working together. Published online March 2024. Disponible en: https://www.fda.gov/media/177030/ Consultado 2 de marzo 2026

10. Parlamento Europeo y Consejo de la Unión Europea. Reglamento (UE) sobre inteligencia artificial (AI Act). Diario Oficial (DO) de la Unión Europea del 12 de julio de 2024. Disponible en: https://artificialintelligenceact.eu/es/el-acto/ Consultado 2 de marzo 2026

11. Youssef A, et al. External Validation of AI Models in Health Should Be Replaced with Recurring Local Validation. Nat Med. 2023;29(11):2686–2687.

12. Gobierno de España. Agencia Española de Supervisión de la Inteligencia Artificial (AESIA). Ministerio de Asuntos Económicos y Transformación Digital. 2024. Disponible en: https://aesia.digital.gob.es/es Consultado 2 de marzo 2026

13. Guias de referencia. Soporte a iniciativas. Estrategia para la promoción y adoción de la inteligencia artificial en el sistema de salud de Catalunya. Disponible en https://iasalut.cat/es/suport-a-iniciatives/ Consultado 28 febrero de 2026.

14. European Commission. European Health Data Space (EHDS). 2023. Disponible en: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en Acceso 20 de febrero de 2026

15. European Medicines Agency. Guideline on computerised systems and electronic data in clinical trials. 2023. Disponible en: www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-computerised-systems-and-electronic-data-clinical-trials_en.pdf Acceso 20 de febrero de 2026

16. Ibeas J. Anàlisi de l'enfocament regulador de l'FDA en intel·ligència artificial aplicada a la salut: una perspectiva des d'Europa. *Annals de Medicina*, *ISSN-e 2013-7109, ISSN 0210-7465, Vol. 108, Nº. 2, 2025, págs. 57-61.*

17. ISO 14971:2019. Medical devices, Application of risk management to medical devices. 2019. Consultable en: https://www.iso.org/standard/72704.html. Acceso 20 de febrero de 2026

18. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group; SPIRIT-AI and CONSORT-AI Steering Group; SPIRIT-AI and CONSORT-AI Consensus Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med. 2020 Sep;26(9):1351-1363.

19. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019 Oct 29;17(1):195.

20. NICE. Evidence standards framework for digital health technologies. 2022. Disponible en: https://www.nice.org.uk/what-nice-does/digital-health/evidence-standards-framework-esf-for-digital-health-technologies Consultado 20 de febrero de 2026

21. Parlamento Europeo y Consejo de la Unión Europea. Reglamento (UE) 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial (AI Act). 2024. Consultable en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689 Consultado 20 de febrero de 2026.

22. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019 Apr 4;380(14):1347-1358.

23. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med. 2020 Feb 6;3:17.

24. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden MD, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson P, Weber W, Wheatstone P, McCulloch P; DECIDE-AI expert group. Reporting

guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022 May;28(5):924-933.

25. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019 Jan;25(1):44-56.

26. Wachter RM, Brynjolfsson E. Will Generative Artificial Intelligence Deliver on Its Promise in Health Care? JAMA. 2024 Jan 2;331(1):65-69.

27. Youssef A, Pencina M, Thakur A, Zhu T, Clifton D, Shah NH. External validation of AI models in health should be replaced with recurring local validation. Nat Med. 2023 Nov;29(11):2686-2687.

28. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. JAMA. 2018 Apr 3;319(13):1317-1318.

II

IA en el ámbito del I+D: el futuro de la investigación

Dr. Julio Mayol
Catedrático de Cirugía y Director Científico del Instituto de Investigación Sanitaria San Carlos - IdISSC

1. Palabras clave y glosario específico

1.1. Palabras clave

Modelos fundacionales; Ensayos clínicos adaptativos; RWD/RWE (datos/evidencia del mundo real); Gemelos digitales; Gobernanza y reproducibilidad algorítmica; Sesgo y validez externa; Integridad científica en la era generativa; Traslación clínica.

1.2. Glosario mínimo (para lectura clínica)

- RWD (Real-World Data): datos generados en la práctica clínica real (HCE, registros, dispositivos, etc.).
- RWE (Real-World Evidence): conclusiones sobre efectividad/seguridad/uso derivadas de RWD mediante un diseño y análisis adecuados.
- Calibración: grado en que las probabilidades predichas por un modelo se corresponden con las frecuencias observadas del evento en poblaciones externas o temporales independientes.
- Drift: pérdida de rendimiento por cambios en población, práctica clínica, tecnología o codificación con el tiempo.
- Validación externa: evaluación del modelo fuera del entorno de desarrollo (otro centro y/o otro periodo y/o otra población).
- AUC: métrica de discriminación; mide capacidad de ordenar riesgos, no garantiza utilidad clínica ni calibración.
- Emulación de ensayo (target trial emulation): aproximación causal con datos observacionales que intenta reproducir un ensayo ideal mediante criterios explícitos.

Predicción, causalidad y decisión: tres preguntas distintas:

Predicción: "¿Qué probabilidad tiene este evento en este paciente/población?" (útil para estratificar, priorizar, asignar recursos).

Causalidad: "¿Qué pasaría si cambiamos la intervención?" (requiere diseño causal explícito: aleatorización o inferencia causal observacional basada en supuestos identificables, control de confusión y análisis de sensibilidad).

Decisión clínica: "¿Qué haré diferente y qué resultado relevante mejorará?" (exige demostrar impacto en proceso o resultados, no solo métricas del modelo).

Regla práctica: un modelo puede predecir bien y, aun así, ser inadecuado para justificar cambios terapéuticos o conclusiones causales.

2. Bloque teórico

2.1. De la investigación "lineal" a la investigación aumentada

Durante décadas, la investigación biomédica se ha organizado como una cadena relativamente estable: observación clínica, hipótesis, diseño experimental, publicación y traslación. La IA introduce un cambio relevante: no solo añade una herramienta, sino que puede intervenir en varias fases del ciclo de I+D.

En una investigación "aumentada", la IA puede sugerir hipótesis, priorizar variables, detectar patrones y automatizar tareas analíticas. Esto desplaza el foco desde el "análisis al final" hacia la orquestación de todo el ciclo de investigación.

En la práctica: cuanto más temprano entra la IA en el ciclo, más importante es fijar reglas de diseño y documentación antes de ver resultados, para evitar sesgos retrospectivos.

2.2. Modelos fundacionales: nueva infraestructura cognitiva para la biomedicina

Los modelos fundacionales (foundation models) aprenden representaciones generales a partir de grandes volúmenes de datos y se adaptan después a tareas específicas. En biomedicina pueden trabajar con lenguaje clínico, biología molecular, señales, imagen o combinaciones multimodales.

Aportan velocidad (reutilización de capacidades), pero elevan la exigencia: cuanto más potente y general es el sistema, más importante es delimitar su uso, auditar sesgos y demostrar generalización.

Preguntas críticas para el lector

- ¿Qué datos y poblaciones han alimentado el modelo, y qué sesgos puede haber heredado?
- ¿Qué depende del contexto local (codificación, protocolos, dispositivos, lenguaje) y puede no transferirse?
- ¿Qué mecanismos permiten reconstruir decisiones del modelo y del pipeline analítico?

2.3. De la evidencia clásica a la evidencia híbrida: RWD/RWE

La IA potencia el uso de datos del mundo real (RWD) y la generación de evidencia del mundo real (RWE) para preguntas de efectividad, seguridad y patrones de uso en poblaciones amplias y heterogéneas.

Pero RWD no equivale a evidencia por sí mismo: los datos clínicos reales suelen ser incompletos, sesgados por la práctica asistencial y sensibles a cambios organizativos. La IA puede extraer señales útiles o amplificar confusión.

Condiciones mínimas para RWE creíble:

- Calidad y trazabilidad del dato (origen, periodo, población, codificación y transformaciones).
- Diseño robusto (comparador, desenlaces, control de confusión; cuando proceda, emulación de ensayos y sensibilidad).
- Verificación clínica y metodológica; validación externa cuando se pretenda generalizar.

En la práctica: si el resultado parece "demasiado bueno", exige revisar confusión, sesgo de selección y definición de cohortes.

2.4. Ensayos clínicos adaptativos y "data-driven": eficiencia sin perder causalidad

Los ensayos adaptativos buscan evidencia causal con menos coste y tiempo. La IA puede ayudar en selección de centros/pacientes, estimación de reclutamiento, monitorización de seguridad y análisis intermedios.

Riesgo central: adaptatividad sin reglas predefinidas y trazabilidad degrada credibilidad causal (multiplicidad, decisiones opacas, sesgo de selección).

Condiciones no negociables:

- Reglas preespecificadas: qué puede cambiar, cuándo y con qué control del error.
- Trazabilidad: reconstrucción completa de decisiones adaptativas (datos y análisis).
- Supervisión: comité con competencia clínica y metodológica.

2.5. Gemelos digitales: simulación individualizada con exigencia clínica

Los gemelos digitales pretenden representar de forma dinámica al paciente o al sistema biológico para explorar escenarios. Son prometedores, pero su validez depende de construcción, evaluación y mantenimiento.

Riesgos científicos típicos

- Confundir predicción con explicación causal.
- Pérdida de generalización cuando cambian protocolos, población o tecnología (drift).
- Complejidad que dificulta auditoría y mantenimiento.

Cómo evaluarlos en práctica

- Validación externa + calibración.
- Pruebas de estrés ante cambios de contexto y análisis por subgrupos clínicamente relevantes.
- Monitorización y recalibración como requisito, no como "extra".

2.6. Sesgo, equidad y validez externa: el cuello de botella

La IA aprende de datos con historia (decisiones previas, desigualdades, prácticas de codificación). Un modelo puede ser preciso y no transferible, o producir desigualdad si falla por subgrupos.

Regla práctica: si no hay validación externa real, no hay evidencia suficiente para generalizar. La validación externa no es solo "otro hospital": también es otro tiempo, otro flujo y, cuando procede, otra población.

- Qué debería incluir una validación externa útil
- Otro centro y/o otro periodo (idealmente ambos).
- Comparación con un baseline clínico razonable.
- Rendimiento por subgrupos relevantes (sexo, edad, comorbilidad; socioeconómico cuando proceda).
- Calibración y consecuencias clínicas de error (falsos positivos/negativos).

2.7. Gobernanza y reproducibilidad algorítmica: el método científico operacional

En IA, el resultado depende de decisiones técnicas que deben documentarse: preprocesamiento, particiones, hiperparámetros, manejo de missingness, selección de variables y versiones de software.

Reproducibilidad algorítmica significa poder reconstruir completamente el pipeline analítico (datos, preprocesamiento, particiones, hiperparámetros, versiones y entorno computacional) y verificar la robustez de los resultados.

Elementos mínimos

- Control de versiones: datos, código, modelo y entorno.
- Trazabilidad del pipeline de extremo a extremo.
- Documentación estándar (propósito, población, limitaciones y riesgos).
- Plan de monitorización (drift, recalibración, incidentes y retirada).
- Roles definidos: quién valida, aprueba, despliega, monitoriza y detiene.

2.8. Integridad científica en la era generativa

La IA generativa multiplica productividad (redacción, síntesis, código), pero también riesgos: errores verosímiles, citas defectuosas, datos sintéticos no validados y p-hacking por automatización.

La respuesta es disciplina: transparencia de uso, verificación humana y trazabilidad de fuentes, datos y decisiones analíticas.

Uso responsable de IA generativa en investigación (mínimo práctico)

- Aceptable como apoyo: borradores, mejora de claridad, ayuda en código con revisión y pruebas, extracción preliminar para verificación posterior.
- No aceptable sin verificación humana independiente: citas bibliográficas, resultados numéricos, afirmaciones fácticas o interpretaciones científicas generadas por el modelo.
- Obligatorio: revisión humana documentada en puntos críticos (datos, métodos, resultados, referencias).
- Transparencia: describir la tarea, la herramienta (y versión si procede) y los controles aplicados.

2.9. Traslación clínica: del hallazgo al impacto

La traslación con IA exige demostrar utilidad clínica, sostenibilidad y legitimidad. Muchos modelos funcionan en laboratorio, pocos cambian la práctica; el fracaso suele ser por integración, métricas inadecuadas, mantenimiento y gobernanza.

Tres pruebas de realidad

- Utilidad clínica: demostrar cambio en decisión/proceso y efecto en outcomes relevantes o seguridad.

- Sostenibilidad: monitorización, recalibración, mantenimiento y recursos asignados.
- Legitimidad: explicabilidad útil, seguridad, equidad y posibilidad real de retirada.

