



Lineamiento para la implementación, monitoreo y evaluación de tecnologías basadas en inteligencia artificial en el Sistema Nacional de Salud

JUNIO 2026

**Dirección General de
Modernización del Sector Salud**



Gobierno de
México

Salud
Secretaría de Salud

Publicado por la Dirección General de Modernización del Sector Salud

1ª edición, 2026

D.R. Secretaría de Salud.

Agrarismo 227, Piso 2.

Col. Escandón II Secc.

DT. Miguel Hidalgo. C.P.

11800, CDMX.

Hecho en México.

Se permite la reproducción total o parcial, sin fines comerciales, citando la fuente.

DIRECTORIO

Dr. David Kershenobich Stalnikowitz
Secretario de Salud

Lic. Eduardo Clark García Dobarganes
Subsecretario de Integración Sectorial y Coordinación de Servicios de Atención Médica

ELABORACIÓN

DIRECCIÓN GENERAL DE MODERNIZACIÓN DEL SECTOR SALUD

Dr. José Enrique Pérez Olguín
Director General de Modernización del Sector Salud

Dra. Pamela Stephany Guiovanna del Moral Villavicencio
Dirección De Desarrollo E Integración De Medicina Basada En La Evidencia

Ing. Adrian Pacheco López
Director de Políticas de Tecnologías para la Salud

Dra. Mariana Beatriz Pineda López
Subdirección de Evaluación de Tecnologías en Salud

Mtro. Fernando Antonio Vázquez Franco
Subdirección de Planeación y Validación de Equipamiento Médico

Dr. Miguel Patiño González
Médico Especialista

Dra. Carolina Alejandra Padilla Mejía
Jefa de Departamento de Evaluación de Medicamentos

Lic. Ana Cristina Gutiérrez Jiménez
Jefa de Departamento de Evaluación de Instrumental y Equipo Médico

Dra. Karina Salas Martínez
Enfermera Especialista en Área Normativa

COMISIÓN FEDERAL PARA LA PROTECCIÓN CONTRA RIESGOS SANITARIOS

Dr. Víctor Hugo Borja Aburto
Comisionado Federal para la Protección contra Riesgos Sanitarios

Dr. Rodrigo Bazúa Lobato
Comisionado de Autorización Sanitaria

COMISIÓN NACIONAL DE BIOÉTICA

Dr. Patricio Javier Santillán Doherty
Comisionado Nacional de Bioética

Dra. Gabriela Pineda Hernández
Directora de Desarrollo Institucional

Mtro. Gustavo Fernando Olaiz Barragán
Subdirector de Políticas Públicas y Bioética

Mtra. Graciela Mendoza de la Rosa
Jefa del Departamento de Análisis de Políticas Públicas

Lic. Rebeca Elizabeth Rivero Riqué
Soporte administrativo - Políticas Públicas y Bioética

Contenido

Glosario General	2
Lista de Siglas y Acrónimos	7
Capítulo I. Disposiciones Generales	8
Capítulo II. Gobernanza y Ciclo de Vida	10
Capítulo III. El Acto Médico Aumentado y la Intersección entre la Pirámide del Razonamiento Clínico y la Pirámide Diagnóstica	14
Capítulo IV. Modalidades Operativas de Implementación Clínica	20
Capítulo V. Base Nacional de Información en Salud e Interoperabilidad de Sistemas	29
Capítulo VI. Clasificación por riesgo	32
Capítulo VII. Mecanismos de Reconocimiento y Validación	39
Capítulo VIII. Vigilancia y Seguimiento	43
Referencias	49

Glosario General

Para efectos de este lineamiento, se establecen las siguientes definiciones:

A

Acto médico: Proceso intelectual, deliberativo y relacional mediante el cual el profesional de la salud y el paciente, en una alianza terapéutica, toman decisiones diagnósticas y terapéuticas de manera compartida. Este proceso reconoce la centralidad de la persona usuaria, atendiendo no solo a la evidencia clínica, sino a las inquietudes, intereses y valores del paciente bajo un enfoque de bioética clínica.

Acto médico aumentado: Modalidad de práctica clínica en la cual un sistema de inteligencia artificial fortalece la base informativa del razonamiento clínico sin sustituirlo, optimizando la recolección y análisis preliminar de datos para que el profesional de la salud concentre su capacidad en los niveles superiores de decisión clínica.

Analítica predictiva de riesgos en el plano clínico-preventivo: Modalidad operativa en la que el sistema estima la probabilidad de que un paciente o grupo desarrolle una complicación o evento clínico, a partir de las fuentes de información disponibles, con el fin de anticipar la intervención preventiva. Su alcance se circunscribe a la dimensión clínico-preventiva.

Asistencia en la documentación clínica: Modalidad operativa en la que el sistema, mediante reconocimiento de voz e inteligencia artificial generativa, asiste en la elaboración de la nota médica o en la síntesis del expediente, bajo revisión y suscripción profesional.

Autoridad regulatoria de referencia: Autoridad sanitaria extranjera cuyas evaluaciones se reconocen conforme al marco de equivalencia vigente

C

Ciclo de vida: Sucesión de fases de un sistema de inteligencia artificial que comprende el desarrollo, validación, autorización, implementación, monitoreo continuo y retiro, sobre la que se ejerce la gobernanza institucional de manera diferenciada, sin que ninguna fase quede exenta de supervisión.

Clasificación por riesgo: Régimen que ordena los sistemas de inteligencia artificial en salud en niveles de riesgo bajo, moderado y alto, determinados por los criterios primarios de significancia de la información y estado de la condición sanitaria, modulados por el nivel de autonomía operativa.

Cuantificación y segmentación avanzada: Modalidad operativa en la que el sistema realiza mediciones y caracterizaciones morfológicas que aportan datos cuantitativos al seguimiento clínico y a la planificación terapéutica, sujetas a validación humana de la segmentación.

D

DataSheet (dataset): ficha técnica u hoja de datos que describe las especificaciones, explica el rendimiento y las características técnicas de un sistema.

Deriva algorítmica: Degradación del desempeño de un sistema de inteligencia artificial a lo largo del tiempo, por cambios en los datos, la población o las condiciones de uso respecto de aquellas en que fue desarrollado y validado. También conocida como deterioro del modelo.

Desviación relevante: Caída sostenida del desempeño de un sistema por debajo de los umbrales definidos en el plan institucional de monitoreo, que activa el protocolo de alerta y respuesta.

Dispositivo médico habilitado con inteligencia artificial: Dispositivo médico, conforme al artículo 262, fracción VI, de la Ley General de Salud, que utiliza tecnología de inteligencia artificial para realizar su función clínica o como componente esencial de ésta. Comprende el software como dispositivo médico y el software embebido en un dispositivo médico de hardware.

E

Evidencia en vida real (Real-World Evidence, RWE): Información clínica generada a partir del análisis de datos obtenidos en condiciones reales de uso, fuera de entornos controlados de investigación, utilizada para evaluar el desempeño, seguridad y efectividad de una tecnología en la práctica clínica habitual. Empleada en la evaluación continua a lo largo del ciclo de vida.

G

Gobernanza de inteligencia artificial: Modelo institucional de evaluación, implementación y monitoreo de las tecnologías de inteligencia artificial a lo largo del ciclo de vida, con componentes de supervisión, gestión del cambio y toma de decisiones adaptativa en sentido institucional.

H

Humano al mando: Principio rector que establece que en toda fase del ciclo de vida de un sistema de IA, la supervisión, interpretación y responsabilidad final de las decisiones clínicas recae en el profesional de la salud como experto técnico y el

paciente como experto en su propia vida. Las recomendaciones algorítmicas tienen carácter auxiliar y no vinculante. Ningún sistema de IA podrá sustituir el juicio clínico del profesional de la salud y las necesidades del paciente.

I

Incidente adverso: A cualquier acontecimiento comprobado que está relacionado con el uso de un dispositivo médico que cuente con pruebas contundentes de la relación causal entre el incidente y el dispositivo médico, y que pudiera ser ocasionado por un mal funcionamiento o alteración de las características del dispositivo médico y que pueda provocar la muerte o un deterioro grave de la salud del usuario. No se considerará incidente adverso a aquel derivado del uso anormal o un uso diferente del recomendado por el titular del registro sanitario del dispositivo médico o su representante legal en México.

Inteligencia Artificial (IA): Sistema computacional que utiliza algoritmos de aprendizaje automático, procesamiento de lenguaje natural u otras técnicas para analizar datos clínicos o administrativos y generar salidas que apoyen procesos asistenciales, diagnósticos, preventivos o de gestión en el ámbito de la salud.

Intención de uso: Finalidad declarada o razonablemente previsible de un sistema de inteligencia artificial, criterio determinante para calificarlo como sistema de inteligencia artificial en salud y para su clasificación por riesgo. Se distingue del uso institucional declarado, que la institución fija para los modelos de propósito general.

Interoperabilidad de intercambio (sintáctica): Capa de interoperabilidad que define la estructura y la transmisión de la información entre sistemas.

Interoperabilidad semántica: Capa de interoperabilidad que define la codificación de los conceptos clínicos para preservar su significado entre sistemas, mediante terminologías de referencia como SNOMED CT y LOINC.

Instancia institucional de evaluación de inteligencia artificial: Instancia multidisciplinaria de cada institución encargada de evaluar la utilidad clínica, la seguridad, la pertinencia y el cumplimiento regulatorio de las tecnologías, autorizar su implementación, supervisar su desempeño y, en su caso, recomendar su suspensión o retiro.

L

Lectura concurrente: En la modalidad de segunda lectura o detección asistida, modo en que el profesional y el sistema analizan el estudio de manera simultánea.

Lectura retrospectiva o de seguridad: En la modalidad de segunda lectura o detección asistida, modo en que el sistema opera como control de calidad posterior a la interpretación profesional.

P

Plan institucional de monitoreo: Plan que la institución establece previo al despliegue de cada sistema, con los indicadores de desempeño, los umbrales aceptables, la periodicidad de evaluación y la responsabilidad del monitoreo, bajo la coordinación de la instancia institucional de evaluación.

Portabilidad del dato clínico: Capacidad de transferir el dato clínico entre sistemas preservando su estructura y significado, como condición de continuidad de la atención y trazabilidad.

R

Razonamiento clínico: Proceso intelectual deliberativo mediante el cual el profesional, a partir de los datos clínicos del paciente y de su conocimiento médico estructurado, formula hipótesis diagnósticas, las contrasta con evidencia adicional y decide una conducta diagnóstica o terapéutica.

Reconocimiento de equivalencia (Reliance): Práctica regulatoria mediante la cual la autoridad sanitaria nacional toma en cuenta y otorga peso significativo a las evaluaciones de autoridades regulatorias de referencia para llegar a su propia decisión, conservando su independencia, autonomía técnica y responsabilidad.

S

Segunda lectura o detección asistida: Modalidad operativa en la que el sistema señala regiones de interés o hallazgos sospechosos que el profesional revisa con atención específica, en procesos de tamizaje o de revisión de alto volumen. El profesional conserva la condición de dictaminador único.

Sesgo de automatización: Tendencia del profesional a aceptar acríticamente las salidas de un sistema automatizado, incluso cuando son incorrectas, particularmente bajo fatiga cognitiva, alta carga asistencial o presión de tiempo.

Sistema Algorítmico: Es una herramienta tecnológica dinámica basada en Inteligencia Artificial y Aprendizaje Automático (AI/ML) que, a diferencia del software tradicional de reglas fijas, posee la capacidad de aprender, iterar y ajustar sus funciones lógicas de forma continua a partir de la exposición a grandes volúmenes de datos biomédicos. Su propósito es procesar flujos complejos de información clínica para generar inferencias, predicciones y diagnósticos auxiliares automatizados,

constituyendo una entidad evolutiva cuyo rendimiento y seguridad deben ser monitoreados activamente a lo largo de todo su ciclo de vida en el mundo real.

Sistema de inteligencia artificial: Sistema basado en máquina, diseñado para funcionar con distintos niveles de autonomía, que puede mostrar capacidad de adaptación tras su despliegue y que, para objetivos explícitos o implícitos, infiere a partir de la información de entrada la manera de generar resultados de salida, predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales.

Sistema de Soporte para la Decisión Clínica (SSDC): herramienta algorítmica que provee información procesada, alertas o recomendaciones al profesional de la salud con el propósito de apoyar, sin sustituir, el juicio clínico en el proceso diagnóstico o terapéutico. Corresponde al término en inglés Clinical Decision Support System (CDSS).

Software como Dispositivo Médico (SaMD): software cuya función prevista es el diagnóstico, prevención, monitoreo, tratamiento o alivio de enfermedades o condiciones de salud, sin formar parte de un dispositivo médico físico. Definición conforme al International Medical Device Regulators Forum (IMDRF).

Sugerencia algorítmica: Salida de un sistema de inteligencia artificial utilizada como insumo del razonamiento clínico, incluyendo diagnósticos diferenciales sugeridos, alertas, recomendaciones terapéuticas, predicciones de riesgo, clasificaciones, reportes de hallazgos priorizados y resultados análogos.

T

Tarjeta de Modelo: Documento técnico estandarizado que describe la procedencia de los datos de entrenamiento, los usos previstos del sistema, las métricas de desempeño, las limitaciones conocidas del modelo y sus condiciones de implementación. Constituye el estándar obligatorio de transparencia y trazabilidad algorítmica en el presente lineamiento.

Tecnovigilancia (vigilancia de la seguridad de los dispositivos médicos): Al conjunto de actividades que tienen por objeto la identificación y evaluación de incidentes adversos producidos por los dispositivos médicos en uso así como la identificación de los factores de riesgo asociados a éstos, con base en la notificación, registro y evaluación sistemática de las notificaciones de incidentes adversos, con el fin de determinar la frecuencia, gravedad e incidencia de los mismos para prevenir su aparición y minimizar sus riesgos. Idealmente, la información del sistema de tecnovigilancia se comparte entre autoridades competentes y fabricantes/distribuidores, a modo de facilitar las actividades en materia de

tecnovigilancia, así como las acciones preventivas y correctivas de cada caso a nivel nacional e internacional que impacten en el territorio nacional.

