

Interacciones robot-paciente: Condicionantes actuales y prospectiva desde la ética del cuidado

Miguel Vieito Villar

Abogado.
Vocal del Comité de Ética Asistencial de la Xerencia de
Xestión Integrada de Santiago de Compostela
Doctorando – Universidad Pública de Navarra
miguelvieitovillar@yahoo.es

Robot-patient Interactions: Current Determinants and Prospective from the Ethics of Care

ISSN 1989-7022

RESUMEN: La innovación es un factor indisolublemente unido a una atención sanitaria de calidad. Sin embargo, y pese a la existencia de controles férreos, incluso a nivel ético, para la irrupción de nuevos medicamentos y productos sanitarios, no se verifica un estándar de exigencia en la inclusión de inteligencia artificial y robótica en salud, siendo que ya asistimos a ejemplos de interacción robot-paciente.

Es en este entorno de retos y nuevos paradigmas en el que proponemos un análisis de los textos internacionales y nacionales que abogan por dotar al uso de las inteligencias artificiales en contacto con personas usuarias de un mínimo común de seguridad y responsabilidad, protegiendo la integridad del paciente, también ante el daño moral.

Queremos también llamar la atención sobre la relevancia e idoneidad de una aproximación desde la ética del cuidado, como marco de referencia para el desarrollo de la práctica generalizada de interacciones máquina-humano en sanidad.

PALABRAS CLAVE: robótica, inteligencia artificial, responsabilidad, interacciones robot-paciente, ética del cuidado

ABSTRACT: Innovation in the field of health services is an indissolubly included factor in quality healthcare. However, and even since the existence of strong controls regarding new medicines and medical devices, there is not a required standard concerning artificial intelligence and its application in health, counting we are facing examples of robot-patient interactions

In this field of new challenges and paradigms, we propose an analysis of both international and national documents which defend giving this use of artificial intelligence in touch with users a minimum of security and responsibility, protecting the patients' integrity, concerning moral damage as well.

We would also like to highlight the relevance and adequacy of an approach coming from the ethics of care, as a framework for the development of this general utilisation of machine-human interactions in healthcare.

KEYWORDS: robotics, artificial intelligence, liability, robot-patient interactions, ethics of care

1. Introducción

Al ser humano no le es ajena la ilusión de una entidad creadora moldeando a su voluntad un ser finalista, útil a sus deseos y carente de libre albedrío. Desde el mito fantástico de Talos o del propio Prometeo, hasta el gólem de Praga o el mismo monstruo del Doctor Frankenstein, ha sido recurrente el imaginar escenarios en los que se domina la fuerza dadora de vida, dotando a elementos inanimados (bronce, barro, o restos inertes de otros seres), de nuevo, de movimiento, acciones y voluntad.

Tampoco fuera de los límites de lo mitológico nos extraña la creación de máquinas complejas. Así, y desde los autómatas de Herón de Alejandría o Apolonio de Perga, en el siglo III a.C., el conocido como "hombre de hierro" de San Alberto Magno (1240 – 1282), los autómatas de Gianello della Torre (1540) o el famoso escribano de Pierre Jacques-Droz (1774), el ser humano ha llevado a la realidad las ensoñaciones mitológicas, buscando emular un poder celestial.

Es, sin embargo, en el momento actual, con vastos conocimientos sobre la psique humana y con los desarrollos en procesamiento de lenguaje natural, aprendizaje profundo (*deep learning*) y redes neuronales artificiales, en el que estamos más cerca de conseguirlo. Se habla así del desarrollo de entidades personales no humanas e inteligencias artificiales que puedan desarrollar una singularidad al margen de sus creadores.



Nos encontramos pues ante virtualidades que pueden desarrollar razonamientos complejos más allá de los programados, que pueden aprender autónomamente, e incluso realidades electrónicas que desarrollan procesos tan humanos como mentir¹, engañar o crear sistemas de comunicación propios². Tanto es así que incluso nos hemos planteado ya el otorgar nacionalidad a entidades robóticas³, reconocerles la titularidad de un pasaporte⁴ y considerar necesaria una profunda revisión de los fundamentos de nuestro Derecho Civil, al hablar de “personalidad electrónica”, como una realidad diferente y extra muros de la personalidad clásica⁵.

Se nos derrumba poco a poco el antropocentrismo exclusivista: ya no solo regulamos (o regularemos) realidades centradas y enfocadas en el ser humano como única entidad con derechos y obligaciones; sino que debemos crear nuevas instituciones, o al menos adaptar las existentes, para acomodar a una entidad capaz no personal, que puede desarrollar un pensamiento superior. En consecuencia, las interacciones que mantenemos en nuestro día a día no serán ya únicamente humano-humano (bien sea con relaciones cara a cara o con mediación de tecnología) sino genuinamente máquina-humano, con cada vez más escaso factor personal en uno de los polos.

2. Definiciones y planteamiento

2.1. *Hacia un concepto de Inteligencia Artificial*

En primer término, debemos poner de manifiesto que el término “robot” es, sin duda, el más inespecífico, no designando una realidad concreta. El primer uso de este vocablo data del año 1921, por parte del autor checo Karel Čapek en su obra R.U.R. (*Rossumovi univerzální roboti*). “Robot”, que podríamos traducir como “esclavo” o “trabajador forzado”, es recogido décadas después por el cine y la literatura, siendo la voz más ampliamente diseminada, aceptada y popularizada.

En segundo lugar, cabe mencionar que no es “robot” el término más antiguo, recayendo este mérito en el vocablo “androide”, recogido ya por San Alberto en 1270, y popularizado por el autor francés Auguste Villiers en su novela *L'Ève future* (1886). Utilizando “androide” aludimos genéricamente a una entidad mecánica de aspecto antropomorfo.

En tercer lugar, en un uso ligeramente más preciso, encontramos el término “sistemas ciber-físicos”, definido por la Unión Europea⁶ como: sistemas robóticos inteligentes, vinculados con el Internet de las cosas, o sistemas técnicos de ordenadores en red, robots e inteligencia artificial, que interactúan con el mundo físico. Nótese que esta definición aporta un paso más al indicar que no se trata de máquinas autónomas aisladas, sino aquellas que entran en interacción directa con el mundo que las rodea, sea a través de sensores (interacción con factores ambientales) o con personas (interacción con factores humanos).