2.10. Límites y condiciones reales de implementación

La aplicabilidad de este marco depende de capacidades locales. Para evitar programas "de escaparate", conviene explicitar límites frecuentes:

- Datos insuficientes o no interoperables (HCE heterogénea, codificación variable, falta de variables clave).
- Dependencia de proveedores: acceso parcial a datos/modelos, cajas negras contractuales o limitaciones para auditar.
- Falta de equipo híbrido: sin metodólogo/estadístico e ingeniería de datos, la IA se convierte en un piloto eterno.
- Coste de mantenimiento subestimado: el trabajo real empieza tras el primer modelo (drift, recalibración, cambios en flujos).
- Entornos pequeños: priorizar casos de uso acotados y colaborar en redes para validación externa.

3. Bloque práctico

3.1. Mapa operativo: dónde aporta valor la IA en I+D (sin perder el control científico)

Para orientar decisiones rápidas, conviene ordenar los casos de uso por momento del ciclo de investigación y por nivel de riesgo:

A) Antes del estudio (ideación y diseño)

- Priorización de preguntas y cohortes: identificar vacíos de evidencia y poblaciones infraestudiadas.
- Diseño metodológico asistido: simulaciones de tamaño muestral, selección de variables, definición de endpoints clínicamente relevantes.
- Feasibility y selección de centros: estimación realista de elegibilidad y reclutamiento.

B) Durante el estudio (ejecución y monitorización)

- Reclutamiento y estratificación: cribado eficiente, enriquecimiento de cohortes, balance de grupos.
- Monitorización de seguridad: detección temprana de señales y desviaciones del protocolo.
- Calidad del dato: detección de inconsistencias, codificación anómala y patrón de datos faltantes.

C) Después del estudio (análisis, validación y traslación)

- Analítica RWD/RWE: emulación de ensayos, análisis de efectividad/seguridad y subgrupos predefinidos.
- Validación externa y recalibración: robustez en otros centros/periodos y análisis de drift.
- Traslación clínica: integración en flujos, evaluación de impacto y plan de mantenimiento.

Regla práctica: la IA puede acelerar (A) y (B), pero solo "cuenta" como evidencia sólida si refuerza (C): validez externa, reproducibilidad y utilidad clínica.

3.2. Checklist de "entregable válido" en un proyecto de IA para investigación

Esta lista permite decidir rápidamente si un proyecto está listo para avanzar o debe rediseñarse. Si falla algún bloque, el proyecto puede seguir como exploratorio, pero no debería presentarse como evidencia sólida ni escalarse.

1) Pregunta y utilidad

- ¿La pregunta es clínicamente relevante y medible (no solo "predice bien")?
- ¿El modelo cambiará una decisión o un proceso concreto?

2) Datos (calidad y trazabilidad)

- Origen, periodo y población descritos con precisión.
- Variables clave definidas (cómo se obtienen y codifican).
- Gestión explícita de datos faltantes (qué falta y por qué).
- Registro de transformaciones/pipeline (reproducible).

3) Diseño científico

- Separación estricta entre entrenamiento/validación/test, evitando fuga de información (data leakage) temporal, institucional o derivada del preprocesamiento.
- Comparador clínico (baseline útil) y métricas acordes al objetivo (incluida calibración cuando proceda).
- Análisis de subgrupos predefinido y clínicamente justificado (evitar "pesca" de resultados).

4) Validez externa y sesgo

- Validación externa real (otro centro y/o otro periodo).
- Evaluación de sesgo y equidad (por sexo, edad, nivel socioeconómico, etc., cuando aplique).
- Calibración del modelo (riesgo predicho vs observado).

5) Gobernanza y reproducibilidad algorítmica

- Control de versiones (datos, código y modelo).
- Documentación mínima: propósito, población, limitaciones y riesgos.
- Plan de monitorización (drift, rendimiento, incidentes).
- Responsables definidos: quién valida, quién aprueba, quién detiene.

6) Integridad científica

- Trazabilidad de resultados y del uso de IA generativa (si se utiliza) en texto y/o código.
- Auditoría interna de coherencia (revisión por pares técnica y clínica).

3.3. Protocolo mínimo en 6 pasos para desplegar IA en I+D (instituto/hospital)

1. Selección del caso de uso: elegir problemas con dolor real (reclutamiento, variabilidad, endpoints difíciles, alta carga analítica, necesidad de estratificación).
2. Equipo híbrido: clínico responsable, metodólogo/estadístico, científico de datos, ingeniería de datos y figura de gobernanza (calidad y cumplimiento).
3. Diseño y preregistro interno (y externo cuando proceda): definir antes de acceder a los resultados población, objetivo, métricas, comparadores, subgrupos, criterios de parada y estrategia analítica.
4. Evaluación técnica y clínica: revisar calibración, rendimiento por subgrupos, errores clínicamente críticos y escenarios donde el modelo no debe usarse.
5. Validación externa + prueba en flujo real: validar fuera del dataset original e integrar en proceso (tiempos, usabilidad, aceptación).
6. Monitorización y mantenimiento: umbrales de alarma, recalibración, auditoría periódica y procedimiento de retirada.

ENDOSKELETON
MONITOR 001
CEREBRUM
22.46
37.54
MONITOR 002
MYOCARDIUM
MONITOR 003
R.A. 4.08
L.A. 15.44
59.91
93.06
R.V.
L.V.
SCAN
CT
EQ
STATS ANALYSIS & DATABASE
3.8
72.37
13.89
55.05
90.11
33.27
CONTROL CENTER
FILES & DATA
DENSITY SCANS
EXIT

3.4. Escenarios prácticos (para anticipar errores típicos)

Escenario 1: RWE que "confirma" demasiado

Un análisis con RWD sugiere que un tratamiento es claramente superior. Al revisar, se detecta que los pacientes del grupo A tenían mejor seguimiento, menos comorbilidad y más acceso a pruebas.

Lección: en RWE, la IA puede amplificar confusión. Se exige diseño causal (p. ej., emulación de ensayo), análisis de sensibilidad y revisión clínica/metodológica.

Escenario 2: Ensayo adaptativo con decisiones opacas

Durante un ensayo adaptativo, el algoritmo propone modificar la asignación a un brazo por una señal intermedia. El equipo no puede explicar la regla exacta ni reconstruir el camino analítico.

Lección: adaptatividad sin trazabilidad erosiona causalidad. Toda decisión debe estar preespecificada, auditada y reproducible.

Escenario 3: Gemelo digital que deja de "ser cierto"

Un gemelo digital predice descompensaciones con gran acierto, pero tras un cambio de protocolo asistencial baja su rendimiento y aumentan falsos positivos.

Lección: drift clínico-organizativo. Sin monitorización y recalibración, la IA produce ruido y fatiga clínica.

Escenario 4: Modelo fundacional que "suena bien" pero falla en el detalle

Un modelo genera resúmenes de literatura y propuestas de hipótesis convincentes. Al verificar citas y datos, aparecen errores sutiles y omisiones relevantes.

Lección: la IA generativa acelera el trabajo, pero obliga a una cultura de verificación. En investigación, lo verosímil no equivale a verdadero.

4. Bibliografía

1. Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., van Calster, B., … TRIPOD+AI group. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ, 385, e078378. https://doi.org/10.1136/bmj-2023-078378

2. Dimairo, M., Pallmann, P., Wason, J., Todd, S., Jaki, T., Julious, S. A., … Mander, A. P. (2020). The Adaptive designs CONSORT Extension (ACE) statement: A checklist with explanation and elaboration guideline for reporting randomised trials that use an adaptive design. BMJ, 369, m115. https://doi.org/10.1136/bmj.m115

3. European Medicines Agency. (s. f.). Real-world evidence. European Medicines Agency. Recuperado el 6 de febrero de 2026, de https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/real-world-evidence

4. Food and Drug Administration. (2018). Framework for FDA's Real-World Evidence Program. U.S. Food and Drug Administration. https://www.fda.gov/media/120060/download

5. Food and Drug Administration. (s. f.). Good machine learning practice for medical device development: Guiding principles. U.S. Food and Drug Administration. Recuperado el 6 de febrero de 2026, de https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

6. Heads of Medicines Agencies & European Medicines Agency. (2024). Big Data Workplan 2023–2025: HMA/EMA joint Big Data Steering Group (Version 1.4, June 2024). European Medicines Agency. https://www.ema.europa.eu/en/documents/work-programme/workplan-2023-2025-hma-ema-joint-big-data-steering-group_en.pdf

7. International Medical Device Regulators Forum. (2025). Good machine learning practice for medical device development: Guiding principles (IMDRF/AIML WG/N88 FINAL:2025). https://www.imdrf.org/sites/default/files/2025-01/IMDRF_AIML%20WG_GMLP_N88%20Final_0.pdf

8. Iqbal, J. D., Krauthammer, M., Witt, C. M., Christen, M., & coautores. (2025). A consensus statement on the use of digital twins in medicine. npj Digital Medicine, 8, Article 484. https://doi.org/10.1038/s41746-025-01897-4

9. Mayol, J., Hicks, C. W., & Wexner, S. D. (2025). *Generative artificial intelligence and scientific publishing: Turning noise into trust. Surgery, 183*, Article 109496. https://doi.org/10.1016/j.surg.2025.109496

10. Moons, K. G. M., Damen, J. A. A., Kaul, T., Hooft, L., Andaur Navarro, C., Dhiman, P., … coautores. (2025). PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ, 388, e082505. https://doi.org/10.1136/bmj-2024-082505

11. Tudor, B. H., Shargo, R., Gray, G. M., Fierstein, J. L., Kuo, F. H., Burton, R., … Ahumada, L. M. (2025). A scoping review of human digital twins in healthcare applications and usage patterns. npj Digital Medicine, 8, Article 587. https://doi.org/10.1038/s41746-025-01910-w

12. World Health Organization. (2021). Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization. https://www.who.int/publications/i/item/9789240029200

13. World Health Organization. (2025). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. World Health Organization. https://www.who.int/publications/i/item/9789240084759

14. Xie, Q., Chen, Q., Chen, A., Peng, C., Hu, Y., Lin, F., … Bian, J. (2025). Medical foundation large language models for comprehensive text analysis and beyond. npj Digital Medicine, 8, Article (online). https://doi.org/10.1038/s41746-025-01533-1

III

Derechos y obligaciones del médico ante la IA

Dr. José Antonio Trujillo
Vicepresidente 1º Colegio de Médicos de Málaga.
Director del Commálaga Health Hub.
Máster en IA aplicada a sanidad.

1. Palabras claves:

Derechos de cuarta generación, Desplegador (deployer), Dignidad algorítmica, Objeción de conciencia tecnológica, Gobernanza algorítmica en la práctica clínica.

2. Bloque teórico:

2.1. Derechos de cuarta generación

La irrupción de la inteligencia artificial (IA) en la práctica médica ha reabierto el debate sobre la actualización de las generaciones de derechos fundamentales. La noción de "derechos de cuarta generación" se utiliza cada vez más para designar un haz de garantías vinculadas a la esfera digital, los datos y los algoritmos, que no sustituyen a los derechos clásicos, sino que los reformulan en un entorno de decisiones automatizadas. En el ámbito sanitario, estos derechos tienen una doble dimensión: la del paciente y la del profesional. Si el siglo XX consagró el derecho a la protección de la salud como derecho social, el siglo XXI exige reforzar la "dignidad algorítmica" de quienes reciben cuidados y de quienes los prestan.

Desde la teoría de los derechos humanos se ha propuesto agrupar bajo esta cuarta generación a los llamados derechos "digitales": sobre los datos, la identidad digital, la explicación de las decisiones automatizadas y sobre la protección frente a formas nuevas de vigilancia y discriminación. En medicina, este bloque se traduce en garantías concretas frente a sistemas de IA capaces de influir de forma decisiva en el diagnóstico, el pronóstico o la priorización de recursos. La cuestión ya no es solo qué derechos asisten al paciente o al médico, sino cómo un modelo algorítmico puede erosionarlos de manera silenciosa si no se fijan salvaguardas jurídicas y organizativas adecuadas.

La regulación europea ha comenzado a articular este marco. El Reglamento General de Protección de Datos (RGPD) reconoce el derecho de la persona a no quedar sometida a decisiones exclusivamente automatizadas con efectos jurídicos o similares, y exige la posibilidad de intervención humana, de expresar el propio punto de vista y de impugnar el resultado. La futura aplicación plena del Reglamento de IA de la Unión Europea (AI Act) refuerza esta exigencia en los sistemas de alto riesgo en salud, imponiendo a proveedores y "desplegadores" (deployers) obligaciones de transparencia, trazabilidad, supervisión humana y registro de incidencias. Estas normas se han elaborado pensando, ante todo, en la protección de los derechos de los pacientes; sin embargo, sus implicaciones para los derechos de los médicos son igualmente profundas.

En conjunto, la introducción de los derechos de cuarta generación permite formular un marco normativo en el que la protección de los pacientes y la protección de los médicos dejan de ser polos opuestos. La IA puede contribuir a reducir inequidades, mejorar diagnósticos y liberar tiempo clínico, pero solo si se articula sobre tres pilares: transparencia útil para decidir, gobernanza clínica robusta, con autonomía profesional reforzada, y responsabilidad proporcionada entre todos los actores del ecosistema algorítmico. El objetivo último es un derecho a la salud aumentada que preserve, y no diluya, el pulso humano de la relación médico-paciente.