U

Umbral de desempeño: Valor aceptable mínimo definido en el plan institucional de monitoreo para cada indicador de desempeño de un sistema; su incumplimiento sostenido constituye una desviación relevante.

Uso previsto: Descripción formal de la función clínica o administrativa para la cual fue diseñado y validado un sistema de inteligencia artificial, incluyendo la población objetivo, el entorno de uso y el tipo de dato procesado.

V

Verificación humana: Acto deliberativo mediante el cual el profesional evalúa críticamente una sugerencia algorítmica antes de integrar al razonamiento clínico, considerando la congruencia con el cuadro clínico, las limitaciones declaradas del sistema y los criterios de la *lex artis ad hoc* aplicable.

Lista de Siglas y Acrónimos

COFEPRIS: Comisión Federal para la Protección contra Riesgos Sanitarios.

FDA: Administración de alimentos y medicamentos, del inglés: Food and Drug Administration.

FEUM: Farmacopea de los Estados Unidos Mexicanos

IA: Inteligencia Artificial

IMDRF: Foro Internacional de Reguladores de Dispositivos Médicos, del inglés: *International Medical Device Regulators Forum*

IECS: Instituto de Efectividad Clínica y Sanitaria.

LGS: General de Salud.

NICE: Instituto Nacional de Excelencia en Salud y Atención del Reino Unido, del inglés: National Institute for Health and Care Excellence.

PACS: Sistema de Comunicación y Archivo de Imágenes, del inglés: Picture Archiving and Communication System.

RWE: Evidencia del mundo Real, del inglés: Real World Evidence

SaMD: Software como dispositivo médico del inglés, del inglés: Software as a Medical Device.

SNS: Sistema Nacional de Salud

SSDC: Sistema de Soporte para la Decisión Clínica

UNAM: (Universidad Nacional Autónoma de México)

Capítulo I. Disposiciones Generales

Con fundamento en los artículos 3, 13, 17 Bis, 71 Bis, 133 y demás disposiciones aplicables de la Ley General de Salud (LGS) (1), así como en las atribuciones conferidas a la Secretaría de Salud en su carácter de autoridad rectora del Sistema Nacional de Salud (SNS), se emite el presente lineamiento con el propósito de establecer criterios técnicos, operativos y orientadores para el uso de Sistemas de Soporte para la Decisión Clínica (SSDC), también conocidos como Clinical Decision Support Systems, y de otras tecnologías basadas en Inteligencia Artificial (IA) aplicadas en procesos clínicos.

El objeto del lineamiento es regular y promover el empleo seguro, bioético, responsable y jurídicamente protegido de los SSDC, garantizando la protección de los pacientes, la salvaguarda jurídica del personal de salud y la responsabilidad institucional, en la habilitación, supervisión y gobernanza de los sistemas algorítmicos. Asimismo, busca asegurar el cumplimiento de la normativa sanitaria y de protección de datos personales vigente, así como la evaluación continua de la seguridad, eficacia y desempeño clínico de las soluciones implementadas.

El documento integra de manera estructurada dos enfoques complementarios: el modelo de gobernanza de tecnologías basadas en IA, sustentado en el ciclo de vida completo, la generación de evidencia y el monitoreo continuo del desempeño; y el marco normativo de los SSDC, que establece la primacía del acto médico, la centralidad del paciente y el principio de “humano al mando”. Esta articulación permite una supervisión tecnológica continua subordinada a la rectoría clínica del profesional de la salud.

Reconociendo la naturaleza particular de la IA, caracterizada por su capacidad de aprendizaje, adaptación y evolución en el tiempo, este instrumento establece un marco técnico, metodológico y operativo para la adquisición, implementación, monitoreo, evaluación y mejora continua de estas tecnologías, atendiendo los desafíos regulatorios, clínicos, operativos y bioéticos que las distinguen de las tecnologías tradicionales para la salud (2,3).

Principios rectores

La interpretación y aplicación del presente Lineamiento se rige por los siguientes principios rectores, que orientan la totalidad de sus capítulos:

- I. Humano al mando. La supervisión, interpretación y responsabilidad final de las decisiones clínicas corresponde al profesional de la salud y a la persona usuaria; las salidas de los sistemas de inteligencia artificial tienen carácter auxiliar y no vinculante.

- II.** Centralidad y autonomía de la persona. La persona usuaria es el centro de la atención y conserva el derecho a conocer el papel que la inteligencia artificial desempeña en su atención y a otorgar su consentimiento informado, en los términos que el presente Lineamiento determine.
- III.** Seguridad, no maleficencia y protección del interés público. Las tecnologías de inteligencia artificial no deben causar daño evitable y deben satisfacer requisitos de seguridad, exactitud y eficacia para usos bien definidos, con control y mejora continua de la calidad.
- IV.** Transparencia, explicabilidad y trazabilidad. Las decisiones apoyadas en inteligencia artificial deben poder reconstruirse y auditarse, con base en documentación suficiente, accesible y comprensible.
- V.** Responsabilidad y rendición de cuentas. La responsabilidad del uso adecuado de las tecnologías de inteligencia artificial corresponde a los actores humanos e institucionales, con líneas claras de atribución.
- VI.** Inclusión, equidad y no discriminación. Las tecnologías de inteligencia artificial deben fomentar el acceso y uso equitativos, minimizar sesgos y atender la representatividad de las poblaciones del país, en particular las subrepresentadas y las de acceso limitado a los servicios de salud.
- VII.** Sostenibilidad y capacidad de respuesta. Las tecnologías de inteligencia artificial deben evaluarse de forma continua durante su uso real y ser consistentes con la sostenibilidad institucional, presupuestal y ambiental del Sistema Nacional de Salud.

Ámbito de aplicación

Las disposiciones contenidas en el presente lineamiento son de observancia obligatoria y serán aplicables como instrumento rector y de referencia técnica para las instituciones públicas, sociales y privadas que integran el SNS y que utilicen, desarrollen, adquieran o implementen sistemas de apoyo algorítmico en procesos clínicos. Su ámbito de aplicación comprende a las entidades responsables de la prestación de servicios de atención médica, la evaluación de tecnologías para la salud, la regulación sanitaria, la adquisición de insumos, la gestión de información en salud, el desarrollo de infraestructura digital y la formulación de políticas públicas, entre otras.

Se establece como prioridad estratégica y técnica la mitigación proactiva de los sesgos algorítmicos, reconociendo que estos no representan únicamente un error técnico, sino un riesgo potencial para la salud pública con capacidad de profundizar disparidades existentes en el acceso y la calidad de la atención. Por tanto, la implementación de estas tecnologías en el SNS tiene como objetivo transversal prevenir daños y garantizar que sus beneficios sean inclusivos, respondiendo a las

necesidades de todas las comunidades, con especial atención a aquellas históricamente subrepresentadas en los sistemas de información en salud (4).

El Lineamiento reconoce que la confianza institucional y pública en los sistemas algorítmicos depende directamente de su transparencia, confiabilidad y explicabilidad. En consecuencia, es obligación de las instituciones asegurar que los algoritmos sean auditables y que sus recomendaciones resulten comprensibles para los profesionales de la salud, garantizando una toma de decisiones informada y la rendición de cuentas en cada etapa del proceso asistencial.

En consecuencia, este lineamiento deberá incorporarse de manera transversal en los procesos institucionales de adquisición, implementación, operación, supervisión, evaluación y mejora continua de tecnologías basadas en IA, independientemente del nivel de atención, modelo de prestación o naturaleza de la institución.

Para efectos del cumplimiento del presente, las instituciones deberán contemplar, como mínimo, los siguientes elementos: mecanismos de gestión de riesgos; transparencia y trazabilidad algorítmica; documentación clínica y técnica; protección de datos personales conforme a la normativa vigente; criterios de validación clínica y técnica; definición de responsabilidades institucionales; y protocolos de respuesta ante eventos adversos o fallos operativos (5,6).

Capítulo II. Gobernanza y Ciclo de Vida

A diferencia de los modelos regulatorios convencionales, centrados principalmente en la autorización previa a la comercialización o a su uso, el lineamiento adopta un enfoque de gobernanza basado en el ciclo de vida completo de la tecnología, en el cual la evaluación no se limita a una fase inicial, sino que se extiende a lo largo de todo el periodo de uso en la práctica clínica. Este enfoque se implementa bajo el principio de “humano al mando”, asegurando que en todas las fases del ciclo de vida se mantenga la supervisión clínica y la responsabilidad del profesional de la salud.

Modelo de gobernanza

En este contexto, el modelo de gobernanza se estructura en cuatro componentes interdependientes:

- Supervisión clínica transversal en todas las fases del ciclo de vida.
- Validación en condiciones reales mediante la generación de evidencia real (*real-world evidence*, RWE).
- Monitoreo continuo del desempeño, con énfasis en la detección de deriva algorítmica, sesgos emergentes y variaciones en el rendimiento.
- Toma de decisiones adaptativa y basada en evidencia.

Este enfoque se alinea con las tendencias regulatorias internacionales que reconocen la necesidad de supervisión continua para tecnologías adaptativas, entre ellas el marco propuesto por la *Food and Drug Administration* (FDA) de los Estados Unidos para el Software as a Medical (SaDM) basado en IA. Dicho marco habilita mecanismos para ajustar, restringir o retirar tecnologías en función de su desempeño observado en condiciones reales, criterio que el lineamiento adopta para el contexto del SNS (9,10).

Se incorpora un enfoque de evaluación centrado en el paciente, en el cual el desempeño de las tecnologías no se valora únicamente en términos de precisión técnica o métricas algorítmicas, sino en función de su impacto real en la toma de decisiones clínicas, los resultados en salud, la seguridad del paciente y la eficiencia del sistema. Dado que la evidencia científica ha documentado que el alto desempeño en entornos controlados no garantiza beneficios clínicos equivalentes en la práctica real (37), las instituciones deberán evaluar estas tecnologías en condiciones operativas y con desenlaces clínicamente relevantes (2,11,12,35).

Fases del ciclo de vida

Para los efectos de este Lineamiento, el ciclo de vida de un sistema de inteligencia artificial comprende las siguientes fases:

- **Fase de desarrollo:** Es la etapa inicial del ciclo de vida donde se traducen los requisitos clínicos y las necesidades del usuario en especificaciones técnicas para construir el software. Abarca desde la planificación de la arquitectura del código y la programación, hasta las primeras pruebas técnicas internas (verificación).
- **Fase de validación:** Es la etapa crítica donde se confirma, mediante evidencia objetiva, que el software cumple consistentemente con su propósito médico previsto (Intended Use) y que es seguro y efectivo para los usuarios (médicos o pacientes).
- **Fase de autorización:** Es la etapa regulatoria y de conformidad en la cual el fabricante recopila toda la documentación técnica (declaración de definición, justificación de la categoría de riesgo e informes de validación) para demostrar que el SaMD cumple con los marcos legales aplicables.
- **Fase de implementación:** Es el proceso de despliegue, instalación y puesta en marcha del software en los entornos operativos de los usuarios finales.
- **Fase de monitoreo continuo:** Es la etapa de supervisión activa una vez que el software ya está siendo utilizado en el mundo real. Consiste en recopilar datos sobre el rendimiento operativo, reportes de errores, vulnerabilidades de seguridad cibernética y retroalimentación clínica. El objetivo de esta fase es realizar mantenimientos correctivos, preventivos o adaptativos para mitigar riesgos emergentes e identificar si una actualización altera las capacidades principales del software, lo que obligaría a reiniciar el ciclo.
- **Fase de retiro:** Es la etapa final en la que el fabricante decide retirar el software del mercado o discontinuar el soporte técnico y las actualizaciones para una versión específica (obsolescencia). Esta fase requiere una planeación estructurada para notificar a los usuarios y garantizar la seguridad de los datos de salud de los pacientes que dependían del sistema, asegurando que la desconexión del software o la falta de parches no genere un peligro latente en el entorno médico donde operaba.

Gestión del cambio y la articulación sanitaria

Para asegurar la sostenibilidad y legitimidad de estas tecnologías, los mecanismos de supervisión deberán ser de naturaleza multidisciplinaria, incorporando expertos técnicos, clínicos, especialistas en bioética y derecho, así como representantes de las comunidades afectadas. Este enfoque participativo tiene por objeto garantizar que los sistemas de IA se alineen con las prioridades de salud pública, permitan identificar sesgos no anticipados en el diseño y promuevan una adopción basada en la pertinencia cultural y la equidad operativa en todos los niveles de atención (7,8,34).

Por lo anterior, se establecen las bases del modelo nacional de gobernanza de Inteligencia Artificial, orientado a que el SNS no solo adopte tecnologías emergentes, sino que regule su uso de manera activa, genere evidencia propia en condiciones reales y gestione los riesgos asociados a su implementación, bajo principios de seguridad, trazabilidad y responsabilidad institucional.

Gestión de actualizaciones y revalidación

Las instituciones del Sistema Nacional de Salud establecerán procedimientos para gestionar las actualizaciones de los sistemas de inteligencia artificial que utilicen, distinguiendo entre:

- Actualizaciones de software en el cual, el modelo permanece bloqueado y se modifica mediante nuevas versiones, cada una de las cuales podrá requerir validación local adicional.
- Modelos adaptativos que se actualizan mediante aprendizaje continuo, los cuales exigen control de versiones, documentación de cambios y evaluación del impacto clínico antes de su puesta en operación.

Toda modificación relevante de un sistema en operación obligará a documentar el cambio, conservar el control de versiones y revalidar localmente el desempeño. Cuando la tecnología constituya dispositivo médico, esta gestión se entenderá sin perjuicio de las obligaciones que correspondan ante la autoridad sanitaria.

Gobernanza institucional

Este apartado define las instancias multidisciplinarias responsables de vigilar todo el ciclo de vida de los sistemas algorítmicos. Por medio de la supervisión constante, la gestión activa del sesgo poblacional y el diseño de protocolos de contingencia, se busca asegurar la equidad, calidad técnica y la continuidad ininterrumpida de la atención médica en beneficio de la población mexicana.

Instancia institucional de evaluación de inteligencia artificial: Cada institución del Sistema Nacional de Salud contará con una instancia de evaluación de las tecnologías basadas en inteligencia artificial, de composición multidisciplinaria, encargada de evaluar la utilidad clínica, seguridad, pertinencia y el cumplimiento regulatorio de dichas tecnologías, así como de autorizar su implementación, supervisar su desempeño y, en su caso, recomendar su suspensión o retiro.