Por último, la fórmula “inteligencia artificial” delimita una realidad capaz de analizar datos y actuar en consecuencia en la consecución de un fin, programado de antemano. Esto es, sumamos a los rasgos ya vistos (movilidad, sometimiento a órdenes e interacción con el mundo físico, o captación de información del medio), la capacidad de análisis, elaboración de razo-

namientos y actuación conforme a los mismos. En la definición más moderna que la Unión Europea⁷ propone para el término, se define del siguiente modo:

“Los sistemas de inteligencia artificial (IA) son sistemas de software (aunque también pueden incluir hardware) diseñados por humanos, dado un objetivo complejo, para actuar en una dimensión digital o física, mediante la percepción de su entorno y a través de la adquisición de información, interpretación de los datos estructurados o no estructurados, el razonamiento con base en ese conocimiento, o el procesamiento de información, como consecuencia de lo cual son capaces de decidir la mejor acción para conseguir el objetivo dado. Los sistemas de IA pueden usar reglas simbólicas o aprender un modelo numérico, y pueden adaptar su comportamiento analizando como el entorno se ve afectado por sus acciones anteriores.

Como disciplina científica, la IA incluye aproximaciones y técnicas, como el *machine learning* (del cual el aprendizaje profundo y el aprendizaje reforzado son ejemplos específicos), el razonamiento mecánico (que incluye planear, programar, representar conocimiento y razonamientos, así como búsqueda y optimización), y robótica (que incluye control, percepción, uso de sensores y mecanismos de acción, así como la integración de todas las otras técnicas en sistemas ciber-físicos)”

El punto de partida de la inteligencia artificial nos permite construir progresiones, también terminológicas, de la misma. Así, es comúnmente aceptado el uso de conceptos como IA general o fuerte (*strong AI*) e IA débil (*narrow AI*), para referirnos a aquella que puede llevar a cabo acciones tradicionalmente asignadas a seres humanos, en el primer caso, y a la que está diseñada solo para realizar un rango muy limitado de acciones, en el segundo.

2.2. El uso de la robótica en las relaciones médico-paciente

El pasado marzo acontecía en el *Kaiser Permanente Medical Center* de Fremont (California) uno de los más monumentales fiascos del uso clínico de la tecnología en salud; y, en concreto, de las interacciones tecnología-paciente.

Tal y como relata la familia de Ernest Quintana, de 78 años e ingresado de urgencia tras el agravamiento de su EPOC (enfermedad pulmonar obstructiva crónica), en la tarde del 3 de marzo un robot entró en su habitación acompañado de una enfermera, y sin mediar aviso de ningún tipo, uno de los profesionales responsables de la asistencia ponía en conocimiento de la familia del paciente, que no del Señor Quintana, pues, debido a su avanzada edad había perdido agudeza auditiva, que sus pulmones no respondían al tratamiento, y que la muerte se produciría de modo inminente. Sin mediar intervención de ningún tipo por parte de la enfermera que acompañaba a la máquina y sin que nada se hubiera explicado al propio Señor Quintana, la máquina abandonó la sala⁸.

Lo acontecido en Fremont se suma a una ya larga lista de aplicaciones directas de robótica en pacientes. Señalando únicamente los supuestos con mayor difusión, destacamos:

- Paro⁹ es un robot-peluche con forma de foca, creado por el ingeniero japonés Takanori Shibata en el año 2002. Anunciado como el primer robot terapéutico, su uso en el acompañamiento de pacientes con alzheimer o demencia es amplio en países como Japón, Reino Unido o Uruguay.

- Doctor Asistente AI¹⁰: desarrollado por la corporación china iFlytek, es un robot que realiza funciones similares a las de un médico de Atención Primaria, pasando consulta y recetando medicamentos a pacientes
- RIBA II (*Robot for Interactive Body Assistant II*) es un androide creado por la compañía japonesa RIKEN¹¹, que se especializa en la movilidad de pacientes, transportándolos de la cama a la silla de ruedas y viceversa.

Es nuestra intención partir de estos y otros supuestos para proponer una reflexión sobre el uso de la inteligencia artificial en las relaciones sanitarias con los pacientes. Tomaremos así en nuestro análisis subsiguiente un concepto amplio de inteligencia artificial, sea general o débil, considerando toda solución tecnológica inteligente que mantenga relaciones con factores humanos en la forma de pacientes o usuarios de servicios de salud. En este ámbito, y a falta de referencias normativas, podemos categorizar a estas inteligencias artificiales en los términos de la norma ISO 13482:201412, considerando la existencia de: robots de servicio móvil, asistentes físicos y robots transportadores de personas¹³. Nos encontramos así ante los denominados robots de atención y/o cuidado¹⁴, carebots o robots asistenciales, nacidos como mecanismos de ayuda o soporte (RIBA II nace como apoyo del personal de enfermería) pero que han evolucionado hacia un uso cada vez más autónomo (Doctor Asistente IA), ante determinantes como la falta de personal derivada de los recortes presupuestarios en salud y el elevado índice de personas en situación de dependencia.

Se trata de realidades múltiples y variadas que tienen una interacción directa con pacientes, pero que no se sujetan, empero, a las normas propias de la profesión médica, ni a las del personal de enfermería, auxiliar o celadores. No son sujetos vinculados a las normas sobre autonomía del paciente, consentimiento informado o estatuto de la profesión sanitaria. Se encuentran sometidos a la regulación de productos y dispositivos, pese a que extralimitan esas fronteras para abarcar funciones propias, hasta ahora al menos, de seres humanos.

3. Límites al uso de inteligencias artificiales en la relación con los pacientes

La observación de los retos éticos que el uso de inteligencias artificiales en la interacción con pacientes implica no ha pasado desapercibido a múltiples entidades y organizaciones que, desde hace años, vienen analizando estas interacciones y estableciendo propuestas de marcos de deliberación. En este sentido, veremos a continuación las aportaciones internacionales y nacionales más relevantes en la materia, incluyendo las más recientes.