2.2. La figura clave del desplegador (deployer) en el ai act.

La gran novedad del AI Act respecto a otros marcos tecnológicos es que no se limita a regular a quienes desarrollan o comercializan sistemas de IA (providers), sino que introduce expresamente la figura del "deployer", que traducimos como *desplegador sanitario*, como sujeto regulado. Esto significa que el uso clínico de la IA deja de ser un espacio neutro: pasa a estar jurídicamente configurado y sometido a obligaciones propias, con impacto directo en la práctica médica cotidiana.

El artículo 3.4 del Reglamento (UE) 2024/1689 define al desplegador como toda persona física o jurídica, autoridad pública, agencia u otro organismo que utiliza un sistema de IA bajo su autoridad, salvo cuando lo haga en el marco de una actividad puramente personal o no profesional.

El artículo 26 del AI Act fija un catálogo de deberes que puede agruparse en tres bloques:

1. Uso conforme y medidas organizativas internas:

- Adoptar medidas técnicas y organizativas para garantizar que el sistema se utiliza de acuerdo con las instrucciones de uso facilitadas por el proveedor.

- Asignar la supervisión humana a personas con la competencia, formación y autoridad necesarias, proporcionando además los apoyos materiales y de tiempo para que esta supervisión sea real y efectiva (no meramente nominal).

- Implantar políticas de *AI literacy* (alfabetización en IA) para el personal que interactúe con el sistema, de forma proporcional al riesgo.

2. Gestión de datos, funcionamiento y registro:

- Garantizar que los datos de entrada bajo su control sean pertinentes, suficientemente representativos y de calidad, de acuerdo con las especificaciones del sistema.

- Monitorizar el funcionamiento del sistema de forma continua, detectando degradaciones de rendimiento, sesgos o comportamientos anómalos, e interrumpir su uso cuando exista un riesgo significativo para la seguridad o los derechos de las personas.

- Conservar los registros (logs) generados automáticamente por el sistema, en la medida en que estén bajo su control, durante el periodo mínimo previsto por el Reglamento y la normativa sectorial aplicable.

- Verificar, cuando proceda, que el sistema está correctamente inscrito en la base de datos europea de sistemas de IA de alto riesgo y abstenerse de utilizarlo si no cumple los requisitos de registro.

3. Transparencia, impacto en derechos y cooperación regulatoria:

- Informar a las personas afectadas cuando se utilicen determinados sistemas de alto riesgo que tomen o apoyen decisiones con efectos significativos sobre ellas, obligación que se añade a los deberes de información propios del RGPD.

- Realizar, en los supuestos previstos en el artículo 27, una evaluación de impacto sobre los derechos fundamentales previa al despliegue de determinados sistemas de alto riesgo, especialmente cuando los utilicen autoridades o se afecte de forma sistemática a personas vulnerables.

- Coordinar estas evaluaciones con las evaluaciones de impacto en protección de datos (DPIA) exigidas por el RGPD, lo que en la práctica obliga a diseñar marcos integrados de gobernanza de datos y de IA en las organizaciones sanitarias.

- Notificar incidentes graves y no conformidades al proveedor y, en su caso, a las autoridades de vigilancia de mercado y de protección de datos, cooperando en la adopción de medidas correctoras.

2.3. Derechos fundamentales de los médicos frente a la ia.

1. Autonomía clínica reforzada

La autonomía clínica es el derecho del médico a tomar la decisión final sobre el diagnóstico y el tratamiento, asistido pero nunca sustituido por la tecnología. El AI Act exige que los sistemas de alto riesgo permitan una supervisión humana efectiva, incluyendo la posibilidad de intervenir, modificar o interrumpir su funcionamiento para evitar daños a la salud o a los derechos fundamentales.

2. Debida diligencia profesional (lex artis) y "última palabra"

La lex artis ad hoc se actualiza para incorporar el uso razonable de la IA, pero no se sustituye por ella. El médico tiene el derecho-deber de usar la IA con la misma exigencia crítica que aplicaría a cualquier prueba diagnóstica, lo que incluye conocer sus limitaciones, la población en la que fue validada y las condiciones de uso seguro.

3. Condiciones de trabajo seguras (IA validada y mantenida)

En IA sanitaria, esto implica que el empleador y el desplegador (*deployer)* deben garantizar que:

- Los sistemas de IA que actúan como Software as a Medical Device (SaMD) cumplen el Reglamento (UE) 2017/745 (MDR): evaluación clínica, gestión de riesgos, vigilancia poscomercialización.

- Existen procesos formales de validación local, detección de *drift (deriva)* y mantenimiento correctivo antes y durante su uso.

Estas obligaciones son “sin perjuicio” de otras que el desplegador ya asume en virtud de la legislación sanitaria, de productos sanitarios, de protección de datos o de seguridad del paciente, lo que refuerza su posición como nodo central de responsabilidad en el ecosistema de IA clínica.

- Se puede suspender el sistema si se detectan problemas de seguridad, sin penalizar al profesional que lo reclama. El AI Act refuerza este derecho al exigir supervisión humana y registros de uso para sistemas de alto riesgo.

4. Derecho a la formación continua y a la alfabetización en IA

El AI Act introduce una obligación específica de alfabetización en IA: proveedores y desplegadores (*deployers)* deben asegurar que el personal que utiliza estos sistemas tenga un nivel adecuado de conocimientos para usarlos de forma segura y responsable.

5. Responsabilidad proporcional y reparto justo del riesgo

El médico tiene derecho a no ser convertido en "chivo expiatorio" de fallos que son, en realidad, de diseño algorítmico, ciberseguridad o implementación organizativa.

Este derecho se traduce en que:

- Los errores atribuidos a defectos del producto (p.ej. un *bug* que infradiagnostica sistemáticamente, datos de entrenamiento gravemente sesgados, fallo de ciberseguridad) se ventilan prioritariamente en el marco de la responsabilidad por producto.

- La organización sanitaria responde por incumplir sus deberes como *deployer* (formación insuficiente, ausencia de monitorización, no realización de evaluaciones de impacto, falta de protocolos).

- El médico solo asume responsabilidad cuando incumple su deber de diligencia (por ejemplo, no revisa una imagen o no explora a un paciente porque "la IA ya lo ha hecho").

6. Libertad académica e intelectual (alertas de seguridad y crítica pública)

Para que la IA mejore, los profesionales deben poder investigar, documentar y comunicar sus fallos, sesgos y riesgos sin temor a represalias. La Directiva (UE) 2019/1937 sobre protección de informantes y su trasposición en España mediante la Ley 2/2023 protegen a quienes, en un contexto laboral, informan de infracciones normativas o riesgos graves para el interés público.

Este derecho implica:

- Posibilidad de notificar internamente y a la autoridad competente los problemas detectados en un sistema de IA sin sufrir represalias disciplinarias o contractuales.

- Capacidad de publicar resultados científicos sobre el rendimiento real, los sesgos o las limitaciones de un sistema, siempre respetando la confidencialidad de pacientes y los canales de seguridad.

- Protección frente a cláusulas de confidencialidad abusivas que intenten impedir la comunicación de riesgos para la salud pública.

7. Protección de los datos personales y de las métricas del profesional

Los sistemas de IA generan de forma rutinaria logs, métricas de rendimiento y patrones de uso que pueden referirse tanto a pacientes como a profesionales. El AI Act exige conservar registros para garantizar trazabilidad y seguridad, pero también obliga a informar a los trabajadores cuando la IA se utiliza en el contexto laboral, especialmente si se trata de sistemas de alto riesgo para gestión de personas.

El médico tiene derecho a que:

- Los datos de uso de la IA y sus métricas individuales no se conviertan en un sistema de vigilancia encubierta ni en base para sanciones automatizadas.
- Cualquier cuadro de mando que utilice indicadores derivados de IA para evaluar su desempeño esté sometido a evaluación de impacto en protección de datos, información previa clara y posibilidad real de impugnar decisiones.
- Se apliquen principios de minimización, anonimización o agregación cuando la finalidad sea la mejora de calidad o la investigación, y no el control individual del trabajador.

2.4. Objección de conciencia e IA.

En el marco de la medicina digital, puede hablarse de objeción de conciencia tecnológica como el derecho del profesional sanitario a negarse, de forma motivada y documentada, a utilizar un sistema concreto de inteligencia artificial o de automatización clínica cuando considere que su uso compromete la seguridad, la dignidad o la adecuada atención del paciente.

No se trata de un rechazo genérico a la tecnología ni de una resistencia irracional al cambio, sino de una forma específica de objeción profesional que se sitúa en la intersección de:

- La autonomía profesional e independencia clínica del médico. ▪ El deber de no maleficencia y de protección de la integridad del paciente, tal como reconoce el Convenio de Oviedo (primacía del ser humano frente al interés de la ciencia o la sociedad).

- El marco regulatorio de la IA, que exige supervisión humana eficaz sobre los sistemas de alto riesgo (AI Act, art. 14 y obligaciones del desplegador o deployer).

Desde la tradición bioética, la objeción de conciencia se ha defendido como un mecanismo para preservar la integridad moral del profesional en contextos como el aborto o la eutanasia. La novedad en el contexto de la IA no es el "derecho a objetar" en sí, sino su objeto: ya no se dirige únicamente contra determinados actos clínicos, sino contra el uso de determinadas herramientas algorítmicas cuando éstas se perciben como inseguras, sesgadas o incompatibles con la lex artis.

En el contexto español, esta figura puede interpretarse como una extensión de la llamada "objeción de ciencia", ya introducida en el debate deontológico, entendida como el rechazo de una norma u orden por razones científicas o profesionales que se oponen a la libertad de método o de prescripción del médico.

La objeción tecnológica sería, en este sentido, la aplicación de esa objeción de ciencia a sistemas de IA y tecnologías automatizadas.

Aunque el término "objeción de conciencia tecnológica" no está todavía positivizado como tal, pueden observarse tendencias convergentes en distintos países de nuestro entorno como Francia, Alemania, Canada o EEUU.

Estas tendencias apuntan a un principio compartido: el médico no puede ser obligado a utilizar una IA que considere incompatible con la buena práctica y la seguridad del paciente, siempre que su negativa sea razonada, proporcionada y compatible con el acceso del paciente a la atención necesaria.

3. Bloque práctico:

A) Desde la perspectiva del médico individual, su nueva condición de **desplegador (*deployer)*** no debe entenderse solo como una carga añadida, sino también como un argumento para reivindicar derechos aseguradores: acceso a información técnica suficiente, formación acreditada, protocolos claros de uso y posibilidad de documentar adecuadamente la supervisión clínica. Sin estas condiciones, el profesional queda expuesto a un riesgo jurídico y asegurador desproporcionado.

En la práctica, esto puede traducirse en varios ajustes contractuales:

1. Cláusulas específicas sobre uso de IA. Muchas pólizas de RCP empezarán a incluir cláusulas que:

- Exijan que los sistemas de IA utilizados estén debidamente certificados/registrados y se empleen en el marco de la indicación autorizada.

- Subordinen la cobertura plena al cumplimiento razonable de las obligaciones del *deployer* (protocolos internos, supervisión humana, documentación de decisiones, conservación de logs).

- Delimiten qué se considera "error médico" y qué se considera "fallo del sistema" atribuible al proveedor, abriendo la puerta a acciones de repetición o coaseguro entre aseguradoras de profesionales, centros y fabricantes.

2. Reforzamiento del componente organizativo. La responsabilidad se desplaza parcialmente del profesional aislado a la organización desplegadora (servicio, hospital, red sanitaria). Esto hace más relevante:

- La existencia de pólizas colectivas de RCP que cubran tanto a la institución como a los profesionales, con límites y franquicias adaptados al nuevo riesgo tecnológico.

- La evaluación de la *gobernanza de la IA* como criterio técnico de suscripción: comités de IA, protocolos de validación clínica local, auditorías periódicas de rendimiento y sesgos, planes de contingencia y cese de uso ante incidencias. Las organizaciones con mejor gobernanza podrán negociar condiciones más favorables (primas, límites, exclusiones).

3. Documentación, trazabilidad y prueba pericial. La obligación del *deployer* de conservar logs y registrar las interacciones con el sistema tiene un efecto inmediato en el terreno probatorio:

- En un eventual procedimiento judicial, estos registros serán piezas clave para reconstruir si el médico actuó conforme a la lex artis y si supervisó adecuadamente la recomendación algorítmica.

- La ausencia de registros o la imposibilidad de reconstruir el flujo de decisión puede interpretarse como indicio de incumplimiento de las obligaciones del *deployer*, lo que tensiona la posición del asegurado frente a su compañía.