La instancia integrará, según la estructura y los recursos de cada institución, perfiles clínicos, de informática médica, protección de datos personales, de calidad y seguridad

del paciente, de bioética y de gestión. La instancia se articulará con el responsable técnico institucional previsto en el Capítulo III.

Supervisión multidisciplinaria y participación: Los mecanismos de supervisión serán de naturaleza multidisciplinaria e incorporarán, cuando proceda, la perspectiva de las comunidades afectadas, con el objeto de alinear los sistemas de inteligencia artificial con las prioridades de salud pública, identificar sesgos no anticipados y promover la pertinencia cultural y la equidad operativa en todos los niveles de atención.

Gestión del sesgo y representatividad: La gobernanza integrará la gestión del sesgo algorítmico en todas las fases del ciclo de vida, desde la formulación del problema y la recolección de datos hasta la validación y el monitoreo posterior a la implementación. Los conjuntos de datos de entrenamiento y validación deberán observar estándares de representatividad que reflejen la diversidad de la población mexicana en términos de sexo, edad, distribución geográfica y factores socioeconómicos, a fin de evitar la pérdida de precisión de modelos entrenados en contextos externos o poblacionalmente limitados al aplicarse en el entorno nacional.

Monitoreo continuo y articulación con la vigilancia: Las instituciones implementarán mecanismos de monitoreo continuo que permitan evaluar el desempeño de los sistemas a lo largo de todo el ciclo de vida, con el propósito de detectar deriva algorítmica, sesgos emergentes y variaciones del rendimiento, así como de asegurar la protección del paciente en condiciones reales de uso. Los mecanismos concretos de vigilancia y reporte se rigen por el Capítulo VIII de este Lineamiento.

Continuidad del acto médico y contingencia: La implementación de tecnologías basadas en IA en el SNS deberá mantenerse subordinada al acto médico y al respeto de los derechos del paciente para garantizar el uso ético, seguro y justo de la tecnología en la medicina. El principio de “humano al mando”, definido en el Capítulo I, implica que ningún sistema algorítmico podrá sustituir el juicio clínico del profesional de la salud, que la interpretación de los resultados generados por la tecnología corresponde al médico tratante, y que la responsabilidad bioética, clínica y legal de las decisiones asistenciales recae en todo momento en quienes intervienen conforme a sus atribuciones profesionales e institucionales. Las instituciones garantizarán protocolos de contingencia ante fallos del sistema, de modo que el acto médico no se interrumpa ni se degrade por problemas de infraestructura o conectividad.

Capítulo III. El Acto Médico Aumentado y la Intersección entre la Pirámide del Razonamiento Clínico y la Pirámide Diagnóstica

La evolución de la práctica médica en el contexto de la salud digital en México demanda una precisión conceptual rigurosa para distinguir entre los procesos automatizados de la tecnología y los procesos cognitivos del profesional de la salud, para ello, es necesario establecer que si bien la IA interviene de manera disruptiva en la gestión de la información, el acto médico permanece anclado en un proceso de razonamiento clínico que es cualitativamente distinto al procesamiento algorítmico (36).

Verificación humana obligatoria

Bajo el marco normativo de la LGS y la interpretación jurisprudencial de la Suprema Corte de Justicia de la Nación, el acto médico se define como: “un proceso intelectual y humano orientado a la toma de decisiones en condiciones de incertidumbre, donde el profesional asume la responsabilidad bioética y legal de sus determinaciones”. Para comprender esta dinámica en la era de la modernización, debemos analizar de forma separada dos estructuras jerárquicas: la pirámide del razonamiento clínico y la pirámide diagnóstica, en cuya cúspide unificada deberá prevalecer la decisión compartida entre el paciente y el médico bajo el principio de “humano al mando”, en el cual, el mando reside en la interacción humana entre el profesional como experto técnico y el paciente como experto de su propia vida y, en donde la IA no debe interponerse, sino facilitar que el encuentro sea más informado, transparente y centrado en las necesidades del usuario. Véase la figura 1.

Pirámide del razonamiento clínico

Describe el proceso intelectual del médico desde el primer contacto con el paciente hasta la consolidación de un plan diagnóstico o terapéutico en el que se utilizan tanto procesos intuitivos rápidos, conocidos como sistema uno, como procesos analíticos deliberados, conocidos como sistema dos. En su base se encuentra la recolección de datos primarios mediante la anamnesis y la exploración física, un proceso que requiere de habilidades semiológicas, escucha activa, empatía y una comprensión integral del individuo. Esta fase es eminentemente humana a la vez que implica la decodificación de un lenguaje subjetivo y la observación de signos clínicos que a menudo son sutiles o ambiguos. Conforme se asciende en esta pirámide, el médico realiza un proceso de síntesis donde los datos se transforman en información clínica relevante, permitiendo la generación de hipótesis diagnósticas.

En la cúspide de esta pirámide se encuentra el juicio clínico integral, donde el médico pondera las hipótesis frente a la realidad biopsicosocial del paciente y decide una

conducta específica. Este proceso es el corazón del acto médico y se rige por principios bioéticos que exigen que la decisión final sea el resultado de una deliberación responsable que considere el respeto a la autonomía del paciente, su preferencia y consentimiento informado cuando resulte aplicable.

Pirámide diagnóstica

Esta pirámide se refiere a la arquitectura técnica y al flujo de procesamiento de la información dentro de los sistemas de IA y el software como dispositivo médico. Su base está constituida por la captura de datos a gran escala, tales como píxeles de imágenes médicas, señales electrofisiológicas digitales o secuencias genómicas. En este nivel, la tecnología actúa con una capacidad de procesamiento masivo que permite la normalización y el filtrado de la información, eliminando el ruido técnico que podría confundir al humano. El cuerpo de la pirámide corresponde a la extracción de características y al reconocimiento de patrones mediante algoritmos de aprendizaje profundo, aquí la IA realiza una tarea descriptiva de alta precisión, identificando biomarcadores o hallazgos radiológicos que pueden ser imperceptibles para el ojo humano y, en la cúspide de esta pirámide se encuentra la sugerencia diagnóstica probabilística o el reporte de hallazgos priorizados. Es fundamental comprender que esta cúspide técnica no constituye un diagnóstico médico, sino una inferencia estadística de alta complejidad que debe ser entregada al profesional de la salud como un insumo para su propio proceso de razonamiento.

El Acto Médico Aumentado

En la intersección entre estas dos pirámides, se sitúa el concepto de acto médico aumentado definido para efectos de este lineamiento como aquel en el cual, la tecnología fortalece la base informativa del razonamiento clínico sin sustituirlo y en donde la tecnología a través de la pirámide diagnóstica, enriquece la base informativa de la pirámide del razonamiento clínico, al proporcionar datos previamente analizados y sugerencias precisas. La IA reduce la carga cognitiva del médico en las fases de recolección y análisis preliminar, permitiendo que se dedique con mayor esfuerzo intelectual a la cúspide de su razonamiento, es decir, al juicio clínico, la comunicación con el paciente y la decisión bioética. Esta integración exige que el médico mantenga una vigilancia crítica constante sobre la tecnología.

La responsabilidad profesional en México dicta que el médico no puede limitarse a ser un receptor pasivo de las sugerencias de los sistemas informáticos por lo que debe ser capaz de auditar el proceso algorítmico, comprendiendo que la sugerencia diagnóstica de la máquina es una probabilidad y no una verdad absoluta. El acto médico es intransferible precisamente porque la decisión clínica en la cúspide de la pirámide del razonamiento clínico involucra una valoración de riesgos, beneficios y valores

humanos que la pirámide diagnóstica, por su naturaleza matemática, es incapaz de procesar (20).

Desde la perspectiva del derecho administrativo y la bioética, la distinción entre estas pirámides protege la integridad del sistema para la salud. Si se permitiera que la pirámide diagnóstica técnica sustituyera a la pirámide del razonamiento clínico, se estaría incurriendo en un reduccionismo biotecnológico que vulneraría la dignidad del paciente y diluiría la responsabilidad legal del Estado y los servidores públicos profesionales de la salud.

Figura 1. Intersección entre Pirámide del Razonamiento Clínico y la Pirámide Diagnóstica



Fuente: elaboración propia DGMoSS, 2026.

Nota: La línea del acto médico delimita el umbral a partir del cual ningún sistema algorítmico podrá sustituir el juicio clínico del profesional de la salud

La Suprema Corte de Justicia de la Nación ha señalado que la *lex artis* no es un conjunto de reglas rígidas, sino un estándar de conducta que se adapta a las circunstancias de modo, tiempo y lugar (19). En el entorno digital, esto significa que el médico debe utilizar las mejores herramientas disponibles en la pirámide diagnóstica para informar su razonamiento clínico, pero la decisión final siempre debe estar fundamentada en su juicio humano. La fundamentación y motivación de la nota médica, exigida por el marco jurídico mexicano, debe reflejar este proceso de validación humana donde el médico explica por qué aceptó o rechazó la sugerencia algorítmica.

El principio de “humano al mando” se convierte en el eje rector que garantiza que la pirámide diagnóstica permanezca al servicio de la pirámide del razonamiento clínico.

Este principio obliga a los desarrolladores de tecnología y a las instituciones para la salud a diseñar sistemas que sean explicables y transparentes. El médico no debe enfrentarse a una caja negra que emite dictámenes sin fundamento comprensible, la IA debe proporcionar los elementos de juicio necesarios para que el profesional pueda integrarlos en su razonamiento analítico.

Capacitación obligatoria del profesional

La formación de los recursos humanos para la salud en México bajo el enfoque del acto médico aumentado debe centrarse en esta dualidad, capacitando a los médicos no solo en la medicina de forma tradicional, sino en la evaluación crítica de la tecnología. Un médico competente en la era digital es aquel que domina su razonamiento clínico y que sabe cómo interrogar a la pirámide diagnóstica para obtener el máximo beneficio para su paciente (21).

Las instituciones del Sistema Nacional de Salud que implementen sistemas de inteligencia artificial deberán garantizar la capacitación documentada del personal profesional que los utilice, la cual, comprenderá al menos:

- I.** Funcionamiento general del sistema e intención de uso declarada.
- II.** Limitaciones técnicas conocidas del sistema.
- III.** Criterios de validación clínica aplicados.
- IV.** Procedimientos de verificación humana de salidas.
- V.** Obligaciones documentales conforme a lo establecido en el presente documento.
- VI.** Procedimientos institucionales de reporte de eventos adversos o desempeño anómalo.

Las instituciones deberán mantener registros de la capacitación recibida por cada profesional así como del periodo de validez de la misma y su actualización, se realizará ante modificaciones sustantivas del sistema, actualizaciones de su intención de uso y conforme al programa institucional periódico.

La implementación de este modelo jerárquico y dual en el Sistema Nacional de Salud requiere de una infraestructura digital sólida que garantice la interoperabilidad y la seguridad de los datos. La Base Nacional de Información en Salud, contemplada en la reforma de 2026 y cuyo marco de interoperabilidad se desarrolla en el Capítulo V, debe ser el terreno fértil donde la pirámide diagnóstica se alimente de datos de alta calidad para proporcionar sugerencias confiables.

Al mismo tiempo, el expediente clínico electrónico debe ser diseñado para facilitar el registro del razonamiento clínico humano, permitiendo que el médico documente su proceso de toma de decisiones de manera estructurada y segura. La soberanía de los datos personales es esencial en este proceso, pues la información del paciente es el

insumo básico que fluye a través de ambas pirámides. El profesional de la salud tiene la responsabilidad bioética de asegurar que este flujo de información se realice bajo los más estrictos estándares de confidencialidad y respeto a la privacidad del individuo.

Es imperativo referir la relación terapéutica en el orden procedente desde un enfoque bioético: la relación paciente-médico. Esto refleja la centralidad de la persona en el proceso de atención. La implementación de IA en el SNS debe garantizar que el médico recupere tiempo de calidad para la escucha activa, permitiendo que la decisión final surja de un diálogo que considere no solo la enfermedad, sino las inquietudes e intereses del paciente. El acto médico es, ante todo, un compromiso humano donde la tecnología es el puente, pero la voluntad del paciente y el juicio del médico constituyen la última y única palabra.

En el ámbito de la evaluación de tecnologías para la salud, es fundamental aplicar criterios de rigor científico para validar cada componente de la pirámide diagnóstica. Los organismos como el NICE y la FDA establecen marcos de evaluación que consideran no solo la precisión técnica del algoritmo, sino su impacto real en el desenlace para la salud del paciente.

La validación clínica de un SaMD debe demostrar que su integración en el proceso de razonamiento clínico mejora la efectividad diagnóstica sin introducir riesgos adicionales. La vigilancia continua del desempeño tecnológico es necesaria para detectar posibles degradaciones del algoritmo o sesgos que puedan surgir con el tiempo, garantizando que la base informativa de la pirámide diagnóstica se mantenga siempre íntegra y veraz.

Mitigación del sesgo de automatización

Las instituciones del Sistema Nacional de Salud implementan medidas para la mitigación, que comprenden al menos:

- I.** Verificación de que la interfaz del sistema comunique explícitamente la naturaleza probabilística de la salida, incluyendo, cuando proceda, el nivel de confianza y las variables determinantes consideradas.
- II.** Capacitación del personal sobre el riesgo del sesgo de automatización y sus factores agravantes.
- III.** Aplicación de las obligaciones documentales establecidas en la documentación del expediente clínico y registros institucionales del presente Capítulo, que materializan el ejercicio de verificación humana.
- IV.** Auditoría institucional periódica que identifique patrones de aceptación acrítica y aplique las medidas correctivas que correspondan.

La responsabilidad del Estado en la gestión de la salud digital implica también la protección de la autonomía del profesional de la salud. El médico debe estar libre de presiones administrativas o comerciales que pretendan imponer el uso de sugerencias algorítmicas por encima de su juicio clínico. La pirámide del razonamiento clínico es el bastión de la libertad profesional, y su defensa es esencial para garantizar un sistema para la salud que sea verdaderamente ético y profesional.