3.1. La posición de grupos de expertos

En el año 2017 tiene lugar en el *Asilomar Conference Grounds* (California) la conferencia auspiciada por la organización *Future of Life Institute*, y a la que acuden líderes mundiales en economía, derecho, ética y filosofía, bajo el rótulo de “inteligencia artificial y beneficencia”. Las conclusiones de este encuentro, plasmadas en la formulación básica de 23 principios de ética robótica¹⁵, son hoy uno de los más relevantes estándares para medir o valorar la corrección ética de una aplicación tecnológica en salud. De entre todos ellos, y para el propósito que nos ocupa, destacaremos los siguientes:

- **Seguridad:** indicándose que todos los sistemas de IA deberán proporcionar un entorno garantista durante el transcurso de su vida útil.
- **Alineamiento en valores:** los sistemas basados en inteligencia artificial con un alto grado de autonomía deben diseñarse de tal modo que sus valores y objetivos se alineen con valores humanos.
- **Valores humanos:** todo sistema de IA debe operar con pleno respeto y compatibilidad con las ideas humanas de dignidad, derechos, libertades y diversidad cultural.
- **No subversión:** el poder conferido a sistemas avanzados de IA debe respetar e incrementar, y no subvertir, los procesos sociales y cívicos de los que depende la salud de nuestra sociedad.
- **Bien común:** la superinteligencia debiera únicamente ser desarrollada bajo la premisa del respeto a valores éticos compartidos y para el beneficio de la humanidad, y no de un Estado o una corporación individuales.

El acercamiento es cauteloso, con un fuerte fundamento de humanización del uso de la inteligencia artificial, apostando por un modelo basado en una idea de riesgo/precaución, de tal modo que únicamente aquella tecnología que pruebe su efectividad y, al tiempo, garantice un respeto de los valores de los sujetos sobre los que se aplica, puede ser aceptable.

En segundo término, podemos citar en este apartado las recomendaciones establecidas por el panel de expertos de la Universidad de Stanford, en su texto *Artificial Intelligence and life in 2030*¹⁶. Este primer documento de conclusiones, dentro del marco de desarrollo del estudio *One Hundred Year Study on AI*, bebe directamente de los avances del marco teórico de Asilomar, y plantea una labor centrada en campos como la prestación de servicios sanitarios y robots de asistencia desde un punto de vista multidisciplinar e interconectado.

A diferencia de Asilomar, la Universidad de Stanford se posiciona en un punto escéptico con respecto al desarrollo de inteligencias artificiales generales, con consciencia de sí mismas y pensamiento complejo. De este modo, considera que estas no suponen hoy una amenaza para la humanidad, y no lo serán en el futuro próximo; es decir, no parece suponer un riesgo que una máquina elija dañar a seres humanos, sino más bien que un agente humano cause daño a sus pares con un mal uso de la tecnología. No obstante, el informe coloca en el punto de mira, en lo que a sanidad se refiere, la necesidad de que la robótica “se gane la confianza” de los pacientes, así como en lo importante de las relaciones interactivas naturales tanto con pacientes como con familiares de estos.

Stanford coloca un elemento interesante de análisis, que eleva los requisitos de Asilomar: no solo es necesario un diseño garantista, *ad initium*, que asegure un respeto de la dignidad humana y de los valores que nos son propios; sino que la relación no puede catalogarse como satisfactoria por el mero hecho de existir, debiendo ser además fluida, directa y natural, del mismo modo que lo es (que debiera serlo) la interacción humano-humano. De este modo, devienen cruciales dos elementos recurrentes en el estudio y que se colocan como una de las fronteras de futuro de la IA: el procesamiento de lenguaje natural y el reconocimiento discursivo, en aras a favorecer una comunicación adecuada entre inteligencia artificial y pacientes.

Se señala como factor relevante la individualización, así como la atención a las características de las personas con quien deben interactuar los sistemas ciber-físicos, esto es:

- Las características generales de todo paciente, y las particularidades de la relación médico-paciente, con atención a cualidades como la confianza, la empatía, la privacidad o el secreto médico.
- Las peculiaridades de los pacientes concretos, de cada una de las especialidades en las que la inteligencia artificial vaya a desarrollar su labor. De este modo, no sobrepasaría el estándar mínimo un robot configurado de igual modo para interactuar con pacientes mayores de edad, que para pacientes pediátricos; o con las mismas características para pacientes cardíacos, que para pacientes de servicios de salud mental o sintomatología neurológica.

En definitiva, Standford nos apunta a una conclusión: la aplicación de la inteligencia artificial en la sanidad, y en los cuidados específicamente, no puede tomar como punto básico la uniformidad y el diseño único.

3.2. La interacción robot-paciente en el nuevo Código de Deontología médica español

Si bien el texto no se encuentra aún en su versión final¹⁷, debemos resaltar que esta actualización muestra una evidente preocupación por el uso de la medicina electrónica en términos generales, al tiempo que sobre el uso de inteligencias artificiales. No en vano les dedica un total de dos capítulos, el XXV y el XXVI, respectivamente.

Afirmando (artículo 104) que a cualquier forma de comunicación electrónica con los pacientes le es plenamente aplicable el propio Código, el artículo 102.2 nos coloca ante la primera de las notas de relevancia para el establecimiento de un marco ético del uso de las nuevas tecnologías basadas en inteligencia artificial o robótica: deben haber demostrado su validez en beneficio de los pacientes; siendo que tanto los profesionales médicos como los de administración deben mantener sus conocimientos actualizados al respecto. El artículo 104.4 es tajante y directo:

- “Son contrarias a la Deontología Médica las intervenciones carentes de fundamento científico o que consistan en charlatanismo”

En esta línea, observamos con claridad que la Organización Médica Colegial exige un estándar alto de cuidado, basado en la evidencia y en la comunicación garantista con los derechos de los pacientes, asegurando un marco de respeto de valores y fundamentos éticos. En esta línea, conforma el artículo 112 del nuevo Código quizá el ejemplo más claro de tal posicionamiento, refiriendo su apartado primero:

- “El médico debe exigir un control ético y finalista de la investigación con inteligencia artificial, basado en la transparencia, la reversibilidad y la trazabilidad de procesos en los que intervenga, para garantizar la seguridad del paciente”

Parece posicionarse así la entidad colegial en la exigencia de un control ético previo de los algoritmos usados en sanidad, en tanto que exige plena transparencia. Podemos también interpretar que las inteligencias artificiales deben contar con una identificación propia como

tales, comunicando precisamente al propio paciente que no son entidades humanas y que la interacción se produce en un entorno electrónico. Por último, la organización se posiciona en la exigencia de un control de los procesos lógicos de la inteligencia en sanidad; esto es, no podrán basarse en procesos de *deep learning* oscuros y en los que sea imposible descifrar la cadena de decisiones que conduce a las consecuencias y elecciones; así mismo, deberá ser exigible la presencia de “cajas negras”, donde se registren las citadas cadenas de procesamiento.