- A la vez, esta trazabilidad permite delimitar mejor la participación causal del sistema de IA, facilitando la pericia forense y el reparto de responsabilidad entre médico, centro y proveedor del sistema.

B) La **objeción de conciencia tecnológica** no debe confundirse con el rechazo genérico a la digitalización. Su espacio natural es la frontera entre el uso razonable y el uso imprudente de la IA. Algunos escenarios paradigmáticos:

3.1. Triaje automatizado con criterios injustificados:

- Un hospital implanta un sistema de triaje de urgencias que prioriza pacientes en función de variables sociodemográficas (código postal, aseguradora, nivel de renta indirecto) sin justificación clínica ni transparencia en el modelo.

- Un médico de urgencias detecta que, de manera sistemática, las personas mayores que viven solas quedan en categorías de menor prioridad, pese a cuadros clínicos graves. Decide **no usar** el algoritmo en su turno y documenta en la historia que aplica triaje clínico tradicional, alegando riesgo de discriminación.

- Aquí, la objeción se fundamenta en la **no maleficencia**, la igualdad de trato y la **prohibición de discriminación algorítmica** prevista en el AI Act.

3.2. Sistema de ayuda al diagnóstico "caja negra" con falsos negativos:

- Un servicio de radiología adopta un sistema de detección de nódulos pulmonares con alta sensibilidad declarada. Estudios internos revelan, sin embargo, un número relevante de falsos negativos y el proveedor se niega a compartir las bases de validación o criterios de funcionamiento.

- Un radiólogo puede alegar objeción de conciencia tecnológica y **exigir realizar su propia lectura sistemática**, utilizando la IA sólo de modo exploratorio o no utilizándola, hasta que se valide localmente el rendimiento.

3.3. Modelos no adaptados al contexto local:

- Un algoritmo de predicción quirúrgica se ha entrenado en hospitales de alto volumen con pacientes de bajo riesgo y tecnología distinta. En un hospital comarcal, los resultados muestran discrepancias importantes (p.ej., infravalora complicaciones en población anciana con comorbilidades).

- El cirujano puede objetar el uso de ese modelo para la planificación en su centro hasta que se realice una **recalibración local** y se establezca una evidencia de validez externa, invocando la obligación de evitar prácticas clínicas no validadas.

3.4. Automatización excesiva en salud mental:

- Una organización implanta un chatbot para el seguimiento de pacientes con ideación suicida, reduciendo significativamente las entrevistas presenciales.

- El psiquiatra responsable puede oponerse a delegar en la IA la valoración de riesgo agudo, manteniendo la exigencia de **contacto humano directo** en pacientes de alto riesgo, y alegando que el modelo no garantiza la contención emocional ni la detección fina de señales de peligro.

En todos estos casos, la objeción no supone un rechazo absoluto de la IA como concepto, sino de su uso concreto en condiciones que el médico considera profesionalmente inaceptables.

4. Bibliografía

1. Trujillo Ruiz JA. Guía básica de Inteligencia Médica Artificial. Independently published; 2023. ISBN: 9798862907230. Amazon España+1

2. Trujillo Ruiz JA. Inteligencia Artificial y Derechos de los Pacientes: el equilibrio necesario. Independently published; 2025. ISBN: 9798313198897. Amazon+1

3. Trujillo Ruiz JA. *Nativos inteligentes. La generación de la IA*. Málaga: Fundación Unicaja; 2026. https://libreria.fundacionunicaja.com/inicio/1769-nativos-inteligentes-la generacion-de-la-ia.html

4. Risse M. *The Fourth Generation of Human Rights: Epistemic Rights in the Digital Age*. Harvard Kennedy School; 2021. appext.hks.harvard.edu

5. Razmetaeva Y. The concept of human rights in the digital era and bio information rights as fourth generation rights. *Asian J Legal Ethics*. 2022. ajee-journal.com

6. International IDEA. *Rights in the Digital Age*. Estocolmo: International IDEA; 2025. idea.int

7. Funer F. Physician's autonomy in the face of AI support. *BMC Med Ethics*. 2024. PMC

8. Jones C, et al. Artificial intelligence and clinical decision support: clinicians' responsibilities in the age of AI. *Med Law Rev*. 2023;31(4):501-524. OUP Academic

9. Größer J, et al. Studying the potential effects of artificial intelligence on physician autonomy and decision-making. *Healthcare Analytics*. 2025. ScienceDirect

10. Šustek P, Šolc M. Civil liability for artificial intelligence in medicine: is there a need for a new paradigm? *The Lawyer Quarterly*. 2025;34(3):464- 484. tlq.ilaw.cas.cz

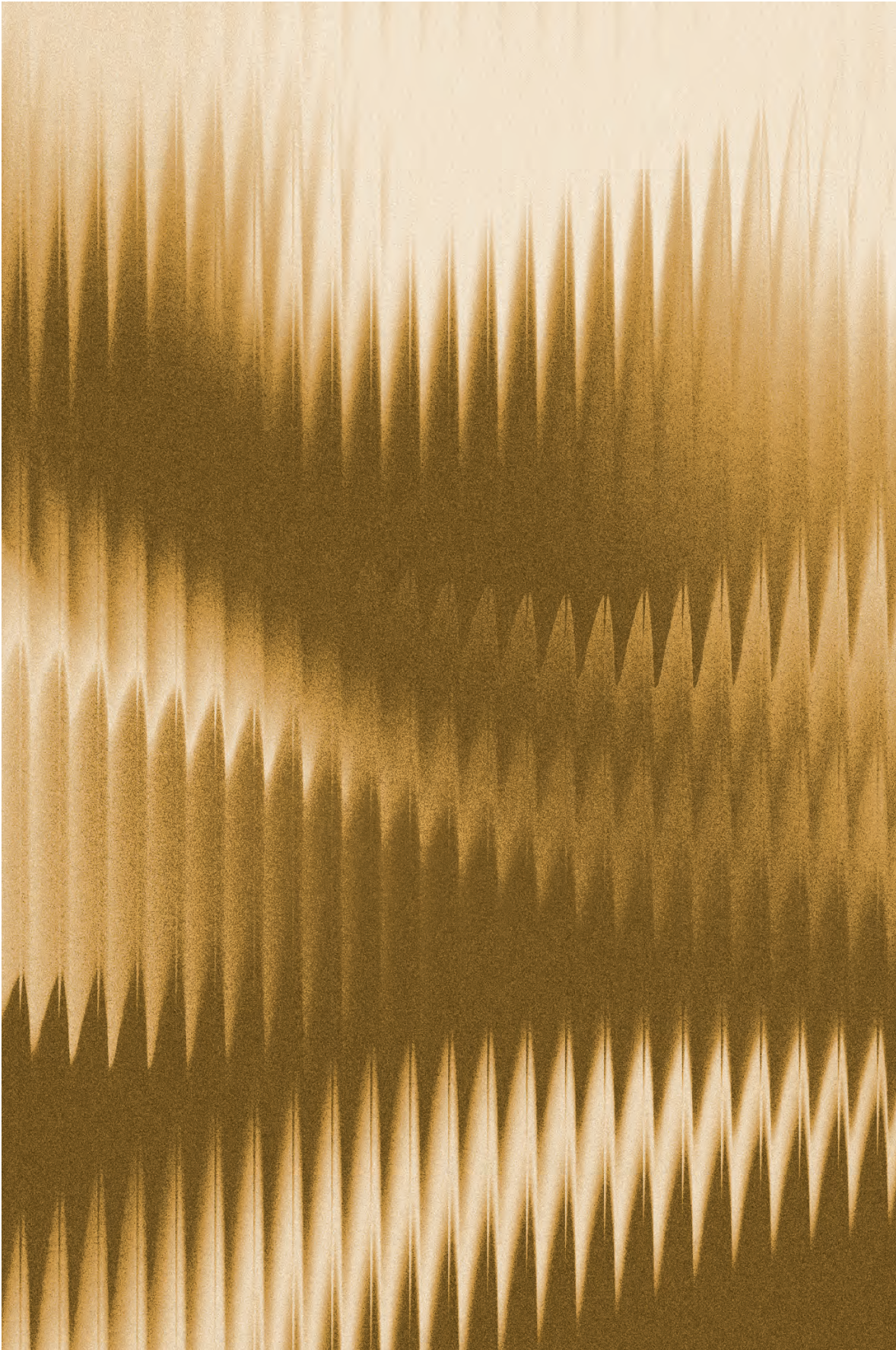
11. World Medical Association. *WMA Statement on Artificial and Augmented Intelligence in Medical Care*, revisada 2024/2025. Asociación Médica Mundial

12. Unión Europea. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (AI Act). DOUE L 2024/1689, 12.7.2024. EUR-Lex+1

13. Reglamento (UE) 2025/327 del Parlamento Europeo y del Consejo, de 11 de febrero de 2025, relativo al Espacio Europeo de Datos de Salud (EEDS). Texto en EUR-Lex (versión en español): https://eur-lex.europa.eu/legalcontent/ES/ALL/?uri=CELEX:32025R0327 EUR-Lex

14. European Commission. *Proposal for a Regulation of the European Parliament and of the Council amending Regulations (EU) 2024/1689 and (EU) 2018/1139 as regards the simplification of the implementation of harmonised rules on artificial intelligence (Digital Omnibus on AI), COM(2025) 836 final.* Brussels, 19 Nov 2025. Disponible en: https://eur-lex.europa.eu/legalcontent/ES/ALL/?uri=CELEX-:52025PC0836 EUR-Lex

15. Osborne Clarke. "EU AI Act's 'deployers' definition has wide-ranging significance for life sciences". 2024. osborneclarke.com



IV

Entre Datos y Diagnósticos: IA Médica con Garantías

Dra. Cristina Gil Membrado
Catedrática de Derecho Civil. Universidad de las Islas Baleares. Investigadora principal del Proyecto El Derecho ante la Salud Digital, Personalizada y Robótica (SALUDPYR). Ministerio de Ciencia, Innovación y Universidades

1. Palabras claves:

Ley de IA, Protección de datos personales, EEDS, Uso responsable de la IA, Derechos de los pacientes.

2. Bloque teórico:

2.1. ¿Por qué hablar de Protección de Datos Personales e IA?

La IA está integrada en:

- Sistemas de apoyo diagnóstico.
- Estratificación del riesgo.
- Priorización asistencial.
- Interpretación de imágenes.
- Modelos predictivos clínicos.

La IA médica funciona con datos de salud, a menudo personales, porque hacen referencia directa o indirectamente a una persona. Estos datos están especialmente protegidos porque son muy sensibles.

Abordamos:

- Las exigencias del Reglamento Europeo General de Protección de Datos (RGPD).
- Las exigencias del Reglamento Europeo de Inteligencia Artificial (Ley de IA).
- Las implicaciones del Espacio Europeo de Datos Sanitarios (EEDS).
- Qué debe saber el médico cuando utiliza la HC.

2.2. ¿Cómo funciona la IA médica?

Para el entrenamiento y uso de un sistema de IA es necesario tratar datos. La IA no razona por sí sola. Compara al paciente actual con patrones aprendidos de otros pacientes. Para ello hay que nutrir a la IA con datos relativos a historias clínicas previas, imágenes radiológicas etiquetadas o no, analíticas con diagnóstico final conocido, evoluciones clínicas, desenlaces y muchos otros.

Durante todo el ciclo de vida de un sistema de IA se utilizan datos, en muchas ocasiones personales, en cuyo caso será de aplicación el RGPD.

Entrenamiento
Aprende con datos históricos de pacientes

Validación
Se ajusta y se corrigen los errores

Prueba
Se comprueba su seguridad y precisión

Uso Clínico
Recibe datos reales del paciente

Mejora Continua
Aprende de los nuevos casos confirmados

2.3. Los datos: el combustible de la IA

Ejemplo: IA diseñada para detectar el cáncer de pulmón en radiografías

- El sistema se entrena con miles de radiografías diagnosticadas por médicos.
- Antes de usarla con pacientes reales, se valida con nuevas radiografías cuyos diagnósticos son conocidos -pero no por la IA-.
- Se compara lo que predice la IA con el diagnóstico médico real.

- Si la IA alcanza una alta precisión y una baja tasa de errores puede aprobarse para pruebas clínicas controladas.
- Tras superar las fases anteriores la IA se utilizará como apoyo a la decisión médica -nunca en sustitución.

Sin datos no podríamos tener garantía de que la IA es segura y efectiva.