El lineamiento que aquí se presenta busca brindar esta seguridad jurídica y técnica al personal de la salud, estableciendo con claridad que la tecnología es un auxiliar y que el acto médico es, y seguirá siendo, una función humana irrenunciable. La modernización no se trata de reemplazar al médico por una máquina, sino de dotarlo de las mejores capacidades analíticas para que pueda ejercer su vocación con mayor precisión.

En conclusión, el acto médico en la era de la IA se define por la interacción dinámica y jerárquica entre la pirámide diagnóstica tecnológica y la pirámide del razonamiento clínico humano. La primera proporciona la infraestructura de datos y el reconocimiento de patrones con una potencia analítica sin precedentes, mientras que la segunda aporta la síntesis interpretativa, el juicio ético y la decisión soberana que constituyen la esencia de la medicina. El juicio clínico integral es donde culmina este proceso unificado, donde el médico, bajo el principio de humano al mando, valida la información tecnológica y la aplica de manera personalizada para el bienestar del paciente.

Atender este proceso de trabajo asegura que la transformación digital del sector salud en México sea seguro, legal, ético y profundamente humano. La preservación de la naturaleza intransferible del acto médico es la mayor garantía de que la tecnología será siempre una herramienta de progreso y no un sustituto de la responsabilidad humana.

Capítulo IV. Modalidades Operativas de Implementación Clínica

La implementación de tecnologías de IA en México no puede ser un proceso uniforme ni estandarizado para todos los escenarios clínicos, ya que el riesgo para la salud del paciente y la exigencia de la *lex artis ad hoc* varían según el propósito del software, el nivel de autonomía del algoritmo y la criticidad del entorno asistencial.

Por lo tanto, se determinan seis modalidades operativas fundamentales que rigen la incorporación de estas herramientas en los sistemas de la salud, asegurando que los profesionales de la salud mantengan la rectoría absoluta del proceso bajo el principio de humano al mando y que cada intervención tecnológica esté fundamentada en el derecho administrativo y la bioética clínica.

Con base en lo anterior el presente capítulo tiene el objetivo de describir las modalidades mediante las cuales los sistemas de inteligencia artificial se incorporan al flujo de trabajo asistencial en el Sistema Nacional de Salud, e identifica, para cada una, su perfil característico, los énfasis de validación que le corresponden y su articulación con el acto médico.

1. Triage y notificación prioritaria

La primera modalidad operativa, representa el uso de la IA como un gestor inteligente de la carga de trabajo y un detector de hallazgos críticos en tiempo real. En el entorno de los servicios de urgencias, áreas de medicina crítica y centros de diagnóstico de alta demanda, el tiempo es un factor determinante del pronóstico para la salud y la vida.

Bajo esta modalidad, el software actúa de manera autónoma en la base de la pirámide diagnóstica, analizando de forma continua cada estudio de imagen, señal electrofisiológica o dato de laboratorio en el momento de su adquisición o procesamiento, incluso antes de que el médico tenga el primer contacto visual con la información.

El objetivo técnico primordial es identificar patrones asociados a condiciones de emergencia médica inmediata, tales como hemorragias intracraneales, neumotórax a tensión, oclusiones de grandes vasos cerebrales, embolia pulmonar o arritmias cardíacas letales. Una vez detectado el hallazgo, el sistema genera una alerta visual o auditiva de alta prioridad que posiciona dicho estudio al inicio de la lista de interpretación del personal de salud, desplazando los casos de rutina o seguimiento. El sistema no emite diagnóstico ni sustituye la interpretación profesional; su función es la gestión del orden y la oportunidad de la revisión.

La jurisprudencia de la Suprema Corte de Justicia de la Nación ha sido enfática al señalar que la omisión en la prontitud de la atención en casos de urgencia médica constituye una violación grave al derecho a la protección de la salud consagrado en el artículo cuarto constitucional (19).

En consecuencia, el uso de IA para el triaje se alinea con el cumplimiento de este mandato al asegurar que los casos de mayor gravedad sean atendidos con la celeridad que el estado de la ciencia actual permite. Desde la perspectiva de la evaluación de tecnologías para la salud, la modalidad de triaje y notificación prioritaria requiere de una validación técnica con énfasis en la sensibilidad del algoritmo. Un falso negativo en esta modalidad implica que una condición potencialmente mortal no fue detectada a tiempo, lo que compromete de forma directa la seguridad del paciente y la integridad del sistema para la salud. Por el contrario, un exceso de falsos positivos puede generar el fenómeno de fatiga por alertas en el personal médico, diluyendo la efectividad real de la herramienta.

Por ello, las instituciones para la salud que decidan implementar esta modalidad deben contar obligatoriamente con un protocolo de validación local que verifique el desempeño del sistema en las condiciones reales de su infraestructura digital y con su perfil epidemiológico específico.

Acto médico

La responsabilidad del médico en esta modalidad se centra en la respuesta oportuna y diligente ante la notificación generada por el sistema. Si el algoritmo emite una alerta, el personal de salud debe validar el hallazgo de manera prioritaria, integrándose de inmediato en su propia pirámide del razonamiento clínico para confirmar el diagnóstico e iniciar la intervención terapéutica necesaria. El acto médico en este escenario se define por la validación del hallazgo crítico y la toma de decisiones subsecuentes, manteniendo siempre la autoridad de mando ante cualquier discrepancia técnica con el sistema algorítmico.

2. Segunda lectura o Detección asistida

La segunda modalidad operativa se sitúa predominantemente en el cuerpo de la pirámide diagnóstica y se enfoca en elevar la sensibilidad y consistencia del diagnóstico médico en procesos caracterizados por una alta fatiga cognitiva o por el tamizaje masivo de población.

En los programas nacionales de prevención para la salud, tales como el tamizaje de cáncer de mama, la detección de cáncer de pulmón en pacientes de riesgo o el diagnóstico de retinopatía diabética, el profesional de la salud se enfrenta a la revisión de volúmenes masivos de estudios donde la probabilidad de hallazgo positivo es baja, pero el impacto de una omisión diagnóstica es catastrófico para el individuo. Bajo la modalidad de detección asistida, la IA actúa como un observador concurrente que

señala regiones de interés, nódulos sospechosos o posibles anomalías que el médico deberá revisar con una atención especial y pormenorizada.

Esta modalidad se divide técnicamente en lectura concurrente, donde el médico y el algoritmo analizan el estudio de forma simultánea y lectura retrospectiva o de seguridad, donde el algoritmo actúa como un control de calidad posterior a la interpretación inicial del médico para detectar posibles discrepancias o hallazgos omitidos.

Administrativamente, la modalidad de segunda lectura requiere una integración técnica fluida con los sistemas de archivo y comunicación de imágenes (PACS) y con el expediente clínico electrónico institucional. Este lineamiento establece que el informe médico final debe indicar de forma explícita, la utilización de una herramienta de detección asistida, detallando la versión específica del software y el tipo de procesamiento realizado sobre la información del paciente. Esta transparencia es esencial para asegurar la trazabilidad del acto médico y para cumplir con los estándares de bioética, ya que permite que todos los profesionales involucrados en la red de atención conozcan el origen y el fundamento de las sugerencias diagnósticas.

Acto médico

En términos de responsabilidad profesional, el médico sigue siendo el dictaminador único y responsable del resultado final. Si el algoritmo señala una región de interés que el facultativo, tras un análisis crítico fundamentado en su razonamiento clínico y en la semiología de la imagen, considera un hallazgo benigno o un artefacto técnico, prevalece siempre el juicio humano sobre la sugerencia algorítmica. La autonomía del profesional para rechazar la sugerencia de la IA es la garantía fundamental de que la tecnología no imponga tratamientos innecesarios, biopsias injustificadas o intervenciones que vulneren el principio de no maleficencia, evitando así el fenómeno del sobrediagnóstico y el consecuente daño psicológico y físico al paciente.

3. Cuantificación y segmentación avanzada

Esta modalidad representa la cúspide de la medicina de precisión dentro de la pirámide diagnóstica técnica, en ella la IA se utiliza para realizar mediciones exactas y caracterizaciones morfológicas que serían imprecisas o demandarían un tiempo excesivo para el ser humano.

Esto incluye la volumetría tridimensional de tumores sólidos, la segmentación automática de cavidades cardíacas para el cálculo preciso de la fracción de eyección, la cuantificación de la carga de lesiones desmielinizantes en enfermedades neurodegenerativas o el análisis detallado de la densidad mineral ósea y la arquitectura tisular. La precisión métrica en estos escenarios es vital para el seguimiento de la respuesta al tratamiento y para la planificación quirúrgica o de radioterapia de alta complejidad.

Los estándares internacionales desarrollados por el Instituto Nacional de Excelencia en Salud y Atención del Reino Unido (National Institute for Health and Care Excellence, NICE) y el Instituto de Efectividad Clínica y Sanitaria (IECS) subrayan que la cuantificación automatizada mejora de manera drástica la reproducibilidad de los resultados clínicos, eliminando la subjetividad inherente a las mediciones manuales y permitiendo una comparación longitudinal veraz de la evolución del paciente (23,24). Para el SNS de México, esta modalidad permite estandarizar la calidad de los reportes médicos en todas las instituciones, independientemente de la carga de trabajo o de la experiencia específica del operador en técnicas de medición complejas, promoviendo así la equidad en la calidad del diagnóstico.

La implementación de la modalidad de cuantificación y segmentación exige una validación analítica extremadamente rigurosa antes de su despliegue clínico. El software debe demostrar que sus mediciones son precisas, consistentes y trazables con los estándares de referencia establecidos en la literatura médica internacional, protocolos y guías de práctica clínica vigentes. En este tipo de aplicaciones, particularmente en escenarios de medicina de precisión, podrá considerarse además el reconocimiento de evaluaciones realizadas por agencias regulatorias internacionales de alta vigilancia, conforme a los mecanismos de equivalencia previstos en el presente lineamiento.

Acto médico

El acto médico en esta modalidad se define por la interpretación experta de las métricas tecnológicas en el contexto integral de la salud del paciente, asegurando que la precisión tecnológica se traduzca de forma efectiva en una decisión clínica acertada, bioética y personalizada. La validación humana de la segmentación es un paso obligatorio, donde el médico debe verificar que los límites anatómicos propuestos por el algoritmo corresponden a la realidad clínica observada, evitando errores derivados de la automatización sin un proceso de revisión crítico.

El médico especialista, al recibir los datos cuantitativos y las segmentaciones generadas por la IA, debe integrarlos en su propia pirámide del razonamiento clínico para darles un significado terapéutico real. Una medición volumétrica, por precisa que sea, no constituye por sí sola una decisión médica; es el juicio clínico humano el que determina si un incremento milimétrico en el volumen de una lesión representa una progresión de la enfermedad que amerite un cambio radical en el esquema de quimioterapia o si entra dentro de la variabilidad biológica aceptable para ese paciente en particular.

4. Soporte propedéutico y diferencial

La cuarta modalidad representa una de las innovaciones más profundas en la asistencia clínica moderna, ya que interviene directamente en la fase de recolección de datos primarios y en la generación temprana de hipótesis diagnósticas.

Esta modalidad se orienta específicamente a la asistencia en la integración de diagnósticos clínicos complejos, la optimización de la semiología médica y la estructuración del interrogatorio y la anamnesis.

A diferencia de las modalidades previas que actúan sobre estudios paraclínicos ya generados, el soporte propedéutico actúa en tiempo real durante el encuentro paciente-profesional de la salud. Utilizando procesamiento avanzado de lenguaje natural, reconocimiento de voz y motores de inferencia clínica de alta fidelidad (sistemas expertos, diseñados para procesar grandes volúmenes de datos de pacientes en tiempo real). El sistema analiza la información vertida durante la consulta y sugiere al profesional de la salud líneas de interrogatorio específicas, síntomas acompañantes que deben ser cuestionados o explorados, así como maniobras de exploración física que podrían haber sido omitidas en el proceso inicial. Esta herramienta actúa como un consultor alrededor del proceso clínico principal que ayuda a reducir de forma significativa el sesgo de confirmación y el establecimiento prematuro del diagnóstico, asegurando que la base informativa de la pirámide del razonamiento clínico sea exhaustiva, estructurada y bien fundamentada.

Desde la perspectiva de la semiología médica, que es considerada por la UNAM como la lógica y la gramática de la medicina, el soporte propedéutico y diferencial potencia la capacidad analítica del médico sin sustituir en ningún momento su interacción humana y empática con el paciente.

En el contexto de la práctica médica en México, marcada por una alta carga asistencial en el primer nivel de atención y en la consulta externa de especialidad, esta modalidad asegura que el interrogatorio sea sistemático y que no se ignoren síntomas prodrómicos o factores de riesgo epidemiológico relevantes para la población.

Al sugerir diagnósticos clínicos diferenciales basados en la fuerza de la evidencia recolectada y en el cruce de información con bases de datos de conocimiento médico jerarquizado, la inteligencia artificial ayuda al médico a navegar por la complejidad diagnóstica con una mayor red de seguridad técnica.

Acto médico

El acto médico en esta modalidad se define por la dirección inteligente del interrogatorio y la validación experta de los signos clínicos detectados, donde el médico utiliza las sugerencias del sistema para profundizar en la historia clínica del

paciente, manteniendo siempre la cúspide del juicio clínico para integrar todos los hallazgos en un diagnóstico integral.

La implementación de esta modalidad en el Sistema Nacional de Salud tiene el potencial de elevar sustancialmente la capacidad resolutive de la medicina general y familiar, racionalizando las referencias a la consulta de especialidad y optimizando el uso de recursos paraclínicos al fundamentar la sospecha clínica inicial desde la primera entrevista.

5. Analítica predictiva de riesgos y medicina de precisión poblacional

La quinta modalidad representa un cambio de paradigma, cambiando de una medicina reactiva a una intervención proactiva y preventiva de alta precisión. Mientras que las modalidades anteriores actúan sobre el paciente que ya presenta una sintomatología o que ya se encuentra en el punto de atención, la analítica predictiva utilizará la potencia de la Base Nacional de Información en Salud para identificar a individuos o grupos poblacionales con una alta probabilidad de desarrollar complicaciones críticas o enfermedades crónicas antes de que estas se manifiesten de forma clínica evidente.