Pero el nuevo código da incluso un paso más. No solo establece los determinantes previos para el uso de inteligencias artificiales, ya de por sí rigurosos, sino que identifica una línea roja de gran importancia: el uso de inteligencias artificiales o sistemas robóticos solamente podrá tener una finalidad colaborativa para con el profesional y nunca podrá dejarse en manos de esta entidad una decisión en el proceso asistencial del paciente¹⁸.

3.3. El posicionamiento de la Organización Mundial de la Salud sobre intervenciones digitales

El pasado 17 de abril de 2019, la Organización Mundial de la Salud afronta el reto que suponen las actividades electrónicas en un entorno de salud, emitiendo la primera guía internacional, de ámbito ONU, en la materia.

Ya desde su introducción, el texto aboga por una aproximación crítica a la tecnología, que no suponga su uso gratuito y caprichoso; en este sentido, se refuerza una vez más la exigencia de evidencia y balance riesgo/beneficio como requisitos *ex ante*¹⁹. Al tiempo, refuerza la idea de que el uso de la tecnología debe redundar en una atención sanitaria inclusiva que “no deje a nadie atrás”, lo que nos habla de un factor eminentemente democratizador y de una aplicación universal. Las tecnologías digitales, afirma el documento, no son (no deben ser) fines en sí mismas; sino herramientas fundamentales para la promoción de la salud, mantener un estándar de seguridad en nuestra sociedad y amparar a los colectivos vulnerables.

De entre los aspectos analizados por la OMS en este informe, nos resulta especialmente oportuno traer a colación los ítems exigidos para una aproximación “de la evidencia a la decisión de implementación”:

- **Efectividad:** debiendo responder la tecnología a preguntas como: ¿cuáles serán los efectos deseados y/o incidentales e indeseados de esta intervención? o ¿cuál es el nivel de evidencia de esta tecnología?
- **Aceptabilidad:** la intervención digital en salud deberá justificarse en función de la demanda, respondiendo a si se trata de una opción aceptada por los pacientes.
- **Viabilidad:** dadas evidencia y aceptación, deben preguntarse los gestores en salud si existen limitaciones materiales, de infraestructuras o de entrenamiento de su propio personal, antes de incorporar una solución tecnológica.
- **Utilización de recursos:** la OMS nos coloca así frente a un criterio de coste/efectividad, que entronca directamente con un análisis en clave de justicia distributiva.
- **Género, igualdad y derechos:** ¿disminuirá esta medida las inequidades del sistema de salud? ¿Qué impacto tendrá esta medida en las desigualdades de los servicios?

Estos serán los criterios sobre la base de los cuales la OMS analizará la evidencia científica existente para cada una de las aplicaciones electrónicas de salud consideradas²⁰. Sin que sea nuestra intención analizar todas las conclusiones del documento, sí creemos valioso traer al foro aquella relacionada con la comunicación entre pacientes y entornos tecnológicos.

Así, para el uso de la telemedicina en las relaciones cliente–proveedor, la OMS indica que el método es potencialmente efectivo. En cuanto a su aceptabilidad, se señala que los trabajadores de los sistemas de salud rechazan la medida, por cuanto consideran que una relación médico-paciente cara a cara redundaría en una mayor calidad de la atención. Por último, y pese a que la medida puede reducir desigualdades en el sistema, no existe evidencia de coste-efectividad en el uso de recursos.

En suma, la OMS recomienda la implantación de mecanismos telemáticos de relación médico-paciente únicamente verificando:

1. Que la medida complementa, y no reemplaza, la relación cara a cara.
2. Que la intervención garantiza la seguridad, confidencialidad de los datos, trazabilidad, responsabilidad por daños y seguridad de los pacientes.
3. Que los pacientes pueden ser monitorizados en un entorno seguro.
4. Que los procedimientos garantizan el consentimiento informado, la protección de datos y el almacenamiento seguro de los mismos.

Tras la exhaustiva consideración de la evidencia científica existente en el *topic* “relación médico–paciente en base electrónica”, se plantea un acercamiento crítico, cauteloso y eminentemente garantista que en ningún caso reemplace a la relación clínica cara a cara tradicional, y que busque decididamente un diseño para todos, con seguridad, efectividad y trazabilidad. Introduce también, de un modo similar al Código de Deontología Médico, la consideración al consentimiento informado, basándose este en un diálogo transparente, veraz, honesto y completo con el usuario. De este modo, y en las relaciones entre robot y sujeto, debe manifestarse abiertamente que es una inteligencia artificial quien dialoga, los términos en los que esta puede hacerlo y cuáles son los potenciales riesgos de esta relación dialogal.

3.4. Las normas éticas europeas en materia de uso de inteligencias artificiales

La creación de un grupo de trabajo sobre robótica en el Parlamento Europeo en el año 2015 ofrecía ya una muestra de la preocupación de la Unión sobre los riesgos transfronterizos de la inteligencia artificial, tomando a la sanidad como uno de los entornos en el que el análisis debía ser mayor. Así, en febrero de 2016 se publica la ya aludida Resolución del Parlamento sobre normas de derecho civil, en las que la salud ocupaba un lugar preeminente, y los “robots médicos” un apartado propio.

Estas normas contenían como uno de sus Anexos un apartado de máxima relevancia, al crearse un Código Ético de conducta para los profesionales de ingeniería a cargo de soluciones de inteligencia artificial. En este se exigía no solo formación en los principios clásicos de la bioética (autonomía, beneficencia, no maleficencia y justicia), sino incluso una aproximación presidida por el principio de precaución, que exigía a los investigadores anticiparse a los eventuales daños que su tecnología pudiese causar. Por último, se incluía también un llamamiento

a los Comités Éticos de la Investigación, como encargados del examen de estos ítems, relacionados con las aplicaciones en salud de la robótica.