2.4. RGPD vs Ley de IA ¿Qué regula cada norma?

Aspecto	RGPD	Ley de IA
Enfoque principal	Protección de datos	Seguridad, ética y control de la IA
Privacidad del paciente	Protege los datos personales del paciente	Exige evaluar los riesgos que la IA puede generar para la salud
Consentimiento	Requiere consentimiento informado para el tratamiento de datos	No se centra en el consentimiento, sino en las condiciones de uso seguro del sistema
Minimización de datos	Obliga a usar solo los datos necesarios	Puede afectar al diseño y entrenamiento del sistema, pero no es su eje principal
Derechos del usuario	Garantiza acceso, rectificación y eliminación de datos	Refuerza la necesidad de transparencia y explicabilidad en sistemas de IA
Transparencia	Informa sobre el uso de los datos personales	Exige explicar cómo y por qué la IA toma determinadas decisiones
Supervisión humana	No es su núcleo central	Exige control y supervisión por profesionales humanos, especialmente en salud
Clasificación del riesgo	No clasifica sistemas por niveles de riesgo	Regula los sistemas de IA según su nivel de riesgo
Objetivo final	Proteger los derechos y datos de las personas	Asegurar un uso seguro, ético y fiable de la IA en medicina

Ideas clave:

- El RGPD protege los datos del paciente y su buen uso por el sistema de IA.
- La Ley de IA controla la seguridad del sistema.
- Ambas normas se complementan.

2.5. ¿Cuándo se pueden utilizar datos personales de salud por sistemas de IA?

El tratamiento de datos personales de salud está prohibido por defecto, salvo que exista una base legal reforzada.

Base legal	Descripción / aplicación en IA médica
Consentimiento explícito	Aceptación directa del paciente para el tratamiento de sus datos de salud.
Interés público	Tratamiento justificado por razones de salud pública o investigación médica de interés general.
Interés vital	Uso de datos en situaciones de emergencia para proteger la vida del paciente.
Obligación legal	Tratamiento necesario para cumplir leyes que obligan a tratar datos de salud.
Interés legítimo	Posible en supuestos muy delimitados, con justificación suficiente y bajo controles estrictos.

2.6. Decisiones automatizadas: límite crítico.

El paciente tiene derecho a no ser objeto de decisiones exclusivamente automatizadas que produzcan en él efectos significativos -cualquier decisión sobre la salud es un caso paradigmático-.

Ejemplo: Sistema de IA de triaje hospitalario que decide si un paciente puede acceder o no a una prueba de diagnóstico urgente.

- En un formulario digital se introducen datos (dolor, fiebre, edad, antecedentes, entre otros).
- La IA analiza los datos y decide automáticamente.
  - No cumple criterios: prueba denegada.
  - Cumple criterios: prueba autorizada.
- Todo ello sin intervención de un profesional sanitario.

Aunque existiera consentimiento explícito del paciente, este tiene derecho a ser informado, a obtener la intervención significativa de un profesional sanitario y a rebatir la decisión.

Elemento	Contenido
Marco normativo	Artículo 22 del RGPD aplicado a la automatización de decisiones médicas por IA.
Principio general	Se limitan las decisiones basadas únicamente en IA cuando pueden tener un impacto significativo en la salud del paciente.
Finalidad de la norma	Proteger los derechos de los pacientes mediante condiciones estrictas para el uso de decisiones automatizadas.
Restricción principal	Está prohibido tomar decisiones médicas solo por IA, sin supervisión humana, cuando puedan producir efectos legales o impactos significativos sobre el paciente.
Medida de protección 1	Realización de evaluaciones de impacto sobre los riesgos para la privacidad.
Medida de protección 2	Información detallada y clara al paciente sobre el tratamiento automatizado.
Medida de protección 3	Derecho del paciente a impugnar y corregir decisiones automáticas.
Supervisión humana 1	Intervención humana para verificar y validar los resultados emitidos por la IA.
Supervisión humana 2	Control periódico para garantizar objetividad y equidad.
Conclusión	El RGPD exige que la automatización de decisiones médicas con IA sea transparente, equitativa y revisada por humanos.

Al ejemplo anterior habría que añadir la revisión de la recomendación de la IA por el profesional sanitario competente para confirmarla, modificarla o rechazarla. La intervención debe ser efectiva, no meramente simbólica. A continuación, nos centramos en ello.

2.7. La supervisión humana del profesional sanitario es imprescindible

Además de lo apuntado por el RGPD, la Ley de IA establece la obligación de garantizar la intervención humana significativa en la IA de alto riesgo -y la de salud lo es-. Exige dotar a los sistemas de IA de herramientas de interfaz humano-máquina adecuadas.

Ejemplo: Sistema de IA que prioriza radiografías de tórax en un hospital para detectar casos graves de neumonía.

- **El radiólogo tiene una interfaz humano-máquina que muestra:**
  - El nivel de riesgo asignado por la IA (alto-medio-bajo).
  - Las zonas de la imagen que han influido en la predicción (mapa de calor).
  - El grado de confianza del sistema.
- **El radiólogo puede:**
  - Confirmar la prioridad.
  - Cambiarla manualmente.
  - Ignorar la recomendación de la IA.
- **El sistema permite;**
  - Pausar o desactivar la IA en cualquier momento.
  - Registrar las decisiones humanas para auditorías.

- **En este caso la vigilancia humana de la IA es efectiva porque el médico:**
  - Entiende la recomendación.
  - Puede intervenir antes de que produzca efectos clínicos.
  - Tiene el control real sobre el resultado.

Elemento	Contenido
Concepto principal	Supervisión humana significativa en sistemas de IA médica.
Finalidad	Garantizar la fiabilidad, la seguridad y la ética en el uso de la IA médica.
Vigilancia constante	Supervisar de forma continua las decisiones adoptadas por la IA.
Intervención experta	Permitir la intervención médica cuando sea necesario para revisar, corregir o complementar la actuación del sistema.
Responsabilidad ética	Asegurar que las decisiones se tomen con criterios éticos y responsabilidad durante todo el proceso.
Mensaje clave	Las personas expertas deben mantener el control sobre la IA médica, corregir errores y asumir decisiones éticas.

2.8. Sesgos en IA: un riesgo clínico real

Los sesgos en los datos de salud aparecen cuando la información con la que se entrena o se evalúa una IA no representa a toda la población o contiene errores sistemáticos.

Ejemplo: si una IA que detecta cáncer de piel es entrenada especialmente con imágenes de personas de piel clara funcionará bien en ese grupo, pero fallará más en personas con piel oscura, porque los datos adolecen de calidad y de representatividad.

Los sesgos pueden provocar más errores en ciertos grupos (edad, sexo, origen étnico), infradiagnósticos o sobrediagnósticos en poblaciones concretas y desigualdades en la atención sanitaria.

La IA no es neutral por sí misma y refleja los patrones y desequilibrios presentes en los datos por lo que se puede afirmar que en un sistema de IA "basura que entra, basura que sale". Es por ello por lo que resulta fundamental usar datos diversos

y representativos, detectar y corregir sesgos antes del uso clínico y mantener la supervisión humana.

Así, los sesgos en datos de salud pueden traducirse en decisiones clínicas menos justas o seguras, y por eso su identificación y corrección es esencial en cualquier sistema de IA sanitaria.

2.9. Derechos del paciente sobre datos personales utilizados por la IA

El RGPD y la Ley de IA contemplan:

- Consentimiento: el paciente debe saber qué hará la IA con sus datos y aceptar el tratamiento de forma consciente y voluntaria.
- Información/Transparencia: implica explicar al paciente, de modo claro y comprensible, cómo y por qué la IA usa sus datos personales (qué datos va a usar; para qué se usan y cómo funciona la IA de modo general; quién los usa, si se comparten y a quién; cuánto tiempo se guardan y los derechos que ostenta).
- Derecho de acceso: conocer si sus datos personales están siendo tratados por un sistema de IA, qué datos, con qué finalidad y cómo influyen en el funcionamiento del sistema. No obliga a explicar el algoritmo con detalle, pero sí los efectos en la atención sanitaria.
- Derecho de rectificación: solicitar la corrección de datos inexactos o incompletos, para evitar que la IA genere resultados erróneos basados en información incorrecta.
- Supresión-Derecho al olvido: pedir la eliminación de los datos personales que no ya no sean necesarios, el tratamiento sea ilícito o se retire el consentimiento -a salvo de las obligaciones de conservación establecidas por la Ley 41/2002-.
- Limitación del tratamiento: exigir que los datos no se utilicen por la IA temporalmente mientras se verifica su exactitud o la licitud del tratamiento.

- Oposición: oponerse al tratamiento de los datos personales especialmente cuando la IA se utilice con fines distintos de la asistencia sanitaria directa.
- Seguridad: el paciente tiene derecho a que sus datos estén protegidos y se usen de forma segura. Ello implica proteger los datos con medidas técnicas, como el cifrado o el control de accesos; limitar quién puede acceder a la información; evitar fugas robos o modificaciones de los datos y revisar y actualizar la seguridad de forma continua.

Derecho del paciente	Contenido / alcance
Consentimiento	Permite un consentimiento libre, informado y claro para el uso de datos personales.
Información y transparencia	Garantiza conocer qué datos se recogen, con qué finalidad y quién los utiliza.
Acceso	Permite saber qué datos se están usando y consultar esa información.
Rectificación	Reconoce el derecho a corregir datos incorrectos o inexactos.
Supresión / derecho al olvido	Permite solicitar la eliminación de datos cuando ya no sean necesarios.
Limitación del tratamiento	Permite restringir el uso de los datos en determinadas circunstancias.
Oposición	Permite oponerse al tratamiento o uso de los datos.
Seguridad	Exige proteger los datos de forma continua mediante medidas adecuadas.

2.10. Uno para todos y todos para uno: el Espacio Europeo de Datos de Salud (EEDS/EHDS)

El EEDS permite, como uso primario, el acceso transfronterizo de los ciudadanos de la UE a sus datos de salud (historias clínicas, informes, recetas electrónicas, entre otros) y al médico el acceso a la información de salud compartida.

Además, prevé la reutilización de estos datos con fines secundarios en beneficio de la sociedad, como la investigación sanitaria, la innovación, la formulación de políticas públicas o la regulación.

Ejemplo: necesidad de entrenar un modelo de IA para predicción de ingresos hospitalarios.

- Los historiales clínicos electrónicos con fines de atención sanitaria se pueden utilizar, con finalidad secundaria, para entrenar un sistema de IA.
- Se solicitará a la autoridad nacional de acceso a los datos el uso para el entrenamiento y la evaluación del algoritmo.
- Antes del uso se eliminarán los identificadores directos (como nombre y DNI), los datos se anonimizarán o se seudonimizarán y se limitarán a las variables necesarias.
- Se utilizarán en un entorno seguro sin poder descargarlos, únicamente con la finalidad de entrenar al modelo.
- En el EEDS, los datos de salud recogidos para el tratamiento de los pacientes se reutilizarán, de modo protegido y controlado, dentro de un entorno seguro para entrenar algoritmos de IA con datos de calidad.
- El paciente puede oponerse a la utilización de sus datos para estos fines mediante un procedimiento *opt-out* (de exclusión voluntaria).

Elemento	Contenido
Marco general	Espacio Europeo de Datos de Salud (EHDS), iniciativa de la Unión Europea para facilitar el intercambio seguro de datos de salud en la UE.
Objetivo 1	Favorecer el acceso transfronterizo a la información médica dentro de la Unión Europea.
Objetivo 2	Impulsar la investigación médica y el desarrollo de tecnologías digitales en salud.
Objetivo 3	Reforzar la seguridad y la privacidad mediante normas estrictas de protección de datos.
Beneficio 1	Mejor atención médica gracias a un acceso más rápido a los datos de salud.
Beneficio 2	Fomento de la innovación en tratamientos y tecnologías médicas.
Beneficio 3	Mayor protección de la privacidad de los pacientes.
Mensaje clave	El EHDS busca mejorar la atención médica y la investigación en salud en Europa garantizando la seguridad y la privacidad de los datos.

2.11. A modo de conclusión

- **El modelo normativo europeo combina:**
  - Protección de los datos (RGPD).
  - Seguridad de los sistemas de IA (Ley de IA).
  - Innovación regulada (EEDS).
- **La IA médica puede mejorar:**
  - Precisión diagnóstica.
  - Priorización clínica.
  - Seguridad del paciente.
- **Pero solo si se integra con:**
  - Supervisión humana profesional real.
  - Gobernanza de datos.
  - Transparencia.
  - Respeto a los derechos y a los principios de la bioética.

El mensaje resumen para el médico que utiliza la IA es:

- La IA no sustituye al médico.
- El médico supervisa la IA.
- La confianza y la seguridad del paciente depende de ello.

3. Bloque práctico

3.1. ¿Qué debe saber el médico cuando utiliza la IA?

1. ¿Es un sistema de alto riesgo?
2. ¿Está validado clínicamente?
3. ¿Conozco sus limitaciones?
4. ¿Sé con qué población fue entrenado?
5. ¿Estoy supervisando activamente el resultado?
6. ¿Los datos de entrada son correctos?
7. ¿Puedo corregir o ignorar la recomendación?
8. ¿Existe trazabilidad?
9. ¿Está integrada en un marco legal válido?
10. ¿Hay protocolo de incidentes?