Esto incluye la predicción de descompensaciones agudas en pacientes con diabetes o hipertensión, la identificación temprana de riesgos de sepsis en áreas de hospitalización mediante el monitoreo continuo de signos vitales y la detección proactiva de brotes epidemiológicos a través de patrones de búsqueda, consumo de medicamentos o reportes de sintomatologías específicas en el primer nivel de atención.

Esta modalidad es el pilar de la salud pública moderna y se fundamenta en el derecho a la prevención y a la protección integral de la salud, permitiendo que el Estado mexicano realice intervenciones oportunas que optimicen la esperanza de vida saludable y aseguren la sostenibilidad financiera de las instituciones al evitar complicaciones costosas y discapacitantes.

Desde el punto de vista del derecho administrativo y la bioética, la analítica predictiva debe ser gestionada con una cautela extrema para evitar el determinismo algorítmico y la discriminación de pacientes basada en perfiles de riesgo probabilísticos. Cualquier hallazgo derivado de la analítica predictiva debe ser validado por un profesional de la salud antes de iniciar cualquier acción médica o administrativa que afecte al paciente.

La medicina de precisión poblacional no busca etiquetar individuos, sino identificar oportunidades de intervención preventiva que mejoren la calidad de vida. Los algoritmos predictivos deben ser transparentes en cuanto a sus variables de entrada

y deben ser auditados periódicamente para asegurar que no perpetúen sesgos socioeconómicos o geográficos.

Acto médico

El acto médico en esta modalidad se manifiesta en la interpretación del riesgo poblacional y en la implementación de planes de cuidado preventivo personalizados, donde el personal de salud utiliza la información predictiva para priorizar recursos y diseñar estrategias de salud pública que sean verdaderamente efectivas y equitativas. La soberanía de los datos personales es aquí un requisito inviolable, asegurando que la Base Nacional de Información sea utilizada exclusivamente para fines de salud pública y bienestar social.

Todo resultado de la analítica predictiva es validado por un profesional de la salud antes de cualquier acción clínica que afecte al paciente. La predicción no etiqueta ni clasifica de manera determinante a las personas; orienta oportunidades de intervención preventiva. El tratamiento de datos observará la legislación nacional aplicable.

6. Asistencia en la documentación clínica y gestión de la información

Esta modalidad aborda uno de los mayores obstáculos para la modernización y para la preservación del acto médico humano, refiriéndonos a la carga administrativa excesiva y el tiempo dedicado a la captura de datos en el expediente clínico electrónico. El uso de IA generativa, modelos de lenguaje de gran escala y reconocimiento de voz para la función de escritura digital permite que el diálogo entre el paciente y el profesional de la salud sea capturado, resumido y estructurado de manera automática en la nota médica correspondiente.

Este esquema devuelve al personal de salud su capacidad esencial de observación, contacto visual y escucha activa, elementos fundamentales de la pirámide del razonamiento clínico que a menudo se ven sacrificados por la necesidad de interactuar con la interfaz digital o el expediente clínico físico durante la consulta. Además de la documentación en tiempo real, esta modalidad asiste en la gestión del conocimiento médico al proporcionar resúmenes ejecutivos de expedientes clínicos extensos y complejos, permitiendo al médico entender la historia del paciente en cuestión de segundos. Esto garantiza la continuidad de la atención entre niveles y reduce significativamente el riesgo de errores por la omisión de antecedentes alérgicos, quirúrgicos o farmacológicos relevantes que podrían estar ocultos en un conjunto de datos no estructurados.

Protección de datos y ciberseguridad

Administrativamente, la asistencia en la documentación exige altos niveles de ciberseguridad y protección de la privacidad, dado que el sistema procesa información nominal y sensible de forma constante.

Se prohíbe el uso de estos sistemas sin la supervisión directa del personal de salud responsable, asegurando que la automatización no degrade la calidad de la información clínica ni la seguridad jurídica del paciente. Esta modalidad habilita la modernización de la atención médica, permite que la tecnología absorba las tareas repetitivas y administrativas, liberando el talento y la humanidad del médico para las tareas de la cúspide de la pirámide, es decir, juicio clínico, empatía y decisión bioética.

Validación, firma y responsabilidad

La responsabilidad del profesional de la salud es siempre el filtro final; el uso de cualquier modalidad operativa no exime al médico de su deber de cuidado ni de su obligación de actuar conforme a la *lex artis* médica. La IA en el sector salud de México es entendida como un potente auxiliar del pensamiento clínico, pero la autoridad de mando, la decisión terapéutica y la responsabilidad bioética permanecen siempre en la figura del profesional de la salud debidamente acreditado.

La validación humana de la nota generada por la IA es un requisito legal y administrativo obligatorio; el médico debe revisar, corregir y firmar el resumen para asegurar que refleje fielmente la realidad del acto médico y las indicaciones terapéuticas. La responsabilidad jurídica del contenido del expediente clínico sigue siendo del profesional de la salud que lo suscribe con su cédula profesional.

Recursos humanos para la salud

La formación de los recursos humanos para la salud debe evolucionar para que el personal sea capaz de navegar con pericia en este nuevo ecosistema digital. Las facultades de medicina y los consejos de especialidades tienen el reto de integrar en sus programas de estudio no solo la ciencia médica tradicional, sino la competencia clínica digital que permita al médico interrogar al algoritmo, detectar sus fallas y validar sus sugerencias con un criterio científico sólido. La alfabetización digital avanzada es hoy una necesidad básica de la competencia profesional.

El fortalecimiento del capital humano es la única garantía de que la pirámide diagnóstica técnica y la pirámide del razonamiento clínico humano converjan de manera segura y eficiente en beneficio de la sociedad mexicana. La modernización busca potenciar las capacidades del médico para que pueda brindar una atención de

excelencia, basada en la ciencia más avanzada pero siempre imbuida de la calidez y el juicio humano que definen a la medicina como una disciplina humanista.

Gobernanza de la implementación

La gobernanza institucional para la implementación de estas seis modalidades operativas debe centralizarse en un proceso administrativo de autorización riguroso y transparente en cada Institución del SNS. Toda unidad de salud que pretenda implementar alguna de estas herramientas debe contar con un expediente técnico y bioético completo que incluya la descripción pormenorizada del software, sus certificaciones de calidad nacionales o internacionales, el protocolo de validación local y el programa de capacitación continua para el personal médico, paramédico y de enfermería.

La evaluación técnica, operativa y funcional de las modalidades de implementación contempladas en el presente lineamiento corresponderá al Comité de Evaluación de Tecnologías de la Dirección General de Modernización del Sector Salud, en coordinación con las áreas competentes de las instituciones participantes y sin perjuicio de las atribuciones conferidas a otras autoridades.

En caso de que se trate de sistemas que no han sido aprobados de forma centralizada, se recomienda que las instituciones consideren la opinión de las instancias colegiadas competentes en materia de bioética, cuando la naturaleza de la tecnología o su impacto clínico así lo ameriten, con el propósito de fortalecer la protección de los derechos de las personas usuarias y favorecer la observancia de los protocolos, procesos, normas y guías de atención correspondientes.

Este enfoque busca favorecer una implementación ordenada, trazable y basada en evidencia de las tecnologías emergentes en salud, promoviendo mecanismos de rendición de cuentas, ciberseguridad, protección de datos, soberanía tecnológica y principios bioéticos aplicables durante todo el ciclo de vida de los sistemas implementados.

El presente lineamiento establece que toda la información procesada por sistemas de IA en las instituciones públicas debe permanecer bajo el control soberano del Estado mexicano, prohibiendo la transferencia de datos nominales a nubes que no cumplan con la legislación nacional en materia de protección de datos personales.

La transparencia algorítmica es un derecho del paciente y del médico, por lo que los sistemas deben ser auditables y explicables en su lógica de funcionamiento.

Las seis modalidades operativas de implementación clínica aquí detalladas representan el marco de referencia integral para la salud digital en México. Al regular

desde el triage hasta la asistencia propedéutica y la analítica predictiva, este lineamiento brinda certeza jurídica, seguridad clínica y eficiencia operativa a todo el Sistema Nacional de Salud.

El compromiso del Estado con la modernización se manifiesta en una regulación que prioriza la innovación tecnológica pero que se mantiene firme en la protección del acto médico y en el respeto a la dignidad de la persona humana. La IA es una herramienta que nos permite aspirar a un sistema para la salud más equitativo, preciso y humano, donde la tecnología esté siempre al servicio de la vida y bajo la rectoría del conocimiento clínico y la responsabilidad bioética de sus profesionales de la salud.

Capítulo V. Base Nacional de Información en Salud e Interoperabilidad de Sistemas

La integración de los Sistemas de Soporte para la Decisión Clínica (SSDC) y de otras tecnologías basadas en IA en el SNS deberán concebirse como parte de una infraestructura estratégica de gobernanza de datos en salud. Su implementación requiere la construcción de un repositorio soberano de activos digitales que incluya conjuntos de datos de entrenamiento etiquetados, modelos fundacionales validados para la población mexicana y entornos de prueba regulatoria que permitan evaluar el funcionamiento de los sistemas previo a su despliegue clínico.

Este enfoque encuentra sustento en el Capítulo VI Bis de la Ley General de Salud, conforme a la reforma de 2026, mediante la cual la salud digital adquiere el carácter de obligación legal para los integrantes del SNS(1). En particular, el artículo 71 Ter establece el mandato de digitalizar la información médica y garantizar su interoperabilidad entre prestadores de servicios de salud, públicos y privados. Bajo esta premisa, la interoperabilidad deja de ser una condición técnica para convertirse en un requisito de cumplimiento normativo vinculante.

Este capítulo establece las obligaciones que los sistemas de inteligencia artificial deben observar en materia de interoperabilidad y calidad del dato clínico para su implementación en el Sistema Nacional de Salud, con el fin de asegurar la continuidad de la atención, la trazabilidad, la calidad de la información y la equidad.

En cumplimiento a lo establecido en la LGS, los desarrolladores y prestadores de tecnologías basadas en IA deberán observar las siguientes directrices:

Garantía de continuidad: Los sistemas de IA deberán ser capaces de procesar y devolver datos en formatos que aseguren la continuidad del expediente clínico del paciente, independientemente de la institución donde se originó el dato.

Portabilidad del dato clínico: Los sistemas deberán permitir la exportación de resultados y hallazgos en estándares abiertos, a fin de que el paciente pueda ejercer su derecho a la portabilidad conforme a la reforma vigente.

Digitalización nativa: Se prohíbe el uso de sistemas de IA que dependan exclusivamente de entradas analógicas o manuales no estandarizadas. Las instituciones deberán promover la adopción de registros digitales nativos conforme al artículo 71 Bis de la Ley General de Salud.

Declaración de metadatos: Los responsables del sistema deberán documentar el origen de los datos de entrenamiento y su concordancia con las unidades del Sistema Internacional de Medidas.

Validación de población: El modelo deberá demostrar que su precisión se mantiene estable al ser evaluado contra datos de validación representativos de la población objetivo, con el fin de prevenir el sesgo algorítmico.

Con el propósito de garantizar la compatibilidad técnica entre sistemas existentes y nuevos, se establece la siguiente jerarquía de cumplimiento en materia de estandarización:

- **Nivel físico:** cumplimiento de la Regla General 16 de la (FEUM) en lo relativo a la clasificación de dispositivos médicos con base en su nivel de riesgo sanitario.
- **Nivel semántico:** Define la codificación de los conceptos clínicos para preservar su significado entre sistemas. Los sistemas deberán emplear las terminologías de referencia adoptadas por la normativa y los catálogos aplicables, entre ellas SNOMED CT para hallazgos, diagnósticos y procedimientos, y LOINC para observaciones de laboratorio y estudios clínicos.
- **Nivel de Interoperabilidad de intercambio:** Define la estructura y la transmisión de la información entre sistemas como condición de buena práctica, se recomienda que los sistemas de inteligencia artificial admitan la especificación HL7 FHIR (Fast Healthcare Interoperability Resources) para el intercambio, por ser el estándar de adopción internacional vigente, en la medida en que resulte compatible con la norma aplicable.
- **Nivel de desempeño:** los sistemas deberán reportar su tasa de error y la detección de deriva algorítmica mediante protocolos de telemetría segura, en los términos establecidos en el presente lineamiento.

Todo sistema de IA que opere en el SNS deberá contar con documentación técnica estructurada disponible para las instituciones, los comités de bioética y los responsables de la toma de decisiones, con el propósito de evaluar la idoneidad, seguridad y pertinencia de cada herramienta antes de su despliegue clínico, esenciales para garantizar la transparencia, la trazabilidad y el uso responsable de los sistemas de IA en salud.

Este proceso permite comprender el origen, las características y las limitaciones tanto de los conjuntos de datos como de los modelos, facilitando su evaluación por parte de equipos técnicos, clínicos y reguladores, y asegurando que las decisiones se basen en información verificable. Por lo que se deberá considerar lo siguiente:

- **DataSheet (dataset):** Ficha técnica u hoja de datos que describe las especificaciones, explica el rendimiento y las características técnicas de un sistema, el cual debe describir el conjunto de datos utilizados, incluyendo su origen y periodo de recogida, los criterios de inclusión y exclusión aplicados, los procesos de limpieza y etiquetado, las variables y codificaciones empleadas, así como la representatividad del dataset respecto a la población de referencia. También debe detallar las limitaciones conocidas y las condiciones de acceso y uso, de forma que cualquier reutilización se realice bajo un marco controlado y conforme a la normativa aplicable.
- **Tarjeta de modelo (Model Card):** Debe documentar el modelo entrenado, especificando el objetivo clínico y la finalidad prevista, la población de referencia, la versión y fecha de liberación, los datos y procesos de preprocesamiento utilizados y las métricas de rendimiento. Además, debe reflejar los resultados por subgrupos relevantes, los riesgos identificados y las salvaguardas implementadas, así como los usos permitidos y prohibidos. Finalmente, debe incluir los requisitos de monitorización y mantenimiento, indicando cómo se controlará el rendimiento en producción y los criterios para su actualización.
- **Actualización y control de cambios:** Garantiza un versionado sincronizado entre dataset y la tarjeta de modelo, con criterios objetivos para identificar cambios significativos. Además, mantiene un registro de versiones que documente la correspondencia entre dataset, tarjeta modelo y servicio desplegado, incluyendo fecha, responsable, motivo del cambio y huella digital. Asimismo, establece un plan de comunicación que asegure que usuarios y evaluadores reciban información clara y oportuna sobre cualquier modificación relevante (38).