Tiempo después, ya en 2016, la Oficina de Prospectiva Científica de la Unión Europea señalaba el reto ético de las inteligencias artificiales en Sanidad, ante preguntas como: ¿Cuándo debe un robot tomar decisiones en lugar de un paciente? ¿Debe existir una actitud paternalista del robot hacia el paciente en las decisiones que afecten negativamente a su salud? ¿Cuándo un sistema ciber-físico debe requerir ayuda de un factor humano? Al tiempo que llamaba la atención, como uno de los factores clave a replantear en el medio plazo, sobre el paradigma de relaciones médico-paciente en el uso de la robótica.

Ya por último, aparece en abril de este mismo año el que es hasta ahora el texto más completo y ambicioso de la Unión en materia de retos éticos del uso de inteligencia artificial. Rotulado *Ethics Guidelines for Trustworthy AI* (Comisión Europea, 2019b), es el resultado final del grupo de expertos de alto nivel de la Comisión Europea y plantea de modo claro que el modelo europeo de inteligencia artificial debe aposentarse en tres pilares fundamentales: uso ético, seguridad e innovación hecha en la Unión. Se posiciona el Viejo Continente, pues, en un valor añadido de control y parámetros de seguridad para los pacientes y a diferencia de una perspectiva con menor grado de exigencia, propia de países como Estados Unidos, Japón o Corea.

El punto de partida del Informe es la confianza en la inteligencia artificial. Esta debe basarse, a su vez, en tres ejes motores, que garanticen un uso sostenible, y minimicen los efectos negativos. En primer término, el uso de la robótica debe ser legal, cumpliendo con las normas y regulaciones existentes. En segundo lugar, debe ser robusto, siendo que debe apoyarse en la evidencia científica y social, indicando el texto que “la IA puede causar daños intencionados, incluso cuando su uso nace de la buena fe”. Por último, debe tratarse de un uso ético, entendiendo por este el que se produce con respeto a principios y valores, en nuestro caso, de los pacientes como sujetos morales con los que operar. Los principios expresados en el texto suponen un avance en la línea de textos que ya hemos observado. Así, se señalan como protagonistas el respeto por la autonomía humana, la prevención del daño, la justicia y la explicabilidad, entendida esta última como transparencia.

Estos principios deben ser analizados en una suerte de evaluación de calidad continua (Vid. Fig. 1), con ponderación entre principios, valores, finalidades y teniendo siempre presentes la seguridad y el análisis de la corrección técnica con base en nuevo conocimiento. Esta evaluación se realizará mediante formularios tipo, con cuestiones detalladas en cada uno de los ítems de evaluación²¹.

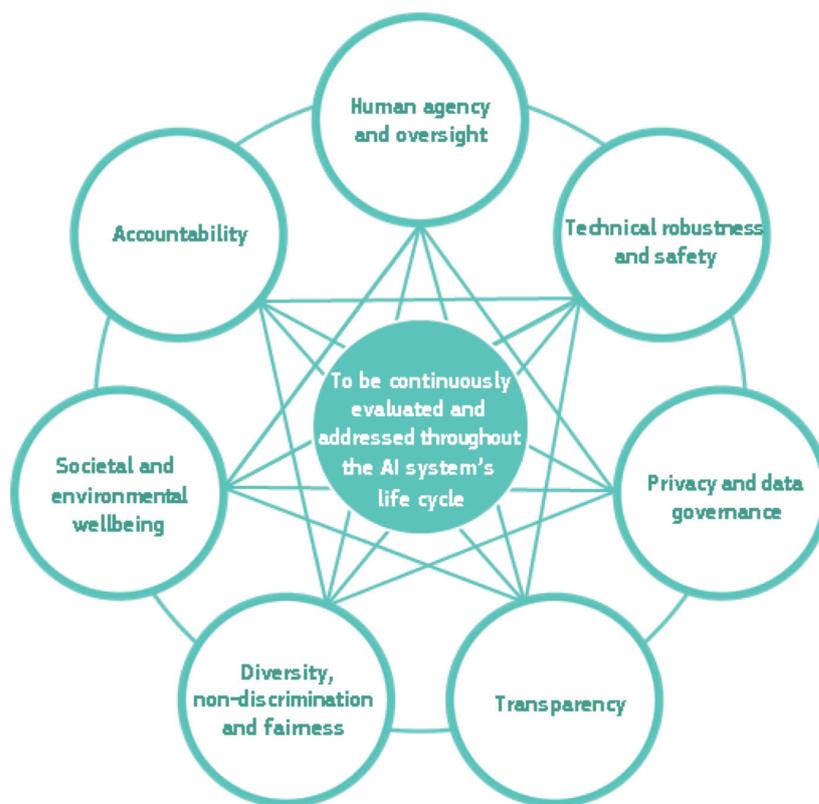


Fig. 1. Sistema de evaluación continua e interrelación de principios

4. Las relaciones robot-paciente desde la ética del cuidado

De todos los textos que hemos podido analizar destaca un posicionamiento global claro y directo: el uso de inteligencias artificiales, sean estas únicamente aplicaciones de software o soluciones de hardware y software, debe sujetarse a condicionantes previos a su puesta en contacto con pacientes, así como debe satisfacer un estándar mínimo en lo que respecta a la relación en sí misma. Un acercamiento ético *by design*, acompañado de una evaluación continua de satisfacción de estándares éticos, siempre que se verifique un entorno seguro, trazable, aceptado por los usuarios y respetuoso con las normas de la profesión médica.

Cabe pues plantear que el uso de robots o inteligencias artificiales en el cuidado de pacientes, y en cualquier otro escenario de interacción con ellos no debe importar acriticamente cualquier nueva tecnología existente: no debemos hacer todo aquello que, técnicamente, podemos hacer²².