3.2. Decálogo para el médico que utiliza la IA

1. La IA es apoyo, no sustituto.
2. Mantenga el juicio clínico.
3. Revise siempre los datos de entrada.
4. No delegue decisiones críticas sin supervisión.
5. Conozca las limitaciones del modelo
6. Identifique posibles sesgos
7. Informe adecuadamente al paciente cuando proceda
8. Documente el uso de la IA
9. Notifique incidentes.
10. Fomente cultura de uso responsable.

4. Bibliografía

1. Trujillo Ruiz, J.A., Gil Membrado, C. (dirs.), Libro Blanco de IA en Medicina, Dykinson, 2026.

2. Gil Membrado, C., "IA y salud cardiovascular en la mujer: una perspectiva biológica y de género", en Gil Membrado, C., Luquin Bergareche, R. (dirs.), *Medicina personalizada genética y criogenización*, Dykinson, 2025.

3. Gil Membrado, C., "La Ley de IA: entre la salud y el Derecho en una nueva era. Un acompasamiento necesario", *Revista Jurídica de las Illes Balears*, núm. 26, 18 de diciembre de 2024.

4. Gil Membrado, C., "Daños producidos por la IA: La opacidad del algoritmo y el efecto caja negra" en *Derecho de contratos, responsabilidad extracontractual e inteligencia artificial*, Asociación de profesores de Derecho Civil, Aranzadi La Ley, 2024.

5. Gil Membrado, C., "Diabetes y Espacio Europeo de Datos Sanitarios", en Gil Membrado, C., Luquin Bergareche, R. (dirs), *Salud Digital. Aplicaciones Móviles, Telemedicina y Chatbots*, Dykinson, 2024.

6. Gil Membrado, C., *Riesgos del uso de algoritmos en el diagnóstico y en la investigación biomédica*, VIII Premio Nacional de Derecho Sanitario, Reus, 2023.

7. Gil Membrado, C., "En el horizonte: la Directiva de responsabilidad extracontractual en materia de IA", en Gil Membrado, C., Luquin Bergareche, R. (dirs), *Derecho y medicina: desafíos tecnológicos y científicos*, Dykinson, 2023.

8. Gil Membrado, C. "La apertura del código fuente y la transparencia en la contratación de servicios. Equidad y transparencia en la prestación de servicios", en Cobas Cobiella, M.E., Guillén Catalán, R., (dirs.), *Equidad y transparencia en la prestación de servicios*, Dykinson, 2023.

9. Gil Membrado, C., "Un régimen europeo de responsabilidad civil para el usuario de la inteligencia artificial", *Direito do consumidor no cenário iberoamericano*, Foco, Brasil, 2023.

10. Gil Membrado, C., "Una nueva era: hacia el robot sanitario autónomo y su encaje en el derecho", en Gil Membrado, C., Luquin Bergareche, R. (dirs.), *M-Health, genética, robótica y criogenización*, Dykinson, 2022.

11. Gil Membrado, C., "Telemedicina y aplicaciones móviles. la m-Health y la seguridad de los datos personales", en Martínez Valencoso, M.L., Sancho López, M. (dirs.), *Protección jurídica de la privacidad: inteligencia artificial, salud y contratación,* Aranzadi La Ley, 2022.

5. Normativa utilizada

1. Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). «DOUE» núm. 119, de 4 de mayo de 2016.

2. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) núm. 300/2008, (UE) núm. 167/2013, (UE) núm. 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial). «DOUE» núm. 1689, de 12 de julio de 2024.

3. Reglamento (UE) 2025/327 del Parlamento Europeo y del Consejo, de 11 de febrero de 2025, relativo al Espacio Europeo de Datos de Salud, y por el que se modifican la Directiva 2011/24/UE y el Reglamento (UE) 2024/2847. «DOUE» núm. 327, de 5 de marzo de 2025.

V

Deontología y ética médica en tiempos de inteligencia artificial

Dr. José María Domínguez Roldán
Presidente de la Comisión de Ética y Deontología Médica de la Organización Médica Colegial. Profesor de la Facultad de Ciencias de la Salud y de la Vida. Universidad CEU Fernando III, CEU Universities

1. Palabras claves

Inteligencia artificial; ética médica; deontología; relación clínica; autonomía; consentimiento informado; secreto profesional; trazabilidad; transparencia; explicabilidad; sesgo algorítmico; justicia; equidad; responsabilidad profesional; gobernanza; seguridad del paciente; datos sanitarios; evaluación clínica; derechos fundamentales; humanismo.

2. Bloque teórico

2.1. Planteamiento: cuando la técnica altera la moral del acto médico

La medicina siempre ha incluido "la técnica" en un sentido amplio: conocimiento aplicado con instrumentos (desde el estetoscopio al laboratorio) y con formas institucionales (hospitales, protocolos, guías). Pero la IA introduce un cambio cualitativo: no es sólo una herramienta, sino un sistema de inferencia que produce recomendaciones y clasificaciones (riesgo, diagnóstico, priorización, pronóstico) a partir de correlaciones aprendidas en datos. Esto desplaza la epistemología clínica desde el razonamiento fisiopatológico explícito hacia patrones estadísticos frecuentemente opacos, y con ello modifica la aproximación humanística del acto médico: ¿qué significa "saber", "explicar", "responder", "decidir" cuando una parte decisiva del proceso es computacional y no intuitivamente auditable?

En ética clínica, el centro de gravedad no está en la potencia de cálculo, sino en la orientación del poder técnico hacia el bien humano. El progreso biomédico ha sido históricamente ambivalente: cura y daña, libera y esclaviza, iguala y discrimina. La IA agudiza esta ambivalencia porque opera a escala (miles de decisiones por minuto), se difunde rápidamente (software replicable) y tiende a naturalizar sus salidas como "objetivas". Por eso, el núcleo ético exige resistirse a dos extremos igualmente inaceptables: (a) la tecnofobia, que se opone a toda innovación por temor; y (b) el *tecnosolucionismo*, que presupone que todo problema clínico y organizativo es reducible a una predicción algorítmica.

V Deontología y ética médica en tiempos de inteligencia artificial

Dr. José María Domínguez Roldán

La deontología no es un freno externo a la innovación, sino una disciplina interna de la profesión. En el contexto de la IA, la deontología es, además, una forma de gobernanza: establece obligaciones exigibles al médico y, por extensión, a las instituciones que lo emplean, frente a sistemas técnicos cuya lógica dominante puede ser económica, industrial o burocrática. En términos clásicos, la medicina no es sólo *episteme* (ciencia), ni sólo *techne* (técnica): es también *phronesis* (prudencia), es decir, deliberación práctica, que integra evidencia, singularidad del paciente, valores, y consecuencias previsibles.

De aquí deriva una tesis que vertebra este documento: la IA puede ser muy útil para la medicina, pero no puede reemplazar su núcleo moral. Y ese núcleo moral se expresa en cuatro afirmaciones operativas:

1. la responsabilidad clínica es indelegable;
2. El paciente no es un conjunto de datos, sino una persona con dignidad;
3. La relación clínica es una relación fiduciaria que no puede reducirse a interacción con sistemas;
4. La justicia sanitaria exige una vigilancia activa contra sesgos, discriminaciones y exclusiones.

2.2. El marco deontológico español: continuidad y actualización

El médico español no parte de un vacío normativo. La profesión dispone de un cuerpo deontológico que, aun siendo anterior a la explosión reciente de la IA generativa, contiene criterios robustos para ordenar tecnológicamente la asistencia. La relevancia del Código de Deontología Médica no reside sólo en "mencionar" la IA, sino en articular un principio decisivo: la tecnología debe estar al servicio de fines éticos, y el médico conserva el deber de control finalista, prudente y responsable.

I. Principio de primacía del paciente y finalidad moral de la tecnología

La primera exigencia deontológica, de resonancia hipocrática, es la orientación al bien del paciente. En IA, esto obliga a distinguir entre la utilidad clínica real y la eficacia aparente. Un sistema puede ser preciso en una métrica y, sin embargo, ser clínicamente inútil o dañino si no mejora resultados relevantes, si introduce sesgos, si aumenta pruebas innecesarias, o si desplaza recursos de manera injusta. El criterio no puede ser "lo que la máquina acierta", sino "lo que la medicina mejora" en salud, sufrimiento, discapacidad, vida digna y equidad.

A esta primacía se opone una tentación contemporánea: convertir al paciente en un medio para alimentar sistemas (la extracción masiva de datos) o para justificar inversiones tecnológicas (proyectos de innovación cuya métrica principal es el despliegue, no el beneficio clínico). La deontología exige invertir el orden: los datos y los algoritmos son medios; la persona es fin.

II. Independencia clínica y prevención de la "obediencia algorítmica"

El riesgo de la IA no es sólo el error puntual, sino la obediencia algorítmica: la tendencia humana a aceptar recomendaciones computacionales como si fuesen neutrales y superiores, especialmente en entornos presionados por tiempo, por gestión, o por miedo a la responsabilidad. Esta obediencia se refuerza cuando la institución convierte la salida del algoritmo en un estándar de calidad o en un requisito administrativo.

Frente a ello, la independencia clínica debe reinterpretarse: no es independencia "contra" la evidencia ni "contra" la estandarización razonable; es independencia contra la sustitución del juicio profesional por una autoridad técnica no moral. La decisión clínica responsable integra guías, datos, preferencias del paciente y contexto; pero no abdica de una decisión automatizada. La IA, por tanto, debe considerarse un instrumento de ayuda cuyo valor depende de su calidad, pertinencia y validez local, así como del control humano.

III. Secreto profesional, confidencialidad y nuevos vectores de vulneración

La confidencialidad es constitutiva de la relación clínica. La IA amplía el perímetro del secreto profesional: ya no se limita al médico y su institución, sino a cadenas complejas de tratamiento de datos (proveedores, subcontratas, nubes, auditorías, telemetría, mantenimiento, reentrenamientos). En este nuevo escenario, la confidencialidad se convierte en un problema de arquitectura técnica y contractual, no solo de virtud profesional.

Deontológicamente, el médico debe comprender, al menos en su lógica, las rutas por las que los datos circulan, y debe poder exigir garantías: minimización, control de acceso, registro de usos, evaluación de reidentificación, y límites a los usos secundarios. El consentimiento informado en la era de la IA no puede convertirse en un "consentimiento por fatiga" (formularios largos e incomprensibles) ni en un formalismo. Debe ser una comunicación moralmente significativa: qué se hace, por qué, con qué riesgos, y qué alternativas hay.

IV. La deontología como criterio de gobernanza institucional

Un error común es creer que la deontología "habla sólo al médico" y no a la organización. En realidad, toda deontología profesional madura se convierte en criterio para gobernar instituciones sanitarias: compra pública, contratación, evaluación de tecnologías y modelos de calidad. Si una institución impone herramientas opacas sin validación, erosiona la independencia clínica y empuja al médico a un conflicto: obedecer al sistema o cumplir su deber con el paciente. Por tanto, el núcleo ético en un manual de buenas prácticas debe

ser también un núcleo de diseño institucional: comités, circuitos de evaluación, auditorías, y mecanismos de escalado de incidentes.

b. Perspectiva actual: derechos fundamentales y ética de la alta escala

Actualmente se considera que la IA no es solo una cuestión de producción de información, sino también un asunto de derechos fundamentales. Esta intuición es coherente con la tradición europea de dignidad humana y con instrumentos jurídicos que, desde la bioética hasta la protección de datos, buscan evitar que el cuerpo y la vida se conviertan en materia prima de procesos despersonalizados.

I. Dignidad, integridad y no instrumentalización

El paciente es portador de dignidad; esto implica que no puede ser tratado como un objeto de optimización. En medicina, optimizar no es maximizar una variable abstracta (eficiencia, flujo, coste), sino realizar una acción prudente en condiciones de incertidumbre y vulnerabilidad. En IA, la instrumentalización puede presentarse de manera sutil: sistemas que orientan al paciente hacia opciones preferidas por la institución; "chatbots" que sustituyen la escucha por guiones prefabricados; o sistemas que penalizan socialmente al enfermo crónico. La ética exige límites: ni manipulación ni explotación de vulnerabilidades ni discriminación automática.

II. Riesgo, proporcionalidad y sistemas de alto impacto

Un principio ético esencial es la proporcionalidad: cuanto mayor sea el riesgo de una tecnología, mayores deben ser los requisitos de seguridad, transparencia y control. En salud, el riesgo tiene particularidades: no es sólo el daño físico, sino también el daño moral (estigmatización), el social (exclusión), el jurídico (pérdida de derechos) y el epistémico (desplazamiento del juicio). Por eso, la IA clínica no puede tratarse como un simple "software de oficina" ni evaluarse únicamente por el rendimiento interno del modelo. La evaluación ética exige considerar: propósito, población, entorno, mecanismos de fallo, y consecuencias distributivas.