La interoperabilidad técnica deberá vincularse de manera inseparable con la calidad y representatividad del dato. Los sistemas deberán ser capaces de capturar, procesar y reportar información desagregada por variables sociodemográficas relevantes, incluyendo sexo, edad, pertenencia étnica, distribución geográfica y nivel socioeconómico, a fin de permitir auditorías de equidad e identificación de sesgos en el rendimiento de los modelos. La ausencia de este nivel de granularidad expone al Sistema Nacional de Salud al riesgo de reproducir o amplificar desigualdades preexistentes bajo una apariencia de neutralidad tecnológica.

En consecuencia, la interoperabilidad deberá entenderse como una condición estructural para garantizar continuidad clínica, trazabilidad, calidad del dato, evaluación de desempeño y equidad en la implementación de tecnologías basadas en IA dentro del SNS.

Capítulo VI. Clasificación por riesgo

La clasificación por riesgo de los SSDC y de las tecnologías basadas en IA, aplicadas en salud se fundamenta en los principios de integridad de los datos, proporcionalidad regulatoria y protección de la seguridad del paciente. Este esquema es consistente con los marcos internacionales de regulación de SaMD, particularmente con las directrices del *International Medical Device Regulators Forum* (IMDRF) y con el marco regulatorio de la FDA para tecnologías basadas en IA (13,14,15).

La validez e integridad del dato constituyen un requisito previo para el funcionamiento de cualquier modelo algorítmico. En consecuencia, toda entrada, proceso de entrenamiento y ejecución deberá garantizar coherencia semántica, calidad técnica y correcta estandarización de unidades conforme al Sistema Internacional de Unidades, en concordancia con la Regla General 16 de la Farmacopea de los Estados Unidos Mexicanos (FEUM) en lo relativo a los datos de laboratorio y medición clínica (16). El presente apartado establece el régimen de clasificación por riesgo aplicable a los sistemas de inteligencia artificial en salud sujetos al Lineamiento, así como los criterios, la matriz de clasificación, las reglas de articulación con el régimen sanitario de dispositivos médicos y los requisitos diferenciados por nivel de riesgo.

Sistemas excluidos

Quedan excluidos del régimen de clasificación por riesgo del presente Capítulo:

- I. Las aplicaciones de bienestar general.
- II. Los sistemas de soporte no clínico.
- III. Los sistemas destinados exclusivamente a investigación científica preclínica o clínica, cuando no se utilicen en la prestación de servicios asistenciales y se encuentren sujetos a los marcos regulatorios específicos de investigación en salud.
- IV. Los sistemas que operan exclusivamente sobre datos disociados o sintéticos sin posibilidad de reidentificación, cuando su uso se limite a finalidades de investigación, desarrollo, validación interna o capacitación, y no se aplique a pacientes en el marco de la prestación de servicios.

La exclusión del régimen de clasificación de riesgo no exime a estos sistemas del cumplimiento de otras disposiciones que les sean aplicables, particularmente las relativas a la protección de datos personales, ética en la investigación, ciberseguridad y propiedad intelectual.

Criterios y matriz de clasificación

La clasificación de los SSDC y demás tecnologías basadas en IA articula dos ejes complementarios y una taxonomía de autonomía, que en conjunto orientan las exigencias regulatorias y de gobernanza proporcionales al riesgo:

Criterios primarios

La clasificación se determina por el cruce de dos criterios primarios:

1. Significancia de la información para la decisión clínica, con los siguientes niveles:

- a) **Nivel A. Informar.** El sistema aporta información que el profesional de la salud puede considerar como uno de varios insumos en su proceso de decisión, sin que la información del sistema condicione la línea de actuación clínica.
- b) **Nivel B. Impulsar.** El sistema aporta información que orienta el manejo clínico del paciente y que, sin determinar por sí sola la decisión, influye en la elección entre alternativas terapéuticas o diagnósticas.
- c) **Nivel C. Determinar.** El sistema realiza la función diagnóstica o terapéutica de manera directa, o aporta información cuya naturaleza es tal que el manejo clínico subsecuente queda definido por la salida del sistema. Este nivel corresponde a los supuestos en que la decisión clínica depende primordialmente de la salida del sistema, en contraste con los niveles A y B, en que el profesional puede revisar de manera independiente los fundamentos de la información y no depende primordialmente de ella.

2. Estado de la condición sanitaria del paciente, con los siguientes niveles:

- a) **Nivel 1. No seria:** Condición que no representa amenaza para la vida ni riesgo de daño permanente o de discapacidad significativa.
- b) **Nivel 2. Seria:** Condición que, sin manejo oportuno y adecuado, puede progresar hacia daño significativo, discapacidad o compromiso funcional relevante, sin que haya una amenaza inmediata para la vida.
- c) **Nivel 3. Crítica:** Condición que representa una amenaza para la vida, riesgo de daño irreversible o necesidad de intervención clínica inmediata.

La matriz se construye sobre el principio de que el riesgo crece de manera monótona conforme aumenta cualquiera de los dos criterios primarios.

Tabla 1. Matriz primaria de clasificación

Significancia ↓ / Condición →	Nivel 1. No seria	Nivel 2. Seria	Nivel 3. Crítica
Nivel A. Informar	Bajo	Bajo	Moderado
Nivel B. Impulsar	Bajo	Moderado	Alto
Nivel C. Determinar	Moderado	Alto	Alto

Fuente: DGMoSS, 2026.

Criterio modulador.

El criterio modulador toma en cuenta el nivel de autonomía operativa y modula el nivel de riesgo dentro del rango establecido por los criterios primarios. Un sistema con mayor autonomía operativa requiere controles regulatorios proporcionalmente más estrictos. Se distinguen tres niveles:

- **Copiloto analítico (Tipo I):** Sistema que provee información o sugerencias bajo supervisión clínica permanente del profesional, quien revisa cada salida y decide su incorporación al razonamiento clínico.
- **Autonomía delegada (Tipo II):** Sistema que opera con mayor independencia en tareas definidas, con revisión humana retrospectiva o por excepción y con mecanismos de interrupción humana disponibles en todo momento.
- **Sistema autónomo (Tipo III):** Sistema que ejecuta funciones sin intervención humana en cada caso individual. La supervisión humana se concentra en el diseño, validación y monitoreo del sistema. Este tipo de operación se admite únicamente en sistemas de muy estrecho propósito previsto, con validación clínica robusta y bajo marcos de responsabilidad compartida entre el desarrollador y la institución usuaria.

Reglas de modulación por autonomía

La modulación del nivel de riesgo por el criterio modulador (autonomía operativa) se aplica conforme a las siguientes reglas:

- I. Los sistemas Tipo I conservan el nivel de riesgo asignado por la matriz primaria sin modulación
- II. Los sistemas Tipo II en celdas de nivel moderado pueden ascender a riesgo alto cuando no existan mecanismos efectivos de revisión humana por excepción documentados.
- III. Los sistemas Tipo III quedan clasificados como de riesgo alto con independencia de la celda primaria, y se sujetan a las disposiciones especiales aplicables.

Reglas de resolución

Para resolver casos limítrofes o de uso heterogéneo se aplican las siguientes reglas:

- I. Prevalencia del propósito declarado: El sistema se clasifica conforme al propósito de uso declarado. Cuando el uso institucional efectivo exceda dicho propósito, la clasificación atenderá al uso efectivo y no al declarado.
- II. Condición más crítica del intervalo poblacional: Cuando la población objetivo comprenda casos con criticidad heterogénea, la clasificación se realiza sobre la base del intervalo más crítico contemplado en el uso previsto.
- III. Modalidad operativa diferenciada: Cuando un mismo sistema admite modalidades operativas distintas, cada modalidad se evalúa por separado y el sistema queda sujeto a las exigencias correspondientes al uso más crítico.

Articulación con el régimen sanitario

El nivel de riesgo establecido en este Capítulo no sustituye la clasificación que la Comisión Federal para la Protección contra Riesgos Sanitarios otorga a los dispositivos médicos. Que un sistema sea considerado "dispositivo médico" (y por lo tanto requiera control y Registro Sanitario) depende exclusivamente de las definiciones de la normativa sanitaria vigente, sin importar el riesgo que le asigne este Lineamiento.

Por lo tanto, un sistema puede tener un nivel de riesgo de acuerdo a este documento sin ser un dispositivo médico. La siguiente tabla de equivalencias es solo orientativa y aplica exclusivamente a los sistemas que sí califiquen como dispositivos médicos:

Tabla 2. Equivalencias con dispositivos médicos

Nivel de riesgo del Lineamiento	Clase de dispositivo médico (COFEPRIS)	Régimen sanitario aplicable
Bajo	Clase I	Régimen ordinario de dispositivos médicos clase I, cuando el sistema constituya un dispositivo médico.
Moderado	Clase II	Régimen ordinario de dispositivos médicos clase II, cuando el sistema constituya un dispositivo médico.
Alto	Clase III	Régimen ordinario de dispositivos médicos clase III, cuando el sistema constituya un dispositivo médico.

Fuente: Datos extraídos Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) 2023.

Cuando un sistema constituya el dispositivo médico, las disposiciones del presente Lineamiento se aplicarán adicionalmente al régimen sanitario correspondiente, sin sustituirlo ni alterar las atribuciones de la autoridad sanitaria.

Disposiciones especiales

Modelos de inteligencia artificial de propósito general.

La clasificación de un modelo de inteligencia artificial de propósito general utilizado en el Sistema Nacional de Salud se realiza sobre el uso institucional específico, no sobre el modelo subyacente. La institución usuaria declara el uso previsto y aplica los criterios primarios del presente Capítulo al uso declarado.

Adicionalmente al nivel de riesgo que corresponda, los modelos de propósito general utilizados en contextos clínicos quedan sujetos a las siguientes obligaciones específicas:

- I. Declaración institucional documentada del uso previsto.
- II. Prohibición de uso fuera del propósito declarado, con controles institucionales que la hagan efectiva.
- III. Verificación humana obligatoria de toda salida con efecto clínico.
- IV. Advertencia explícita al usuario sobre la naturaleza del propósito general del modelo y sus limitaciones.
- V. Restricción para decisiones determinantes (Nivel C) en condiciones serias o críticas sin validación específica documentada.
- VI. Régimen reforzado de protección de datos personales conforme a la modalidad de implementación tecnológica.

Los sistemas de inteligencia artificial en salud desarrollados por instituciones del Sistema Nacional de Salud para uso interno se sujetan al régimen de clasificación por riesgo del presente Capítulo. La institución desarrolladora documenta la validación interna conforme a criterios proporcionales al nivel de riesgo asignado y notifica el sistema a la instancia institucional correspondiente cuando se clasifique como de riesgo alto.

Requisitos diferenciados por nivel de riesgo

Los pilares transversales son seis: gestión de sesgos, ciberseguridad y soberanía de datos, transparencia y explicabilidad, documentación técnica, vigilancia post-implementación, y rendición de cuentas. La intensidad de aplicación de cada pilar se modula según el nivel de riesgo

La clasificación de riesgo determina el nivel de exigencia regulatoria que aplica a cada sistema, pero existen requisitos cuya exigencia es transversal: aplican a todo sistema de inteligencia artificial en salud sujeto a la siguiente tabla.

Tabla 3. Pilares transversales por nivel de riesgo

Pilar	Riesgo bajo	Riesgo moderado	Riesgo alto
Gestión de sesgos	Caracterización general de datos y declaración de población representada.	Lo anterior, más desempeño desagregado por subpoblaciones y mitigaciones documentadas.	Lo anterior, más validación clínica con datos representativos de la población mexicana y revisión continua.
Ciberseguridad y datos	Controles básicos de seguridad, cumplimiento de protección de datos personales.	Lo anterior, más controles reforzados y declaración de modalidad de implementación.	Lo anterior, más continuidad operativa, plan de contingencia y restricciones por soberanía de datos.
Transparencia y explicabilidad	Declaración de uso previsto y limitaciones.	Lo anterior, más explicabilidad operativa de variables determinantes.	Lo anterior, más explicabilidad clínica y documentación del razonamiento del sistema.
Documentación técnica	Tarjeta de Modelo básica.	Tarjeta de Modelo completa con métricas desagregadas.	Tarjeta de Modelo completa, validación clínica y expediente técnico institucional.
Vigilancia post implementación	Registro básico de incidencias.	Monitoreo periódico de desempeño y calibración.	Monitoreo continuo, reporte anual de desempeño y reportabilidad reforzada de eventos adversos.
Rendición de cuentas	Designación de responsable técnico institucional.	Lo anterior, más responsabilidad compartida desarrollador institución documentada.	Lo anterior, más plan de retiro y mecanismos reforzados de auditabilidad.

Fuente: Elaboración propia DGMoSS, 2026.

Disposiciones específicas para riesgo alto

Los sistemas clasificados como de riesgo alto se sujetan a:

- I.** Validación clínica documentada con datos representativos de la población mexicana objetivo.
- II.** Reporte anual de desempeño a la instancia institucional correspondiente.
- III.** Reportabilidad reforzada de eventos adversos, conforme al régimen de tecnovigilancia aplicable.
- IV.** Plan de retiro documentado, que comprenda condiciones de activación, mecanismos de transición y continuidad asistencial.

Capítulo VII. Mecanismos de Reconocimiento y Validación

La integración de tecnologías de IA y SaDM en el SNS se fundamenta en un modelo que articula la eficiencia administrativa con el rigor científico, reconociendo la interdependencia de los marcos regulatorios a nivel global. El mecanismo de reconocimiento de equivalencia, denominado *reliance*, se define como el acto administrativo y técnico mediante el cual la autoridad sanitaria nacional asume la validez de las evaluaciones de seguridad, eficacia y desempeño realizadas por agencias regulatorias de referencia internacional, como fundamento para agilizar o sustentar el proceso de registro sanitario en México. Este enfoque cuenta con el respaldo metodológico de la Organización Mundial de la Salud y del *International Medical Device Regulators Forum* (IMDRF), y permite que la innovación tecnológica sea incorporada de manera oportuna sin comprometer los estándares nacionales de protección a la salud (13,14). El alcance del mecanismo no se agota en la obtención del registro sanitario, sino que marca el inicio de una fase crítica de supervisión operativa y técnica, en la que deberá coordinarse el monitoreo del desempeño clínico y la integridad algorítmica en las unidades de salud.