Siendo esto así, el primer acercamiento a estas tecnologías sanitarias, desde un enfoque ético, debe pasar por desterrar completamente el planteamiento de que la medicina es una ciencia exacta, tendente únicamente a curar al paciente (Aguilar, 2015). La interacción con personas es un escenario eminentemente cambiante y generador de nuevas situaciones, tanto desde un punto de vista físico (pacientes con convulsiones, molestias, dolor que varía en intensidad, aparición de nuevos síntomas o manifestación súbita de cuadros clínicos urgentes

o críticos, entre otros factores), pero también emocional, debiendo respetarse el paradigma de consentimiento informado y una relación paciente-servicios de salud basada en la codecisión y en un modelo autonomista.

De este modo, la ética del cuidado es, en nuestro parecer, el modelo que satisface con mayor grado de precisión los estándares que vienen exigiéndose en la literatura internacional y nacional al respecto. Esta visión nos ofrece un juicio contextual, aproximándonos a la posición de aquella persona sobre la que, en este caso, la inteligencia artificial, se va a aplicar; y colocando en valor la responsabilidad en las relaciones médicas (Marín, 1993).

Tomando como referencia el texto del *Hastings Center (Bioethics Research Institute)*²³ americano, cuatro serían las finalidades de la medicina:

1. Prevención de enfermedades y lesiones, y promoción y mantenimiento de la salud.
2. Alivio del dolor y el sufrimiento causado por patologías.
3. Atención y curación de personas usuarias, y el confort de aquellas personas que no pueden ser curadas.
4. Evitación de la muerte prematura y garantía de una muerte tranquila.

Siendo que debemos colocar a los robots en interacción con los pacientes en alguna de estas categorías, y que estos deben cumplir con un estándar de respeto de derechos, no discriminación y consentimiento informado, resulta ineludible que deban incorporar, de un modo u otro, la empatía (Anderson, 2007) como factor operativo e integrante del concepto mismo de "inteligencia", desterrando así la operación lógica/matemática de búsqueda de soluciones eficientes bajo criterios únicamente objetivos. La ética del cuidado, como parámetro deliberativo y valorativo flexible en el que el sentir obtiene prioridad sobre el razonar (García Moyano, 2015), nos exige tomar en consideración los valores del paciente dentro de los algoritmos, así como un control ético previo de estos, en aras a garantizar la ausencia de sesgos que puedan derivar en daños morales al paciente.

Así, y con especial atención al cuidado de pacientes, el texto apunta una reflexión que consideramos plenamente aplicable al uso de sistemas robóticos como los comentados en la sección introductoria de este trabajo:

"En su entusiasmo por la cura de enfermos, en ocasiones, la medicina moderna ha desatendido su función humanitaria. Parecería afirmar que, si se puede encontrar la curación, ¿quién necesita cuidados? Esta opinión no podría estar más equivocada. En muchos casos, por supuesto, una técnica completamente impersonal resulta aceptable, e incluso deseable, como en traqueotomías de urgencia, reanimaciones cardiopulmonares y muchos tipos de cirugía de alta tecnología; pero lo más común es que se requieran también cuidados. Los cuidados no consisten simplemente en manifestar preocupación, conmiseración y disposición a hablar con el paciente; comprenden también la capacidad para hablar y escuchar de un modo que demuestra conocimiento de los servicios sociales y asistenciales necesarios para ayudar al paciente y a su familia a afrontar una amplia gama de problemas no médicos que pudieran acompañar, como normalmente ocurre, a su padecimiento"

La consideración a la ética del cuidado deberá nacer, en todo caso, de su inclusión ya en las etapas de diseño, a través de una investigación respetuosa y responsable hecha por la sociedad y para ella (en términos de Carsten y Coeckelbergh (2016)), incluyendo factores como el acceso abierto, la igualdad de género, el respeto ético y la gobernanza. Mas, en lo que se refiere a su puesta en marcha y mantenimiento, y al examen continuo en un marco presidido por el principio de precaución, debemos exigir a la robótica asistencial que asuma las características predicables de un “bueno cuidado” o de un cuidado respetuoso y de calidad.

Señala García Moyano (2015), citando a Carol Taylor que son seis las características de un buen cuidador:

“el afecto (experiencia de un sentimiento positivo hacia el paciente), la cognición (capacidad intelectual y técnica de identificar lo esencial para el bienestar del paciente), la volición (compromiso interno para movilizar los recursos personales y técnicos que aseguren el bienestar del paciente), la imaginación (emplear creativamente los recursos personales para relacionarse empáticamente con el paciente y entender su situación existencial particular, así como sus necesidades vitales), la motivación (flujo de energía aplicado al acto de cuidar, refleja el altruismo en el ejercicio de la profesión y el deseo de actuar en consecuencia), y finalmente la expresión (demostrar genuinamente la intención de ayuda mediante el uso de la estrategia de comunicación verbal y no verbal”

Creemos que, en rigor, pocos o ninguno de estos elementos pueden ser predicables del uso de inteligencias artificiales en la relación médico-paciente; al menos en lo referente a las acciones de cuidado y bienestar de las personas usuarias. La falta de emotividad, capacidad de generación de sentimientos y empatía, la fijación imperfecta de cuáles han de ser los ítems para valorar el bienestar del paciente, así como la ausencia de expresiones no verbales o motivación cuestionan, en nuestro parecer, la idoneidad de sustituir, e incluso complementar, la relación médico-paciente con el uso de entidades no humanas.

Como corolario, creemos que las exigencias a la inteligencia artificial pueden subsumirse en la exigencia de dos polos principales, que derivan directamente de un análisis en términos de una ética del cuidado aplicada a la robótica: aceptabilidad y adaptabilidad²⁴. En lo que respecta al primero, el uso de una entidad robótica en las relaciones médico-paciente debe ser técnicamente aceptable, en el sentido de proveer suficiente evidencia científica, garantizar un correcto equilibrio de coste-efectividad y un balance positivo de riesgo-beneficio (Rigby, 2019)²⁵. Al tiempo, debe garantizarse la aceptabilidad por parte de los usuarios, enmarcando esta en el consentimiento informado como expresión final de un proceso informativo transparente, veraz y honesto sobre propio uso de la tecnología. La adaptabilidad nos transporta hacia un examen, en primer lugar, de consideración de las circunstancias del medio; sean las propias del paciente como sujeto moral y, por tanto, planteando el respeto máximo a su dignidad, sean las del mundo físico/matemático en la que el robot ha de operar. En segundo término, exige de estos robots la capacidad de responder a las necesidades súbitas de las personas usuarias de un modo garantista, rápido, ágil y embebido de empatía y confianza.