Protección de datos: más allá de la privacidad como secreto profesional

La protección de datos en salud no es meramente confidencialidad; también es justicia. Quien controla los datos controla la capacidad de clasificar, priorizar y vigilar. Un ecosistema de IA puede convertirse en una infraestructura permanente de vigilancia sanitaria si no se establecen límites. Por ello, la gobernanza de datos debe incluir: finalidad específica, minimización, transparencia en usos secundarios, control de cesiones, y garantías de que el paciente no será penalizado por su patobiografía o por correlaciones inferidas.

En este punto, la ética requiere un esfuerzo intelectual adicional: la IA permite inferir rasgos sensibles incluso cuando no se recopilan explícitamente. Es decir, la privacidad no se viola sólo por "filtración", sino también por inferencia. La clínica debe asumir que un sistema puede inferir el riesgo de depresión, la vulnerabilidad social o la probabilidad de consumo de sustancias a partir de patrones indirectos. Esto obliga a interrogar qué inferencias son legítimas y con qué garantías.

c. Principios clásicos, problemas nuevos: relectura de la bioética clínica

Los principios de la bioética clínica (autonomía, beneficencia, no maleficencia, justicia) siguen siendo útiles, pero requieren una traducción operativa para la era de la IA.

I. Autonomía: consentimiento informado, explicabilidad.

La autonomía no se reduce a firmar un documento de consentimiento. En la era de la IA, la autonomía exige la capacidad de comprender, deliberar y decidir.

La explicabilidad de los sistemas es imprescindible dado que muchos sistemas de IA han sido desarrollados mediante una "caja negra". ¿Qué significa explicar al paciente por qué se le recomienda una intervención, por qué se le prioriza

o por qué se le deniega una prueba? La obligación de información no exige necesariamente revelar el código, pero sí exige una explicación clínica inteligible: variables relevantes, incertidumbre, límites, alternativa de segunda opinión y posibilidad de revisión por un profesional.

Cuando el paciente interactúa con un sistema conversacional digital, la autonomía puede erosionarse si este dirige la conversación hacia opciones cerradas y no detecta matices existenciales (como miedo, violencia o precariedad). El consentimiento informado debe adaptarse: no solo informar "se usa IA", sino también aclarar el rol de la IA en la toma de decisiones y el margen de intervención humana.

Además, la autonomía se vincula con la confianza. Si el paciente percibe que "una máquina decide", puede sentir desposesión y ello afecta a la adherencia y a la alianza terapéutica. La buena práctica exige comunicar con honestidad: la IA puede ayudar, pero no reemplaza el compromiso del médico con la persona.

II. Beneficencia: utilidad clínica, evidencia, validez local

La beneficencia obliga a demostrar beneficio. En IA esto es complejo porque el rendimiento del modelo no equivale a un beneficio clínico. Un algoritmo puede predecir la mortalidad con alta precisión, pero no mejorarla si no hay intervención efectiva o si su uso altera conductas clínicas de forma inadecuada.

La beneficencia exige, por tanto:

- Validación clínica, no solo técnica (con impacto en las decisiones y los resultados).
- Validez local: un modelo entrenado en un contexto puede fallar en otro por diferencias demográficas, de práctica clínica, de codificación, o de prevalencia.
- Monitorización continua: la medicina cambia (nuevos tratamientos, nuevas pandemias, nuevas prácticas), y los modelos envejecen.

El médico debe interiorizar que la IA es un "dispositivo epistémico" que requiere mantenimiento moral: revisión de supuestos, reevaluación de desempeño, y prudencia frente a la innovación.

III. No maleficencia: seguridad del paciente y fallos sistemáticos

Los posibles problemas derivados del uso de la IA en medicina no siempre son dramáticos ni inmediatos. Podemos considerar tres tipos de daño que son especialmente relevantes:

1. **Daño por error**: falsos negativos (no detectar cáncer), falsos positivos (intervenciones innecesarias) y errores de distribución (funciona en varones, pero falla en mujeres).

2. **Daño por automatización**: el clínico reduce la vigilancia, se produce complacencia, o se inhibe la discrepancia.

3. **Daño por sistema**: el algoritmo es correcto en promedio, pero introduce una política implícita (p. ej., priorizar la alta probabilidad de beneficio y relegar a los pacientes complejos), lo que altera el cuidado y la vigilancia.

La no maleficencia requiere instaurar barreras: controles humanos, umbrales de alerta, doble verificación en situaciones críticas, y cultura de notificación de incidentes algorítmicos. La seguridad del paciente debe incluir la "seguridad algorítmica".

IV. Justicia: sesgo, equidad, acceso y reparto de cargas

La justicia es quizá el principio más desafiado. La IA aprende a partir de datos de grandes bases de datos de salud. Si el algoritmo nace de una base sesgada y el pasado fue injusto, puede perpetuar injusticias bajo la apariencia de neutralidad.

Los sesgos pueden aparecer en:

- Datos: infrarrepresentación de ciertos grupos, registros incompletos, codificación distinta.
- Etiquetas: diagnóstico histórico sesgado; "verdades fundamentadas" inciertas.
- Objetivos: optimizar costes o reingresos puede penalizar la pobreza o la cronicidad.
- Despliegue: usar el modelo en poblaciones para las que no fue validado.

La justicia exige auditorías de equidad, estratificadas por sexo, edad, origen, discapacidad, nivel socioeconómico y otros determinantes. Y exige una idea esencial: la equidad no es "igual tratamiento", sino un tratamiento justo según la necesidad. En medicina, esto es crucial: un sistema que mejora el promedio puede empeorar a minorías y ser éticamente inaceptable.

d. Dilemas específicos de la IA clínica y la IA generativa

La conversación ética actual debe incluir no solo la IA clásica (clasificadores, modelos de riesgo), sino también la IA generativa (modelos de lenguaje), que introduce riesgos particulares: alucinaciones, persuasión y opacidad conversacional.

I. IA como "segunda opinión" y riesgo de sustitución

A menudo se propone la IA como "segundo lector" o "segunda opinión". Esto es valioso si se integra como contraste y no como autoridad. El riesgo surge cuando la organización convierte la recomendación en un estándar y penaliza al médico por discrepar de él. Deontológicamente, discrepar puede ser obligación, no capricho, cuando el médico detecta incongruencias clínicas, valores del paciente, o limitaciones del modelo.

La buena práctica requiere documentar la discrepancia razonada: por qué se aceptó o se rechazó la recomendación, qué variables clínicas no capturó el sistema, y qué incertidumbre se consideró. Esto protege al paciente y al profesional: convierte el acto en una deliberación trazable.

II. "*Chatbots*" clínicos: promesa de accesibilidad, riesgo de deshumanización

Los sistemas conversacionales pueden facilitar la información, el acompañamiento y el seguimiento. Pero la ética exige distinguir: informar no es cuidar. El cuidado requiere el reconocimiento del otro, la escucha, la interpretación del sufrimiento y la responsabilidad. Un *chatbot* no "responde" moralmente: produce texto.

Los riesgos específicos incluyen:

- Confundir al paciente sobre si interactúa con un humano.
- Dar consejos clínicos sin contexto (comorbilidades, medicación, signos de alarma).
- No detectar urgencias.
- Erosionar la confidencialidad al operar en plataformas gestionadas por terceros.

El criterio deontológico es claro: si un sistema conversa en un entorno clínico, debe estar sometido a supervisión y a límites estrictos de competencia (qué debe y qué no debe hacer) y debe derivar de inmediato a un profesional en situaciones de riesgo. Además, debe evitarse que el *chatbot* sustituya el acceso a un médico por razones económicas.

AV

III. IA y triaje: el dilema de priorizar en sistemas saturados

El triaje es un punto moralmente sensible: asigna tiempos y recursos. La IA puede mejorar la eficiencia, pero también puede introducir injusticias si prioriza según variables correlacionadas con "privilegio" (como la capacidad de describir síntomas, el acceso digital, el lenguaje). En urgencias y atención primaria, el triaje algorítmico debe ser especialmente conservador: tolerar más falsos positivos (*sobretriaje*) puede ser éticamente preferible a falsos negativos en colectivos vulnerables.

El triaje también puede ocultar decisiones que favorecen a los gestores. Un sistema que optimiza recursos puede convertirse en un mecanismo silencioso de racionamiento sin deliberación pública. La deontología obliga a la transparencia institucional: los criterios de priorización deben ser discutidos y legitimados, no sólo codificados.

IV. IA en imagen, patología y laboratorio: precisión y desplazamiento del oficio

En radiología, anatomía patológica y laboratorio, la IA puede alcanzar rendimientos extraordinarios. Sin embargo, una profesión que delega en exceso en su núcleo cognitivo puede atrofiar las competencias. Esto afecta la seguridad a largo plazo: cuando el sistema falla o se cae, ¿quién sostiene el conocimiento?

Deontológicamente, la profesión debe preservar competencias humanas relevantes. La IA debe integrarse de modo que aumente la capacidad diagnóstica y reduzca los errores, pero manteniendo un entrenamiento que evite la dependencia absoluta. Esto implica rediseñar formación y evaluación de especialistas.

V. Responsabilidad: del "quién decide" al "quién responde"

La responsabilidad es el punto neurálgico del núcleo ético. En medicina, decidir y responder están unidos. La IA tiende a separarlos: "decide" (recomienda) sin "responder" (no tiene agencia moral). Por eso, el médico y la institución deben construir una arquitectura de responsabilidad identificable.

VI. Responsabilidad clínica indelegable y deber de supervisión

El principio rector es que el médico responde de sus actos. Pero esta afirmación, por sí sola, puede resultar injusta si se utiliza para cargar sobre el clínico la opacidad o la mala calidad del sistema digital que utiliza. Deontológicamente, se requiere un equilibrio: el médico debe supervisar y no obedecer ciegamente, pero la institución debe garantizar herramientas validadas, transparentes y auditables. No puede pedirse al médico que "responda" por lo que no puede comprender ni controlar.

Por tanto, el manual de buenas prácticas debe afirmar: la responsabilidad es compartida en niveles, sin diluirse. El médico conserva la decisión final, la institución conserva la obligación de gobernanza, y los proveedores conservan obligaciones de calidad y seguridad del producto. Este enfoque promueve sistemas más seguros.

VII. Trazabilidad: condición ética de la rendición de cuentas

La trazabilidad no es un lujo técnico: es una condición moral. Sin registro de versiones del modelo, de datos de entrenamiento relevantes, de cambios y de contexto de uso, no puede haber aprendizaje institucional ni justicia ante daños. La trazabilidad permite:

- Investigar incidentes.
- Ajustar el modelo.
- Determinar si el error fue clínico, técnico u organizativo.
- Prevenir repetición.

Además, la trazabilidad favorece un consentimiento más honesto: si el sistema tiene límites conocidos, estos deben comunicarse.

VIII. La tentación del "cumplimiento" sin ética

Existe un riesgo contemporáneo: convertir la ética en *checklists.* Esto ocurre cuando la institución se centra en los documentos de conformidad y no en las prácticas reales: el algoritmo cumple con los formularios, pero el paciente sufre sesgos. La deontología debe resistir este vaciamiento. La ética es práctica reflexiva: deliberación, prudencia, revisión, y corrección. Un manual colegial debería insistir en cultura profesional: capacidad de decir "no" a sistemas no fiables, y de detener despliegues ante señales de daño.

e. Conflictos de interés, industria y mercantilización del cuidado

La IA en salud es un mercado. Donde hay mercado, pueden surgir conflictos de interés. La ética médica es particularmente sensible a la mercantilización del cuerpo. Los riesgos incluyen:

- Adopción de sistemas por presión comercial y no por evidencia.
- Dependencia de proveedores debido a infraestructuras cerradas.
- Opacidad contractual respecto del uso de los datos y de la propiedad de los modelos.
- Incentivos para aumentar las pruebas o las derivaciones.

La deontología exige transparencia sobre los conflictos y un criterio de sobriedad: introducir tecnología cuando aporta valor clínico probado y no cuando aporta sólo prestigio institucional o "marca innovadora". El médico, como profesional, debe conservar la libertad para cuestionar herramientas, aun cuando estén integradas en sistemas corporativos.

f. Formación y virtudes profesionales en la era algorítmica

El futuro ético de la IA depende de las competencias y las virtudes. No basta con enseñar "cómo usar" los sistemas; hay que enseñar "cómo pensar con ellos". Dos competencias resultan esenciales:

1. **Alfabetización algorítmica clínica**: entender nociones básicas de sesgo, validación, sensibilidad/especificidad, y límites de generalización.
2. **Hermenéutica clínica**: saber traducir los datos generados en significado humano, integrando los valores, el contexto y la narrativa del paciente.

Junto a las competencias, hay virtudes profesionales específicas:

- **Prudencia**: resistir decisiones precipitadas por la urgencia tecnológica.
- **Humildad epistémica**: reconocer los propios límites y los de la máquina.
- **Coraje moral**: discrepancia razonada ante el sistema que empuja a decisiones injustas.
- **Compasión**: recordar que el sufrimiento no se reduce a variables.