Autoridades regulatorias de referencia.

El sustento jurídico de este mecanismo en el derecho administrativo mexicano emana del mandato constitucional de garantizar el derecho a la protección de la salud mediante el acceso a los avances científicos, conforme al artículo 4° de la Constitución Política de los Estados Unidos Mexicanos y a la reforma a la Ley General de Salud de enero de 2026 (1). Para efectos del lineamiento, se reconocen como agencias regulatorias de referencia para la aplicación del mecanismo de *reliance* la *Food and Drug Administration* (FDA) de los Estados Unidos de América, la *European Medicines Agency* (EMA) y la *Medicines and Healthcare products Regulatory Agency* (MHRA) del Reino Unido, sin perjuicio de que la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) pueda ampliar esta lista mediante resolución técnica fundada.

El presente capítulo establece los mecanismos de reconocimiento y validación aplicables a la incorporación de tecnologías basadas en inteligencia artificial, incluido el software como dispositivo médico (Software as a Medical Device, SaMD), al Sistema Nacional de Salud, así como el régimen de gestión de sus modificaciones posteriores a la autorización.

Efectos del reconocimiento de equivalencia.

El reconocimiento de equivalencia no exime a la tecnología del registro sanitario ante COFEPRIS, pero habilita un procedimiento abreviado de evaluación en el que la documentación de la agencia de origen sustituye total o parcialmente los estudios de sustento requeridos en el proceso ordinario. Una vez otorgado el registro, la tecnología se incorpora al ecosistema digital del SNS bajo un esquema de monitoreo de ciclo de vida que considera tanto la estabilidad del software como su adaptación a las condiciones de la población nacional.

A diferencia de tecnologías con arquitectura estática, los algoritmos de aprendizaje automático poseen una naturaleza dinámica susceptible de presentar deriva algorítmica, entendida como la pérdida de desempeño del modelo al interactuar con datos distintos a los de su entrenamiento original. En consecuencia, el reconocimiento de equivalencia obliga al establecimiento de un monitoreo continuo que verifique si el dispositivo mantiene los parámetros de sensibilidad y especificidad que justificaron su registro. Esta vigilancia constituye un proceso colaborativo en el que las unidades de salud informan sobre el desempeño real de la herramienta, permitiendo que la autoridad competente identifique patrones de error o degradación que pudieran representar un riesgo para la seguridad clínica en el territorio nacional.

Vigilancia posterior a la autorización.

En el marco de la vigilancia post-autorización, las empresas fabricantes y distribuidoras asumen responsabilidades jurídicas y técnicas intransferibles. Dichas responsabilidades comprenden: la obligación de mantener un plan de seguimiento clínico que recolecte y analice de forma sistemática los datos de uso real en las unidades de salud mexicanas; la habilitación de canales de comunicación bidireccional para el reporte inmediato de cualquier mal funcionamiento, falla algorítmica o incidente adverso; la garantía de integridad de las actualizaciones del software (27); y la protección permanente de los datos personales de los pacientes. La ausencia de un sistema de vigilancia proactivo por parte del proveedor contraviene las disposiciones del presente lineamiento y podrá derivar en la revocación del permiso de uso institucional, con independencia de la vigencia del registro sanitario.

Las instituciones de salud podrán considerar la participación de las instancias colegiadas competentes en materia de bioética como mecanismos de acompañamiento, orientación, educativas o consultivas para la identificación de posibles impactos relacionados con el uso de tecnologías basadas en inteligencia artificial, particularmente en aspectos vinculados con la equidad en la atención, el respeto a la autonomía de las personas usuarias y la interacción entre el profesional de la salud y los sistemas tecnológicos.

En su caso, dichas instancias podrán emitir recomendaciones no vinculantes respecto a posibles sesgos algorítmicos, riesgos bioéticos o áreas de mejora identificadas durante la operación de estas tecnologías.

La integración, análisis y seguimiento de hallazgos derivados del monitoreo y vigilancia de tecnologías basadas en IA corresponderá al Comité especializado de cada institución, conforme a los mecanismos institucionales establecidos sin perjuicio de las atribuciones de otras autoridades competentes.

La gestión de modificaciones al software tras su autorización constituye uno de los componentes técnicos más relevantes de la vigilancia post-autorización, por lo que se deberán considerar como modificaciones menores aquellas que comprenden correcciones de errores de programación que no alteran la función médica prevista, actualizaciones de seguridad informática y mejoras en la interfaz de usuario que no modifiquen la interpretación de los datos clínicos.

Estas modificaciones deberán gestionarse mediante un protocolo de notificación administrativa que permita su implementación oportuna, asegurando que las unidades de salud operen con las versiones más estables y seguras del sistema. Para el caso de modificaciones consideradas mayores serán aquellas que impliquen una alteración sustantiva en el uso previsto del dispositivo, un cambio en la arquitectura del algoritmo de decisión, o el reentrenamiento del modelo con nuevas fuentes de datos que modifiquen de manera significativa su desempeño clínico. Estas modificaciones permiten que la instancia institucional de evaluación interna lo catalogue como mayor si detecta un impacto clínico relevante, y no precisamente estará vinculado al registro sanitario original. La responsabilidad de clasificar la magnitud de cada modificación recae en el fabricante, quien deberá notificar a las autoridades competentes y a las instituciones usuarias sobre cualquier cambio que afecte las funciones de soporte a la decisión clínica. Ninguna modificación mayor podrá implementarse bajo la apariencia de una actualización rutinaria.

La clasificación de cada modificación corresponde, en el plano del producto, al fabricante, quien notificará a las autoridades competentes y a las instituciones usuarias cualquier cambio que afecte las funciones de soporte a la decisión clínica. Dicha clasificación no vincula a la institución: la instancia institucional de evaluación prevista en el Capítulo II podrá tratar como mayor para efectos internos, una modificación con impacto clínico relevante. Ninguna modificación mayor podrá implementarse bajo la apariencia de una actualización rutinaria.

Tratándose de sistemas registrados mediante reconocimiento de equivalencia, el titular del registro mantendrá la correspondencia documental entre la versión autorizada por la autoridad de referencia y la versión en operación en México, a fin de

preservar el presupuesto de identidad de características esenciales en que se funda el reconocimiento.

La variabilidad de los entornos operativos en las unidades de salud del SNS incluyendo diferencias en la calidad de la infraestructura, la interoperabilidad con los sistemas de información hospitalaria y la heterogeneidad de los protocolos de atención locales justifica la necesidad de una vigilancia robusta y diferenciada. La Secretaría de Salud deberá utilizar los datos generados por este sistema de vigilancia para emitir recomendaciones técnicas orientadas a mejorar el uso de la IA en el SNS y para retroalimentar a COFEPRIS sobre la seguridad de las tecnologías autorizadas mediante *reliance*.

Responsabilidades del fabricante y del proveedor.

El esquema de vigilancia post-autorización establecido en el presente capítulo tiene por objeto garantizar que el reconocimiento de equivalencia internacional no sea un mecanismo de ingreso sin supervisión, sino el punto de partida de un proceso continuo de evaluación del desempeño en condiciones reales. La responsabilidad compartida entre fabricantes, autoridades regulatorias e instituciones de salud constituye el fundamento de la seguridad jurídica y técnica necesaria para que la incorporación de tecnologías basadas en inteligencia artificial se traduzca en beneficios clínicos comprobables para la población del SNS.

En el marco de la vigilancia posterior a la autorización, los fabricantes, distribuidores y proveedores de tecnologías basadas en inteligencia artificial asumen, frente a las instituciones del Sistema Nacional de Salud en que sus sistemas operen, cuando menos las responsabilidades siguientes:

- Mantener un plan de seguimiento clínico que recolecta y analiza de forma sistemática los datos de desempeño en condiciones reales de uso en las unidades de salud mexicanas en que el sistema opere;
- Habilitar canales de comunicación bidireccional para el reporte y la atención oportuna de fallas, malfuncionamientos, degradación del desempeño e incidentes adversos;
- Garantizar la integridad, trazabilidad y documentación de las actualizaciones del software.
- Cumplir las obligaciones aplicables en materia de protección de datos personales durante todo el ciclo de vida del sistema.

El incumplimiento de estas responsabilidades habilita a la institución a restringir, suspender o retirar el sistema, mediante la instancia institucional de evaluación prevista.

Validación técnica formal para la reactivación de sistemas suspendidos.

La reactivación de un sistema suspendido requerirá una validación técnica formal, coordinada por la instancia institucional de evaluación prevista, que comprenderá cuando menos: la identificación y corrección documentada de la causa que motivó la suspensión; la verificación de que el desempeño del sistema satisface los umbrales del plan institucional de monitoreo y la integración de la evidencia correspondiente al expediente técnico institucional.

Cuando la suspensión derive de una modificación del sistema, la reactivación se sujetará además al régimen de requisitos transversales de la vía administrativa. Cuando la tecnología constituya un dispositivo médico, la validación institucional se entenderá sin perjuicio de las obligaciones que correspondan ante la autoridad sanitaria. La decisión de reactivación se documentará en el expediente técnico institucional y se notificará a la Secretaría de Salud cuando la suspensión hubiere sido notificada.

Capítulo VIII. Vigilancia y Seguimiento

La implementación institucional de tecnologías basadas en IA en el SNS podrá sujetarse a mecanismos institucionales de evaluación técnica, operativa y bioética de cada sistema antes de su despliegue clínico. Esta vía tiene por objeto asegurar que la adopción tecnológica ocurra de manera ordenada, transparente y con plena responsabilidad institucional, bajo el marco normativo de la reforma a la Ley General de Salud de enero de 2026 (1) y conforme a los principios bioéticos de autonomía, beneficencia, no maleficencia y justicia distributiva.

Para efectos operativos, la vía administrativa se estructura en las siguientes fases secuenciales y obligatorias:

Fase 1. Evaluación institucional

Cuando la naturaleza de la tecnología, su impacto clínico o sus implicaciones sobre derechos de las personas usuarias así lo ameriten, las instituciones podrán solicitar la opinión de las instancias colegiadas competentes con el propósito de fortalecer el análisis de aspectos relacionados con autonomía, transparencia, explicabilidad, equidad y supervisión humana efectiva. Estas opiniones tendrán carácter orientador y no sustituirán las atribuciones de las autoridades sanitarias competentes ni de los mecanismos institucionales de evaluación tecnológica.

La evaluación técnica, operativa y funcional de las tecnologías contempladas en el presente lineamiento corresponderá al Comité especializado de cada institución en coordinación con las áreas institucionales competentes de la Secretaría de Salud.

Fase 2. Validación técnica de interoperabilidad y seguridad

La segunda fase consiste en la verificación de que el sistema cumple con los estándares de interoperabilidad establecidos en el Capítulo V y con los requisitos de seguridad de la información aplicables. En esta etapa deberán auditarse los protocolos de cifrado, los mecanismos de control de acceso, la trazabilidad de los datos procesados y la compatibilidad con el expediente clínico electrónico institucional. Esta validación tiene por objeto garantizar que la incorporación del sistema contribuya a la integración del SNS y no a su fragmentación informática.

Fase 3. Consentimiento informado digital

Las instituciones deberán incorporar en el proceso de atención una cláusula específica de consentimiento informado que notifique al paciente o a su representante legal sobre el uso de sistemas de soporte algorítmico en su diagnóstico o tratamiento. Dicha cláusula deberá redactarse en lenguaje accesible, precisando que la tecnología actúa como auxiliar del médico y que la decisión clínica final es responsabilidad exclusiva del profesional de la salud. El consentimiento deberá dejar constancia de que el paciente

tiene derecho a solicitar una explicación sobre cualquier sugerencia emitida por el sistema y de que su negativa a ser evaluado mediante estas herramientas no condicionará la prestación de los servicios de salud. Este consentimiento deberá armonizarse con los requisitos establecidos en la NOM-004-SSA3 relativa al expediente clínico y con la Ley Federal de Protección de Datos Personales en Posesión de los Particulares, según corresponda.

Fase 4. Vigilancia continua y reporte de incidentes

La vía administrativa no concluye con el inicio de operaciones del sistema, sino que se extiende a lo largo de todo su ciclo de vida. Las unidades de salud deberán reportar periódicamente el desempeño de los sistemas de inteligencia artificial conforme a lo establecido en este documento, e informar de manera inmediata cualquier desviación técnica o bioética detectada, incluyendo la aparición de sesgos discriminatorios o errores sistemáticos. El responsable de la unidad de salud deberá garantizar la existencia de protocolos de contingencia ante fallos del sistema, de modo que el acto médico no se vea interrumpido ni degradado por problemas de infraestructura o conectividad. Las omisiones en la supervisión de estas tecnologías quedan sujetas al marco sancionatorio de la Ley General de Responsabilidades Administrativas.

Evaluación periódica

Se realizará con una frecuencia modulada conforme a la clasificación de riesgo propuesta en este documento, para los sistemas de riesgo alto, cuando menos semestral; para los sistemas de riesgo moderado y bajo, con la periodicidad que determine el plan institucional de monitoreo, bajo el criterio de que a mayor riesgo corresponde mayor frecuencia. Adicionalmente, se realizará evaluación extraordinaria cuando se detecten variaciones significativas en las características de la población atendida, en los datos de entrada o en las prácticas clínicas institucionales.

Los procedimientos de evaluación contemplarán la recalibración o actualización del modelo cuando sea técnicamente necesario, así como criterios cuantitativos predefinidos para su suspensión en caso de pérdida significativa de desempeño. Las modificaciones derivadas de este proceso se clasificarán y gestionarán conforme al régimen de modificaciones menores y mayores establecido en el Capítulo VII, que es rector en la materia; al presente capítulo corresponde únicamente la detección de la necesidad de modificación y la activación institucional del proceso.