Bibliografía

- Agencias Aranda de Duero, (2018). "Un robot pasa consulta y extiende recetas en un ambulatorio de China". La Voz de Galicia https://www.lavozdeg Galicia.es/noticia/sociedad/2018/03/06/robot-pasa-consulta-extiende-recetas-ambulatorio-china/0003_201803G6P26993.htm (última consulta 30.04.2019).
- Aguilar Fleitas, B. (2015). "Los Fines de la Medicina". *Revista uruguaya de cardiología*. 30(1), pp. 8–12.
- Anderson, E. E. (2007). "What We Talk about When We Talk about Goals". *Virtual Mentor - AMA Journal of Ethics*. 9(6), pp. 407 – 409.
- Cables New Network (CNN) (2019). "A doctor in California appeared via video link to tell a patient he was going to die. The man's family is upset". <https://edition.cnn.com/2019/03/10/health/patient-dies-robot-doctor/index.html> (última consulta 29.04.2019).
- Carsten Stahl, B y Coeckelbergh, M. (2016). "Ethics of healthcare robotics: Towards responsible research and innovation". *Robotics and Autonomous Systems*. Vol. 86, pp. 152–161.
- Comisión Europea - High-Level Expert Group on Artificial Intelligence (2019a). *A definition of AI: Main Capabilities and Disciplines*.
- Comisión Europea - High-Level Expert Group on Artificial Intelligence (2019b). *Ethics Guidelines for Trustworthy AI*. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (última consulta 30.04.2019).
- Domínguez-Alcón, C. (2017). "Ética del cuidado y robots". *Cultura de los Cuidados*. 21(47), pp. 9-13.
- Fundació Víctor Grífols i Lucas (2007). "Los Fines de la Medicina. Els Fins de la Medicina". http://www.hospital-santjoan.cat/wp-content/uploads/2018/09/fins_medicina.pdf (última consulta 30.04.2019).
- Future of Life Institute (2017). "Asilomar AI Principles". <https://futureoflife.org/ai-principles/?cn-reloaded=1> (última consulta 29.04.2019).
- García Moyano, L. (2015). "La ética del cuidado y su aplicación en la profesión enfermera". *Acta bioethica*. 21(2), pp. 311–317.
- García Portero, R. (2018). "Los robots en la sanidad", en Barrio Andrés, M. (Dir.). *Derecho de los robots*. Madrid, Wolters Kluwer, pp. 205–228.
- Gilligan, C. (2013). "El daño moral y la ética del cuidado", en Gilligan, C. *La ética del cuidado*. Cuadernos de la Fundació Víctor Grífols i Lucas. Núm. 30, pp. 10 – 40.
- Griffin, A. (2017). "Saudi Arabia grants citizenship to a robot for the first time ever". The Independent. <https://www.independent.co.uk/life-style/gadgets-and-tech/news/saudi-arabia-robot-sophia-citizenship-android-riyadh-citizen-passport-future-a8021601.html> (última consulta 29.04.2019).
- International Organization of Standardization (2014). *Robots and robotic devices -- Safety requirements for personal care robots*.
- Lin, P., Abney, K y Bekey, G. A. (ed.) (2012). *Robot ethics. The ethical and social implications of robotics*. Massachusetts Institute of Technology Press.
- Malta AI Taskforce (2018). "Malta.AI – Towards a national AI Strategy". <https://malta.ai/wp-content/uploads/2018/10/Malta-AI-Vision.pdf> (última consulta 30.04.2019).
- Marín, G. (1993). "Ética de la justicia y ética del cuidado". Asamblea de mujeres de Elx. http://www.feministas.org/IMG/pdf/etica_de_la_justicia_y_etica_del_cuidado_-_gloria_marin.pdf (última consulta 30.04.2019).
- Mackenzie, A. (2010). "Japan develops robotic seals to comfort sick and elderly". British Broadcasting Corporation. <https://www.bbc.com/news/av/health-11459745/japan-develops-robotic-seals-to-comfort-sick-and-elderly> (última consulta 30.04.2019).
- Mitri, S., Floreano, D. y Keller, L. (2009). "The evolution of information suppression in communicating robots with conflicting interests". *Proceedings of the National Academy of Sciences*. 106 (37), pp. 15786–15790.

- Oficina de Prospectiva Científica (STOA) del Parlamento Europeo (2016). *Ethical Aspects of Cyber-Physical Systems. Scientific Foresight Study*.
- Organización Mundial de la Salud (OMS) (2019). *WHO Guideline: recommendations on digital interventions for health system strengthening*.
- Parlamento Europeo (2017). *Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))*.
- Rigby, M. J. (2019). "Ethical Dimensions of Using Artificial Intelligence in Health Care", en Rigby, M. J. (ed). *Artificial Intelligence in Health Care*. American Medical Association Journal of Ethics.
- RIKEN (2011). *RIBA II, the next generation care-giving robot*. http://www.riken.jp/en/pr/press/2011/20110802_2/ (última consulta 30.04.2019).
- Sánchez García, A. M. (2018). "Robótica y ética", en Barrio Andrés, M. (Dir.). *Derecho de los robots*. Madrid, Wolters Kluwer, pp. 229–256.
- Stanford University (2016). "Artificial Intelligence and Life in 2030". https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0831fnl.pdf (última consulta 30.04.2019).
- Veruggio, G y Operto, F. (2016). "Roboethics: a bottom-up interdisciplinary discourse in the field of applied ethics in robotics". *International Review of Information Ethics*. Vol. 6, pp. 3–8.
- Walker, J. (2017) "Researchers shut down AI that invented its own language". Digital Journal. <http://www.digitaljournal.com/tech-and-science/technology/a-step-closer-to-skynet-ai-invents-a-language-humans-can-t-read/article/498142> (última consulta 30.04.2019).