Estas virtudes no son ornamentales: son barreras de seguridad.

3. Parte Práctica

3.1. Hacia una ética operativa: criterios de buena praxis para "medicina inteligente, decisiones humanas"

Un capítulo de núcleo ético debe ofrecer criterios que puedan traducirse en una gobernanza efectiva. Diez criterios deontológicos operativos (no como *checklist* burocrático, sino como "prueba" moral de aceptación) relevantes son:

1. **Finalidad clínica legítima**: el objetivo debe ser mejorar la salud y el cuidado, no la vigilancia ni el racionamiento encubierto.
2. **Evidencia de beneficio**: no basta con la precisión técnica; se requieren impacto clínico y evaluación de daño.
3. **Explicación clínica suficiente**: el médico debe poder justificar ante el paciente la recomendación y su incertidumbre.
4. **Consentimiento significativo**: información clara sobre el papel de la IA, las alternativas y la supervisión humana.
5. **Equidad y auditoría**: evaluación de desempeño por subgrupos y corrección activa de sesgos.
6. **Privacidad por diseño**: minimización de datos, control de accesos, y límites a usos secundarios.
7. **Trazabilidad y registro**: versiones, cambios, incidencias y mecanismos de revisión.
8. **Supervisión humana real**: posibilidad efectiva de anular, corregir y escalar las decisiones.
9. **Responsabilidad distribuida sin dilución**: roles claros de médico, institución y proveedor.
10. **Preservación del vínculo clínico**: la IA debe liberar tiempo para el cuidado, no sustituirlo.

Estos criterios encarnan el lema del proyecto: la inteligencia puede ser tecnológica; la decisión debe seguir siendo humana, es decir, deliberada, responsable y compasiva.

3.2. Prospectiva: escenarios futuros y vigilancia ética

Mirando hacia el futuro inmediato, pueden anticiparse cuatro escenarios plausibles, cada uno con riesgos éticos específicos:

1. **IA ubicua e invisible**: integrada en historias clínicas, prescripción, triaje y documentación automática. Riesgo: automatización silenciosa y pérdida de agencia del médico. Respuesta ética: transparencia interna y formación.

2. **Medicina predictiva intensiva**: estratificación de riesgo poblacional y medicina preventiva personalizada. Riesgo: estigmatización, discriminación aseguradora o laboral, y "culpabilización" del paciente. Respuesta ética: límites a usos no clínicos y defensa de derechos.

3. **IA generativa como interfaz universal**: el paciente accede al sistema a través de *chatbots* y asistentes. Riesgo: deshumanización y sustitución del acceso al cuidado médico. Respuesta ética: supervisión, límites, y garantía de acceso humano.

4. **Reconfiguración del oficio**: las especialidades con un fuerte componente cognitivo y de imagen se transforman. Riesgo: dependencia y atrofia de competencias. Respuesta ética: rediseño formativo y preservación de habilidades críticas.

4. Bibliografía

1. Consejo General de Colegios Oficiales de Médicos de España (CGCOM). Código de Deontología Médica. Madrid: Organización Médica Colegial de España; 2022. Disponible en: https://www.cgcom.es

2. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021. ISBN: 978-92-4-002920-0. Disponible en: https://iris.who.int/handle/10665/341996

3. European Parliament and Council. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. 2024;L 2024/1689 (12 July 2024). Disponible en: http://data.europa.eu/eli/reg/2024/1689/oj

4. European Parliament and Council. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Official Journal of the European Union. 2016;L119:1-88. Disponible en: http://data.europa.eu/eli/reg/2016/679/oj

5. Council of Europe. Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (Oviedo Convention). ETS/CETS No. 164. Oviedo; 4 April 1997. Disponible en: https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=164

6. World Medical Association. WMA International Code of Medical Ethics. Adopted by the 3rd WMA General Assembly, London, October 1949; revised by the 73rd WMA General Assembly, Berlin, Germany, October 2022. Disponible en: https://www.wma.net/policies-post/wma-international-code-of-medical-ethics/

7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7

8. Char DS, Shah NH, Magnus D. Implementing machine learning in health care, addressing ethical challenges. N Engl J Med. 2018;378(11):981-983. doi:10.1056/NEJMp1714229

9. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. PLoS Med. 2018;15(11):e1002689. doi:10.1371/journal.pmed.1002689

10. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med. 2018;169(12):866-872. doi:10.7326/M18-1990

11. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337-1340. doi:10.1038/s41591-019-0548-6

12. Amann J, Blasimme A, Vayena E, Frey D, Madai VI; Precise4Q Consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20(1):310. doi:10.1186/s12911-020-01332-6

13. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195. doi:10.1186/s12916-019-1426-2

14. Price WN 2nd, Cohen IG. Privacy in the age of medical big data. Nat Med. 2019;25(1):37-43. doi:10.1038/s41591-018-0272-7

15. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics. 2020;46(3):205-211. doi:10.1136/medethics-2019-105586

Glosario y conceptos clave

Alejandro Trujillo González
Experto y líder de innovación en IA en Fresh Health

Glosario y conceptos clave (Alejandro Trujillo, experto y líder de innovación en IA en Fresh Health)

1. Fundamentos tecnológicos y clínicos

- **Inteligencia Artificial (IA):** Rama de la informática dedicada a crear sistemas capaces de realizar tareas que tradicionalmente requieren inteligencia humana (aprendizaje automático, reconocimiento de patrones).
- **Inteligencia Aumentada:** Enfoque de la IA diseñado para potenciar las capacidades humanas en lugar de reemplazarlas, actuando como un copiloto para el profesional de la salud.
- **Diagnóstico asistido por IA:** Uso de algoritmos, especialmente de visión artificial, para analizar datos médicos (radiografías, biopsias) y ayudar a detectar anomalías o clasificar enfermedades.
- **Sistema de Soporte a la Decisión Clínica (CDSS):** Herramientas informáticas que cruzan los datos de un paciente con bases de conocimiento médico para generar alertas, sugerencias de diagnóstico o recordatorios.
- **Producto Sanitario Software (SaMD):** Software destinado a fines médicos (diagnóstico, tratamiento) que funciona por sí mismo y está sujeto a normativas y certificaciones de seguridad estrictas (ej. marcado CE).
- **Historia Clínica Electrónica (HCE):** Registro digitalizado y estructurado de la salud de un paciente. Es la fuente principal de datos para entrenar y aplicar modelos de IA médica.
- **Interoperabilidad Sanitaria:** Capacidad de diferentes sistemas, dispositivos y aplicaciones de salud para conectarse e intercambiar datos de manera fluida y comprensible.
- **Traslación Clínica:** El proceso de llevar un descubrimiento científico o un modelo algorítmico desde el laboratorio hasta su implementación real en los flujos de trabajo del hospital.

Tumor growing
Swollen Thyroid
CONTROL PANEL - ASSISTANT
CONTROL CENTER
FILES & DATA

2. Modelos Avanzados e Investigación

- **Medicina Basada en Datos:** Enfoque que utiliza el análisis de grandes volúmenes de información (genómica, HCE, biosensores) para personalizar tratamientos y predecir riesgos.
- **Modelos Fundacionales:** Sistemas de IA a gran escala entrenados con datos masivos no etiquetados que sirven como base versátil para adaptarse rápidamente a múltiples tareas clínicas concretas.
- **Ensayos Clínicos Adaptativos:** Estudios que permiten modificaciones planificadas (cambios de dosis o muestras) mientras están en curso, utilizando IA para analizar los datos en tiempo real.
- **RWD / RWE (Datos y Evidencia del Mundo Real):** Datos recopilados en la práctica médica diaria (RWD) y el conocimiento clínico obtenido al analizarlos (RWE) para entender cómo funciona un tratamiento en la población general.
- **Gemelos Digitales:** Simulaciones virtuales precisas de un paciente u órgano que se alimentan de datos del mundo real para probar virtualmente tratamientos antes de aplicarlos físicamente.
- **Gestión de Datos Sanitarios:** Procesos y normativas para recopilar, almacenar y analizar información de salud, garantizando su calidad, privacidad y utilidad para la investigación.
- **Integridad Científica en la era generativa:** El reto de mantener el rigor ético y la honestidad, evitando la fabricación de datos por parte de la IA y declarando su uso de forma transparente en la investigación.

3. Ética, Humanismo y Relación Clínica

- **Humanismo:** Enfoque que sitúa a la persona y su dignidad en el centro, exigiendo que la IA se utilice para potenciar el cuidado, la empatía y el acompañamiento humano.
- **Ética Médica y Deontología:** Principios (beneficencia, no maleficencia) y deberes profesionales que deben guiar el uso de la IA, asegurando que las decisiones algorítmicas respeten la moralidad clínica.
- **Relación Clínica:** El vínculo de confianza entre médico y paciente, donde la IA debe actuar como una herramienta que libere tiempo para la atención humana, no como una barrera.
- **Dignidad Algorítmica:** Principio que defiende que los pacientes no deben ser tratados como meras predicciones estadísticas, respetando su vulnerabilidad e individualidad.
- **Autonomía y Consentimiento Informado:** El derecho del paciente a tomar decisiones libres sobre su cuerpo y a ser advertido explícitamente si una IA va a participar en su diagnóstico o tratamiento.
- **Objeción de Conciencia Tecnológica:** Derecho emergente de los profesionales a negarse a usar un sistema de IA si consideran que carece de evidencia, perpetúa sesgos o pone en riesgo al paciente.

4. Marco Legal, Derechos y Privacidad

- **Ley de IA (AI Act):** Normativa europea que clasifica la IA médica como de "alto riesgo", exigiendo evaluaciones estrictas, supervisión humana y trazabilidad antes de su uso.
- **Derechos de los Pacientes y de Cuarta Generación:** Derechos en el entorno digital, incluyendo el derecho a no ser objeto de decisiones 100% automatizadas, a la privacidad de neurodatos y a la transparencia algorítmica.

- **Protección de Datos y Secreto Profesional:** Obligación de aplicar técnicas como la anonimización a los datos de salud ("categoría especial") para entrenar la IA sin vulnerar la privacidad del paciente.
- **Espacio Europeo de Datos de Salud (EEDS):** Iniciativa de la UE para crear un ecosistema seguro que permita el uso de datos de salud para atención médica (uso primario) y para investigación y entrenamiento de IA (uso secundario).
- **Desplegador (Deployer):** Entidad (como un hospital) que utiliza un sistema de IA bajo su autoridad, asumiendo la responsabilidad de formar a su personal y monitorizar el software.

5. Seguridad, Calidad, Sesgos y Gobernanza

- **Seguridad del Paciente:** Disciplina que busca prevenir daños. En IA, implica asegurar que los algoritmos no induzcan a errores diagnósticos.
- **Supervisión Humana (Human-in-the-loop):** Principio que exige que un experto humano valide e interprete los resultados de la IA antes de tomar una decisión clínica.
- **Sesgo Algorítmico y Validez Externa:** Errores sistemáticos en la IA por datos de entrenamiento poco diversos (sesgo) y la necesidad de que el modelo mantenga su precisión en diferentes poblaciones (validez externa).
- **Justicia y Equidad:** Exigencia de que la IA se diseñe para cerrar brechas de salud y funcionar correctamente para todos los grupos demográficos, sin discriminar.
- **Transparencia y Explicabilidad:** La necesidad de no ocultar el uso de IA y la capacidad de comprender *por qué* el algoritmo llegó a una conclusión específica (evitando la "caja negra").

- **Trazabilidad y Reproducibilidad:** Capacidad de registrar todo el ciclo de vida del algoritmo y garantizar que arroje resultados consistentes si se aplica a los mismos datos en otro momento.
- **Evaluación Clínica:** Proceso científico riguroso para demostrar empíricamente que un software de IA es seguro y eficaz en el entorno médico real.
- **Responsabilidad Profesional (Accountability):** La obligación de definir quién asume la responsabilidad (médico, hospital o fabricante) si un diagnóstico asistido por IA causa daño.
- **Gobernanza Algorítmica y Uso Responsable:** Conjunto de reglas, protocolos y auditorías institucionales para gestionar y vigilar continuamente la seguridad, el impacto ético y el rendimiento de la IA en la práctica clínica diaria.

El “Manual de Buenas Prácticas para la Inteligencia Artificial en Medicina” es una iniciativa de la Organización Médica Colegial (OMC) y forma parte de su proyecto estratégico para establecer un marco técnico, ético, y deontológico en la práctica clínica, en el contexto de la incorporación de la salud digital a la medicina. Todo ello guiado por un principio rector innegociable: Medicina inteligente, decisiones humanas.

OMC
ORGANIZACIÓN MÉDICA COLEGIAL DE ESPAÑA
CONSEJO GENERAL DE COLEGIOS OFICIALES DE MÉDICOS