Vigilancia de eventos adversos y seguridad

Las instituciones deberán implementar un sistema de vigilancia que permita la identificación, registro y análisis de eventos adversos, errores del sistema, sesgos detectados y cualquier situación que pueda comprometer la seguridad del paciente o la calidad de la atención (6,12,32,33). Para ello, deberán habilitarse canales formales de notificación accesibles al personal de salud, con procedimientos institucionales definidos para el análisis, clasificación y escalamiento de los eventos identificados. Los

reportes de eventos adversos graves deberán remitirse a la Comisión Federal para la Protección contra Riesgos Sanitarios (COFEPRIS) y a la Secretaría de Salud en un plazo no mayor a 72 horas a partir de su identificación, conforme a los mecanismos de tecnovigilancia que establezca la autoridad sanitaria competente.

Requisitos transversales de la vía administrativa

Con independencia de la fase en que se encuentre el proceso, los siguientes requisitos son de observancia permanente:

- **Transparencia algorítmica:** Los proveedores de tecnología deberán entregar a las instituciones documentación que detalle las bases de datos de entrenamiento, los límites de precisión del sistema y los riesgos de sesgo conocidos, en los términos de la Tarjeta de Modelo establecida en el Capítulo V. Esta información deberá estar disponible para los comités correspondientes y para el personal clínico que operará el sistema. Todo sistema que no cumpla con este requisito no podrá ser habilitado en las unidades del SNS.
- **Equidad en la implementación:** La Secretaría de Salud deberá supervisar que la adopción de tecnologías basadas en IA no exacerbe la brecha digital en salud. Se dará prioridad a la implementación en unidades que atienden a poblaciones con limitaciones de acceso a servicios especializados, a fin de que la tecnología opere como mecanismo de nivelación del estándar de atención en todo el territorio nacional.
- **Capacitación de los comités:** La Secretaría de Salud podrá promover, en coordinación con instituciones académicas y organismos especializados, acciones de capacitación orientadas al fortalecimiento de capacidades para el análisis bioético, clínico y jurídico de tecnologías basadas en inteligencia artificial. El objetivo no es la formación técnica en ingeniería de software, sino la capacidad de interrogar la tecnología, identificar sus limitaciones y detectar riesgos para la dignidad del paciente y la autonomía del profesional de la salud.
- **Expediente técnico institucional:** Toda unidad de salud que implemente tecnologías basadas en IA deberá integrar y mantener actualizado un expediente técnico que incluya los resultados de la validación de interoperabilidad, el modelo de consentimiento informado utilizado, los registros de capacitación del personal y los reportes de desempeño del sistema. Este expediente estará disponible para auditoría por parte de la autoridad sanitaria competente.

Los presentes lineamientos establecen que la vía administrativa aquí descrita constituye el mecanismo institucional para garantizar que la incorporación de IA en el SNS sea un proceso ordenado, trazable y sujeto a rendición de cuentas, en el que la protección del paciente y la integridad del acto médico prevalezcan sobre cualquier consideración de eficiencia operativa o velocidad de adopción tecnológica.

Monitoreo continuo

Para garantizar la seguridad y eficacia de los sistemas de IA incorporados en el SNS las instituciones deberán implementar mecanismos de monitoreo continuo que permitan evaluar el desempeño a lo largo de todo el ciclo de vida, con el propósito de asegurar la protección del paciente y el funcionamiento adecuado en condiciones reales de uso.

A diferencia de otras tecnologías para la salud, como ya se mencionó en capítulos previos, los sistemas de IA son susceptibles a variaciones en su desempeño derivadas de cambios en los datos de entrada, en las características de la población o en los procesos clínicos, lo que puede generar degradación del rendimiento o desviaciones en sus resultados. Los marcos regulatorios internacionales han establecido la necesidad de adoptar un enfoque de vigilancia continua, adaptado a las particularidades de los sistemas basados en software, análogo en su estructura a la farmacovigilancia (6,31,32).

El monitoreo continuo se implementará mediante los siguientes componentes operativos:

1. Monitoreo continuo del desempeño

Las instituciones deberán evaluar el desempeño de los sistemas de inteligencia artificial de manera periódica o en tiempo real, mediante indicadores que incluyan, como mínimo: precisión, sensibilidad, especificidad, tasas de error, concordancia con la decisión clínica e impacto en resultados clínicos y operativos (29,30,32). Cada institución deberá establecer, previo al despliegue del sistema, los umbrales de desempeño aceptables para cada indicador. La caída sostenida por debajo de dichos umbrales constituirá una desviación relevante que deberá activar el protocolo de alerta y respuesta establecido en el numeral 3 del presente capítulo.

2. Detección y gestión de la deriva algorítmica

Los sistemas deberán contar con mecanismos específicos para identificar la degradación del desempeño del modelo a lo largo del tiempo. Las instituciones deberán establecer procedimientos para la evaluación periódica del modelo con una frecuencia mínima semestral, y de manera extraordinaria cuando se detecten variaciones significativas en las características de la población atendida, en los datos de entrada o en las prácticas clínicas institucionales. Dichos procedimientos deberán contemplar la recalibración o actualización del modelo cuando sea técnicamente necesario, así como criterios cuantitativos predefinidos para su suspensión en caso de pérdida significativa de desempeño (31,32).

3. Protocolo de alerta y respuesta

Toda institución deberá contar con un protocolo de alerta y respuesta que defina, de manera explícita, las acciones a seguir ante la detección de desviaciones en el

desempeño del sistema. Dicho protocolo deberá establecer al menos tres niveles de respuesta:

- **Nivel de monitoreo reforzado:** ante desviaciones menores, se incrementa la frecuencia de evaluación y se notifica al responsable técnico del sistema.
- **Nivel de intervención:** ante desviaciones moderadas o sostenidas, se restringe el uso del sistema a modalidades supervisadas, podrá solicitarse acompañamiento de instancias colegiadas competentes cuando se identifiquen posibles implicaciones bioéticas relevantes y se iniciará un proceso de recalibración o revisión técnica.
- **Nivel de suspensión:** ante desviaciones graves o pérdida significativa de desempeño que comprometan la seguridad del paciente, el responsable de la unidad de salud deberá suspender el sistema de manera inmediata, documentar la decisión en el expediente técnico institucional y notificar a la Secretaría de Salud y a COFEPRIS en los plazos establecidos. La reactivación del sistema requerirá validación técnica formal conforme al proceso establecido en el Capítulo VII.

4. Trazabilidad de decisiones y resultados

Las instituciones deberán garantizar el registro sistemático del uso de los sistemas de inteligencia artificial, incluyendo las recomendaciones emitidas, las decisiones clínicas tomadas por el profesional de la salud y los resultados obtenidos. Este registro deberá integrarse al expediente clínico electrónico y estar disponible para auditorías, evaluaciones de desempeño y procesos de rendición de cuentas institucional (5,12). La trazabilidad constituye un requisito de transparencia algorítmica indispensable para verificar que el principio de humano al mando se ejerce de manera efectiva en la práctica clínica.

5. Integración de resultados en la toma de decisiones institucional

Los resultados del monitoreo del desempeño deberán integrarse de manera sistemática en los procesos de toma de decisiones institucionales respecto a la continuidad, modificación, restricción o retiro de los sistemas de inteligencia artificial. Las instituciones deberán presentar a la Secretaría de Salud un reporte anual de desempeño de los sistemas de IA en operación, con los indicadores establecidos en el numeral 1 del presente capítulo, el registro de eventos adversos del periodo y las acciones correctivas implementadas. La Secretaría de Salud consolidará esta información a nivel nacional y la utilizará para actualizar el presente lineamiento, retroalimentar a COFEPRIS sobre la seguridad de las tecnologías autorizadas y orientar las prioridades de evaluación de tecnologías para la salud en el SNS.

El enfoque de ciclo de vida adoptado en el presente lineamiento establece que la vigilancia y el seguimiento no son procesos aislados ni opcionales, sino componentes

integrados e indispensables de la gobernanza de tecnologías basadas en inteligencia artificial. Su implementación constituye un componente esencial para la continuidad del permiso de uso institucional de cualquier sistema de IA en el SNS (6,31,32).

Referencias

1. Cámara de Diputados del H. Congreso de la Unión. Ley General de Salud: reformas integrales en materia de salud digital y la Base Nacional de Información en Salud. Diario Oficial de la Federación. 2026.
2. Ouanes K, Farhah N. Effectiveness of artificial intelligence (AI) in clinical decision support systems and care delivery. *J Med Syst*. 2024;48:74.
3. Brunyé TT, Mitroff SR, Elmore JG. Artificial intelligence and computer-aided diagnosis in diagnostic decisions: 5 questions for medical informatics and human-computer interface research. *J Am Med Inform Assoc*. 2026;33(2):543-550.
4. Academia Nacional de Medicina de México. Inteligencia artificial en medicina: desafíos éticos, legales y profesionales en el contexto mexicano. *Gac Med Mex*. 2023;159(3):245-258.
5. Van Baalen S, Boon M, Verhoef P. From clinical decision support to clinical reasoning support systems. *J Eval Clin Pract*. 2021;27(3):520-528.
6. Warraich HJ, Tazbaz T, Califf RM. FDA perspective on the regulation of artificial intelligence in health care and biomedicine. *JAMA*. 2025;333(3):241-247.
7. World Health Organization. Ethics and governance of artificial intelligence for health: guidance for regional and national implementation. Ginebra: WHO Press; 2021.
8. World Health Organization. Regulatory considerations on artificial intelligence for health. Ginebra: WHO Press; 2023.
9. Fraser AG, Biasin E, Bijmens B, et al. Artificial intelligence in medical device software and high-risk medical devices: a review of definitions, expert recommendations and regulatory initiatives. *Expert Rev Med Devices*. 2023;20(6):467-491.
10. Ayorinde A, Mensah DO, Walsh J, et al. Health care professionals' experience of using AI: systematic review with narrative synthesis. *J Med Internet Res*. 2024;26:e55766.
11. Vasey B, Ursprung S, Beddoe B, et al. Association of clinician diagnostic performance with machine learning-based decision support systems: a systematic review. *JAMA Netw Open*. 2021;4(3):e211276.
12. Cresswell K, Rigby M, Magrabi F, et al. The need to strengthen the evaluation of the impact of artificial intelligence-based decision support systems on healthcare provision. *Health Policy*. 2023;136:104889.
13. International Medical Device Regulators Forum. Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations. IMDRF/SaMD WG/N12. 2014.
14. International Medical Device Regulators Forum. Software as a Medical Device (SaMD): Application of Quality Management System. IMDRF/SaMD WG/N23. 2015.
15. Food and Drug Administration. Software as a Medical Device (SaMD): Clinical Evaluation. Guidance for Industry and FDA Staff. Silver Spring: FDA; 2017.
16. Secretaría de Salud. Norma Oficial Mexicana NOM-024-SSA3-2012: Sistemas de información de registro electrónico para la salud. Diario Oficial de la Federación. 2012.
17. Jain SS, Elias P, Poterucha T, et al. Artificial Intelligence in Cardiovascular Care-Part 2: Applications: JACC Review Topic of the Week. *J Am Coll Cardiol*. 2024;83(24):2487-2496.

18. Astărăstoae V, Rogozea LM, Leășu F, Ioan BG. Ethical Dilemmas of Using Artificial Intelligence in Medicine. *Am J Ther.* 2024;31(4):e388-e397.
19. Suprema Corte de Justicia de la Nación. Tesis aislada sobre la responsabilidad profesional médica en el uso de tecnologías de diagnóstico avanzado. *Semanario Judicial de la Federación.* 2024.
20. Zayas-Mugica R. El acto médico frente al reto de la inteligencia artificial: una perspectiva desde la salud pública mexicana. *Rev Salud Publica Mex.* 2025;67(1):45-60.
21. Universidad Nacional Autónoma de México. Bioética y derecho en la era de la salud digital: hacia una regulación humanista. México: Instituto de Investigaciones Jurídicas; 2024.
22. Comisión Nacional de Bioética. Código de Bioética para el personal de salud. México: Secretaría de Salud; 2026.
23. Johns Hopkins University. Precision Medicine and Artificial Intelligence: A Guide for Clinical Implementation. Baltimore: Johns Hopkins University Press; 2024.
24. National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies. Londres: NICE; 2022. Disponible en: <https://www.nice.org.uk/corporate/ecd7>
25. Reverberi C, Rigon T, Solari A, et al. Experimental evidence of effective human-AI collaboration in medical decision-making. *Sci Rep.* 2022;12:14952.
26. Health Level Seven International. HL7 FHIR Release 5: Standard for trial use. 2023. Disponible en: <https://www.hl7.org/fhir/>
27. ISO - International Organization for Standardization. ISO 13485:2016. Medical devices — Quality management systems — Requirements for regulatory purposes. Ginebra: ISO; 2016.
28. World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Ginebra: WHO Press; 2024.
29. Angus DC, Khera R, Lieu T, et al. AI, health, and health care today and tomorrow: the JAMA summit report on artificial intelligence. *JAMA.* 2025;334(18):1650-1664.
30. Odone A, Barbati C, Amadasi S, Schultz T, Resnik DB. Artificial intelligence and infectious diseases: an evidence-driven conceptual framework for research, public health, and clinical practice. *Lancet Infect Dis.* 2026;26(3):e152-e167.
31. Guan H, Bates D, Zhou L. Keeping Medical AI Healthy and Trustworthy: A Review of Detection and Correction Methods for System Degradation. *IEEE Trans Biomed Eng.* 2025;PP. Epub ahead of print.
32. Lin JC, Jain B, Iyer JM, et al. Benefit-Risk Reporting for FDA-Cleared Artificial Intelligence-Enabled Medical Devices. *JAMA Health Forum.* 2025;6(9):e253351.
33. Labkoff S, Oladimeji B, Kannry J, et al. Toward a responsible future: recommendations for AI-enabled clinical decision support. *J Am Med Inform Assoc.* 2024;31(11):2730-2739.
34. Tschochohei M, Adams LC, Bressemer KK, Lammert J. KI-gestützte klinische Entscheidungsunterstützungssysteme: Herausforderungen und Potenziale [AI-enabled clinical decision support systems: challenges and opportunities]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2025;68(8):872-879. [German].
35. Jackson NJ, Brown KE, Miller R, et al. Factors influencing the effectiveness of artificial intelligence-assisted decision-making in medicine: a scoping review. *J Am Med Inform Assoc.* 2026;ocag002. Epub ahead of print.:1-11.