Notas

1. En los años 2007 y 2009, los profesores Floreano, Keller y Mitri, del Laboratorio de Sistemas Inteligentes de la Escuela Politécnica Federal de Lausanne y el Departamento de Ecología y evolución de la Universidad de Lausanne, llevan a cabo un experimento con 10 inteligencias artificiales. Estas consiguen puntos cuanto mayor sea el tiempo que permanezcan cerca de una zona denominada "alimento", existiendo también una zona de penalización, llamada "veneno". Las inteligencias podían comunicarse entre sí. Al cabo de 50 generaciones, los robots empiezan a informarse unos a otros sobre donde encontrar el "alimento"; mas no solo eso, sino que, para sorpresa de los investigadores, algunos adquieren el rol de héroes, informando donde encontrar el "alimento" para después morir al impedir que otros se acerquen al "veneno". Otros incluso comunican a sus pares que la zona de "veneno" corresponde con la de "alimento", mintiendo para sobrevivir. Cfr. (Mitri, Floreano y Keller, 2009).
2. Durante el mes de julio del año 2017, Facebook se ve obligado a interrumpir forzosamente la comunicación entre dos sistemas electrónicos. Programados para interactuar en lengua inglesa, habían creado un sistema lingüístico propio, más eficiente para ellos que la lengua de programación, comunicándose de un modo que los diseñadores no podían entender ni descifrar. Vid. (Walker, 2017).
3. Sofía®, una inteligencia artificial con forma humana creada por Ben Goertzel para Hanson Robotics, fue en 2017 la primera androide a la que se reconoció ciudadanía de un Estado; en este caso, de Arabia Saudí. Cfr. (Griffin, 2017).
4. La iniciativa "Malta.ai – Towards a National AI Strategy", del Gobierno maltés, plantea decididamente el otorgamiento de ciudadanía a las inteligencias artificiales avanzadas, barajando la expedición de pasaportes. Para mayor información: <https://malta.ai/wp-content/uploads/2018/10/Malta-AI-Vision.pdf>
5. El 16 de febrero de 2017, el Parlamento Europeo emite una Resolución "con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))", cuya sección 59.f explicita la recomendación a la Comisión:

“crear a largo plazo una personalidad jurídica específica para los robots, de forma que como mínimo los robots autónomos más complejos puedan ser considerados personas electrónicas responsables de reparar los daños que puedan causar, y posiblemente aplicar la personalidad electrónica a aquellos supuestos en los que los robots tomen decisiones autónomas inteligentes o interactúen con terceros de forma independiente”.

6. Vid. (Oficina de Prospectiva Científica (STOA), 2016).
7. Vid. (Comisión Europea, 2019a). El Informe aludido determina la inteligencia en un examen de racionalidad, refiriéndose este a la capacidad de elegir el mejor curso de acción para conseguir un fin, dados ciertos criterios que deben ser optimizados (y que son preconfigurados *ex ante*), y la disponibilidad de recursos.
8. <https://edition.cnn.com/2019/03/10/health/patient-dies-robot-doctor/index.html>
9. (Mackenzie, 2010)
10. Cfr. (Agencias Aranda de Duero, 2018).
11. RIKEN (2011).
12. En el momento de redacción del presente trabajo, la norma se encuentra en proceso de revisión. Cfr. International Organization of Standardization (2014).
13. Los de servicio móvil serían aquellos capaces de desplazarse autónomamente, transportando objetos o intercambiando información. Por su parte, los robots asistentes serían los que colaboran con un usuario, complementando o aumentando sus capacidades. Por último, los *carebots* transportadores son capaces de trasladar a una persona a un destino prefijado.
14. Vid. García Portero (2018).
15. Los principios se agrupan en tres categorías diferentes: cuestiones relacionadas con la investigación, ética y valores, y escenarios a largo plazo. Pueden consultarse en el siguiente enlace: <https://futureoflife.org/ai-principles/?cn-reloaded=1>
16. Puede consultarse en el siguiente enlace: https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0831fnl.pdf
17. En el momento de redacción del presente trabajo, la última versión publicada es la de 15 de octubre de 2018.
18. Cfr. Art. 113.2: “El médico no debe permitir que ningún sistema mecánico ni robótico tome una decisión médica en el proceso asistencial a un paciente”
19. En la introducción, afirma el Doctor Tedros Adhanom Ghebreyesus (Director General de la OMS): “*Countries must be guided by evidence to establish sustainable harmonized digital systems, not seduced by every new gadget*”
20. Para esta versión de la guía, que prevee ampliarse en el futuro, las aplicaciones consideradas son: notificación de nacimientos a través de dispositivos móviles, notificaciones de fallecimiento a través de dispositivos móviles, operaciones logísticas por vía electrónica, telemedicina en la relación “cliente-proveedor”, telemedicina en las relaciones entre proveedores, comunicación con segmentos específicos de pacientes, seguimiento de pacientes y servicios por vía digital, toma de decisiones electrónica por parte del personal de salud y formación de profesionales por vía electrónica.
21. El propio texto de la Unión propone una *Assesment Tool*. Cfr p.26 *supra*
22. Citando a Shay, señala Gilligan que “(La) visión de una vida buena para un ser humano es una elección ética y no se puede coaccionar. Sólo puede surgir del diálogo, la educación y la atracción intrínseca”. Consideramos trasladable su reflexión a una responsabilidad del uso de la inteligencia artificial, que debe identificarse con un carácter receptivo, y que tenga en consideración la identificación de características emocionales en los pacientes.
Cfr. Gilligan, C. (2013).
23. Hastings Center (Bioethics Research Institute) (1996). *The goals of Medicine. Setting new priorities*. El texto se ha traducido al castellano, por la Fundació Víctor Grifols i Lucas. Puede encontrarse en el siguiente enlace: http://www.hospitalsantjoan.cat/wp-content/uploads/2018/09/fins_medicina.pdf

- 24.** Sobre la idea de adaptabilidad de las inteligencias artificiales, cfr. Domínguez-Alcón (2017).
La autora plantea un análisis de la robótica en consideración de la atención, la responsabilidad, la competencia y la capacidad de respuesta, como marco de referencia de la ética del cuidado.
- 25.** Apunta Rigby lo fluctuante y delicado del equilibrio de dos fuerzas. Riesgo, por un lado, por los retos éticos que coloca el uso de tecnologías, y beneficios, por otro, detectados en el incremento de eficiencia de los sistemas de salud.